

Relation between syllable count judgments and durations of English liquid rimes

Abigail C. Cohn and Sam Tilsen

Department of Linguistics, Cornell University
acc4@cornell.edu, tilsen@cornell.edu

ABSTRACT

Judgments of the number of syllables in a given word are usually consistent, but there is a class of English words with liquid codas and tense vowel/diphthong nuclei for which speakers have variable syllable count intuitions (e.g. *fire*, *feel*). This variation has previously been hypothesized to be associated with differences in subsyllabic structural organization, which predicts that speaker judgments of syllable count should correlate with produced rime durations. To test this hypothesis, 34 native speakers of English participated in sequential and parallel syllable count judgment and word production tasks. Durations of diphthong/tense vowel-liquid rimes associated with disyllabic judgments were found to be significantly longer than those associated with monosyllabic judgments. In addition, substantial variation in syllable-count judgments between and within speakers was observed. This relationship between syllable-count judgment and rime duration has important implications for our understanding of the relation between phonological representations and production.

Keywords: Laboratory phonology, syllables, liquids, duration, speech production

1. INTRODUCTION

For most monomorphemic words in English, native speakers have robust, consistent intuitions regarding the number of syllables that comprise the word. However, there is a small class of words, consisting of a diphthong or high/mid tense vowel nucleus and liquid coda (e.g. *pile*, *pail*, *pool*, *fire*, *fail*, *fool*) for which speakers do not exhibit consistent syllable-count judgments (henceforth " σ -count judgments"). The same variation is not observed with low or lax vowel nuclei, nor with non-liquid sonorant codas. This raises the question of why σ -count judgments are variable only for words with the aforementioned rimes. We will subsequently refer to these as "variable-count words", because of the inconsistency in speakers' σ -count judgments. While some speakers judge the variable words as comprised of one syllable ($=1\sigma$), others judge them as comprised

of two syllables, or as more than one but not quite two syllables ($>1\sigma$).

Previous research [5] has suggested a relation between the phenomenon of variable σ -count intuitions and sub-syllabic structural representation. Specifically, variable-count words can be analysed as having a trimoraic syllable structure. (Although the term "mora" is used here, the hypotheses do not require a commitment to moraic theory per se, rather, only to a notion of structure that organizes segments within syllables.)

The current study tests the hypothesis that variation in σ -count judgments is *structurally conditioned*, i.e. attributable to differences in sub-syllabic structural representations. This hypothesis predicts that phonetic aspects of productions, such as rime duration, should be related to σ -count judgments. Specifically, the hypothesis predicts that variable-count words judged as $>1\sigma$ will be produced with longer rime durations than words judged as $=1\sigma$. Part of the rationale for this prediction comes from previous observations of phonetic differences relating to sub-syllabic (or moraic) structure [1, 3, 4], as reviewed in [2]. The null hypothesis is that variation in σ -count judgments is driven entirely by other, non-structural factors (e.g. metalinguistic processes that do not influence production), and hence is not predicted to correlate with phonetic aspects of production.

To evaluate the structural conditioning hypothesis, an experiment was conducted involving sequential and parallel σ -count judgment and word production tasks, as detailed in Section 2. The results of the experiment provide support for the hypothesis that variation in σ -count judgments derives from variation in structural representations: rime durations of variable-count words produced by speakers with $>1\sigma$ judgments were indeed found to be longer than those produced by speakers with $=1\sigma$ judgments.

2. METHOD

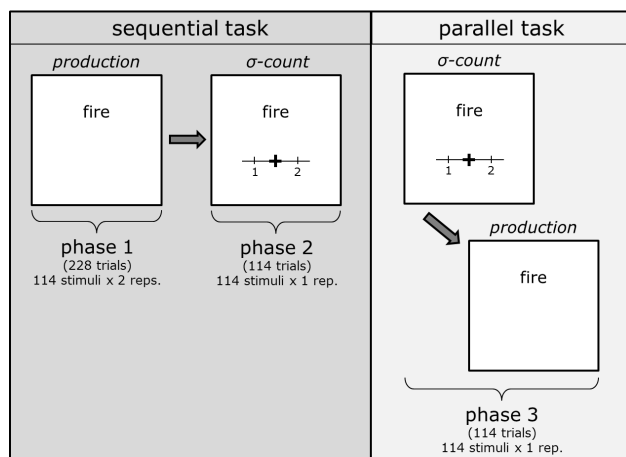
2.1. Participants and task

Thirty-four native speakers of English with no known speech or hearing problems participated in the experiment (18 male, 16 female; ages 18-29,

median 20). Seventeen of the participants had resided in the Eastern U.S. the majority of their life, 12 in the Midwest or Western U.S., and 5 outside the U.S. During the experiment, participants were seated in a sound-proof booth in front of a computer monitor and wore a head-mounted microphone.

The experimental session was organized into three phases as schematized in Fig. 1. Before each phase, participants read instructions on the computer monitor. Participants were instructed to produce each word that appears on the screen in the phrase *I say ___ sometimes*, not emphasizing any word nor hesitating within the phrase. The entire stimulus set was produced twice in the first phase (114 stimuli x 2 reps = 228 trials), judged once in the second, and judged/produced once in the third phase. Within each phase, stimuli were presented in pseudo-randomized order, with target words never occurring on consecutive trials.

Figure 1: Sessions were organized into 3 phases: production, σ -count judgment, and production with σ -count judgment. Phases 1 and 2 comprise the sequential task, phase 3 the parallel task.



Because previous studies suggested that variable intuitions may be associated with the impression that a word contains more than 1 syllable but not quite 2 canonical syllables, σ -count judgments were elicited with mouse clicks on a continuous horizontal scale. The scale ranged from 0.5 to 2.5 and the integers 1 and 2 were labelled with tick marks (see Fig. 1). At the start of each trial the pointer appeared at a value of 1.5. For the σ -count judgment tasks, participants were given instructions that read as follows: "In this part of the experiment, you will decide whether there are one or two syllables in a word. Note that in some cases there is no right answer: people disagree on how many syllables are in some words. In addition, sometimes people feel that the number of syllables in a word is between whole numbers." Furthermore, participants were instructed that when a word

appeared on the screen, they should silently read the word before responding. Each word was displayed on the screen for 1.5 seconds before disappearing, at which point the scale appeared. Participants were also explicitly instructed not to rely on how words are spelled, and told that they should rely on what they perceive when they imagine saying the word.

In the third phase, participants performed the production and σ -count judgment in tandem once for each stimulus. On each trial, they first made a σ -count judgment for a word, and then produced that same word in the carrier phrase. After completing all three phases, participants filled out a survey on their language background, geographic residence history, linguistic educational background, and familiarity with less common (low-frequency) target items in the experiment.

2.2. Stimuli

A list of target stimuli was constructed for elicitation of σ -count judgments and word production. Target stimuli were phonotactically licit combinations of the vowels $\{/i, i, a, ai/\}$ and codas $\{\emptyset, /d, n, l, r/\}$ (\emptyset = no coda), as shown in Table 1. Lax vowel $/i/$ does not occur in open syllables and the tense/lax contrast in high front vowels is merged before $/r/$. To facilitate automated analyses, all words had a singleton labial consonant onset (i.e. $/p/, /b/, /f/, /v/$), or in the absence of viable candidates meeting this criterion, a singleton alveolar stop onset, either $/t/$ or $/d/$. No stimuli were morphologically complex. Thirteen of the 50 target stimuli were expected to be variable-count forms (Table 1, shaded cells). All of the variable-count stimuli were required to be 4 graphemes long.

Table 1: Target stimuli for each coda and nucleus.

nucleus	coda				
	\emptyset	d	n	l	r
i		bid vid	pin, bin fin	pill, bill fill	beer fear pier
i	bee, fee, pea	bead feed	bean teen	peel, feel veal	
a	pa, bah, fa	pod bod	bon Von	pall, ball fall, doll	par, bar, far
ai	pie, buy vie	bide tide	pine, fine vine	pile, bile vile, file	pyre, fire tire

Where possible, words with CELEX log-frequencies in the 25-75% percentile range were used for both target and non-target items. To avoid creating an experiment-wide σ -count response bias, non-target items (n=64) were selected so as to balance the total number of unequivocally monosyllabic and disyllabic stimuli and minimize the correlation between grapheme length and syllable count across the stimulus set.

2.3. Data processing

Responses in the σ -counting tasks were highly multimodal, with most participants exhibiting either bimodal distributions with modes near 1 and 2, or trimodal distributions with modes near 1, 1.5, and 2. For subsequent analyses, response values were mapped to a binomial variable with values of $=1\sigma$ and $>1\sigma$. Data from 6 of the 34 participants were excluded because these participants produced a high proportion of non-standard σ -count judgments for unequivocally mono- and disyllabic words. These participants were likely either not attending closely to the experimental task or were overly reliant on grapheme length in their σ -count judgments.

3. RESULTS

For all variable-count targets, mean rime durations were significantly longer in words judged as $>1\sigma$ than in words judged as $=1\sigma$, in both the sequential and parallel tasks. This finding supports the main hypothesis that variation in production is structurally conditioned. In addition, substantial interspeaker variation was observed in σ -count judgments of diphthong-liquid rimes (/ail/, /air/), and to a lesser extent in high-front/tense vowel-liquid rimes (/il/, /ir/). Word-specific and task-related variation were also observed for both diphthong and high-front/tense vowel rimes in σ -count judgments, although the task effects showed no systematic pattern. These results are described in more detail below.

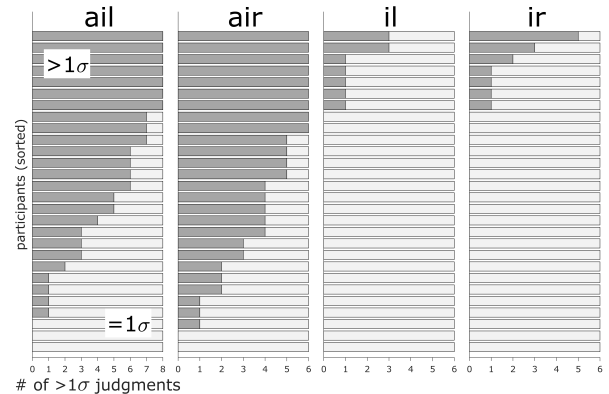
3.1 Variation in σ -count judgments

Analysis of σ -count judgments for each participant revealed a striking pattern of interspeaker variation, particularly for the diphthong-liquid rimes, as shown in Fig. 2. Ten of 28 participants judged all or most (all but one) of the /ail/ rimes as $>1\sigma$, while 7/28 participants judged all or most of these rimes as $=1\sigma$. The remaining 11 participants exhibited intermediate proportions of $>1\sigma$ judgments, reflecting within-participant, word- and/or task-specific variation. A similar pattern was observed for /air/ rimes.

Word-specific variation in σ -count judgments was also observed, with some effect of token frequency. For diphthong rimes, less frequent words such as *bile*, *vile*, and *pyre* were associated with a greater number of $>1\sigma$ judgments than their more frequent counterparts *file*, *pile*, *fire*, and *tire*. Similar frequency-judgment relations were observed with high-front/tense monophthong rimes. A stepwise linear regression of the experiment-wide proportions of $>1\sigma$ judgments by word with rime and log-frequency as predictors showed that lower frequency

words were associated with a higher proportion of $>1\sigma$ judgments ($F = 7.9$, $p = 0.02$).

Figure 2: Counts of $>1\sigma$ judgments for variable-count rimes by participant, sorted for each rime by within-participant proportion.

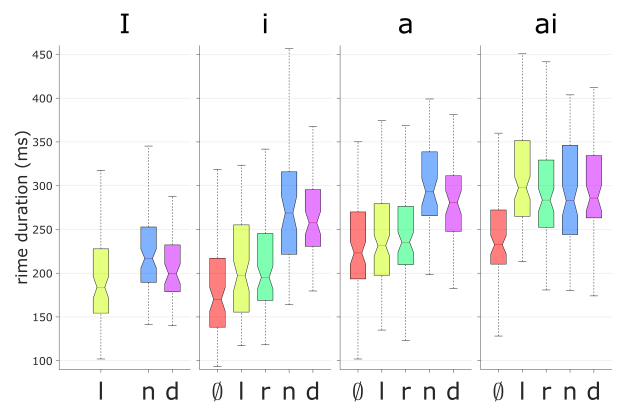


No general trends were evident in the direction of changes of σ -count judgments between tasks. While the majority of σ -count judgments were consistent across tasks, some task-specific variation was also observed. See [6] for discussion.

3.2 Nucleus and coda effects on rime durations

Rime (VC) durations in target word productions were strongly influenced by word identity and rime composition (cf. Fig. 3). Effects of VOWEL, CODA, WORD, and a VOWEL-CODA interaction were all significant in an ANOVA of rime duration (VOWEL: $F(3, 4009) = 603.1$, $p < 0.001$; CODA: $F(4, 4009) = 200.0$, $p < 0.001$; WORD: $F(42, 3967) = 15.5$, $p < 0.01$; VOWEL-CODA: $F(8, 3359) = 43.3$, $p < 0.001$).

Figure 3: Rime durations by vowel and coda. Error bars: 5-95% percentiles, boxes: 25-75% percentiles; notches: ± 2.0 standard error.



For the diphthong nucleus /ai/, the liquid codas, /n/, and /d/ all contribute a substantial amount of duration, resulting in rime durations that are significantly longer than those in an open syllable

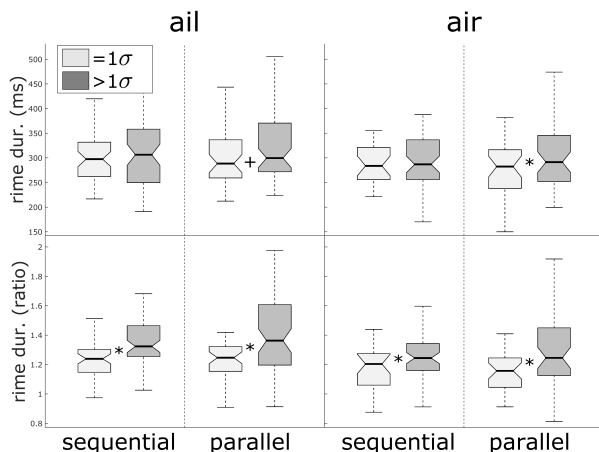
(cf. Fig. 3); however, in rimes with a low vowel /a/ nucleus, the liquids do not contribute a substantial amount of duration to the rime: only the /n/ and /d/ codas result in significantly greater rime duration compared to the open syllable /a/. A similar effect is observed with the high-front/tense vowel.

3.3 Rime durations and σ -count judgments

Crucially, words associated with $>1\sigma$ judgments were produced with greater mean rime durations than words associated with $=1\sigma$ judgments. This finding supports the structural conditioning hypothesis.

Two measures of rime duration were analysed: absolute rime duration and normalized rime duration, expressed as the ratio of absolute rime duration to the within-speaker average rime duration of a nucleus-matched open syllable. Two-sample one-sided t-tests for each speaker and rime category showed that for all rimes/task the ratio measure was significantly higher for tokens associated with $>1\sigma$ judgments than for those associated with $=1\sigma$ judgments. The raw measures were also significantly different for /i/ and /r/ rimes in both tasks, but for the diphthong rimes in the parallel task only. Fig. 4 represents distributions of the measures for the diphthong nuclei.

Figure 4: Effects of σ -count judgment on rime duration for raw and normalized rime durations. * $p < 0.05$, + $p < 0.10$.



4. DISCUSSION

Syllable-count judgments showed substantial interspeaker, word-specific, and task-specific variation for diphthong-liquid rimes (/ail/, /air/); more limited variation was observed in high-front/tense vowel-liquid rimes (/il/, /ir/). Crucially, normalized rime durations were significantly longer in words judged as $>1\sigma$ than in words judged as $=1\sigma$

in both sequential and parallel tasks. Hence the results support the hypothesis that variation in σ -count judgments is structurally conditioned.

Previous investigations have observed interspeaker variation in σ -count judgments for variable-count words [5]. The current study, using a much larger sample of participants, replicates this interspeaker variation, while also revealing a more a complex picture that includes word- and token-specific variation. Although the current study examined only a subset of the variable-count rimes (those with the diphthong /ai/ and high-front/tense vowel /i/), these findings are expected to extend to other variable-count rimes.

The relation between σ -count judgment and rime duration indicates that differences in sub-syllabic representations—not merely variation in the intuition formation process—are responsible for variation in σ -count judgments. These results thus indicate that the process of forming an intuition involves engaging a motor representation in some fashion. A better understanding of the factors influencing σ -count intuitions and production of variable-count word forms should be sought through future studies; understanding these factors will shed light on the nature of phonological representations. The current findings ultimately highlight the importance of studies that address the relation between cognitive processes and representations in behavioural tasks involving explicit judgments of phonological structure.

7. REFERENCES

- [1] Broselow, E., Chen, S.-I., & Huffman, M. 1997. Syllable weight: convergence of phonology and phonetics. *Phonology*, 14(1), 47–82.
- [2] Cohn, A. C. 2003. Phonological Structure and Phonetic Duration: The Role of the Mora. *Working Papers of the Cornell Phonetics Laboratory*, 15, 69–100.
- [3] Duanmu, S. 1994. Syllabic weight and syllabic duration: A correlation between phonology and phonetics. *Phonology*, 11(1), 1–24.
- [4] Ham, W. H. 2001. *Phonetic and phonological aspects of geminate timing*. Routledge.
- [5] Lavoie, L., & Cohn, A. C. 1999. Sesquisyllables of English: the structure of vowel-liquid syllables. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 109–112).
- [6] Tilsen, S., Cohn, A. C. & Ricciardi, E. 2014. Syllable count judgments and durations of liquid rimes in English. *Cornell Working Papers in Phonetics and Phonology 2014*, 1-23.