


RESEARCH ARTICLE

Speech cycling experiment on Finnish rhythm

Joonas Vakkilainen 

Faculty of Information Technology and Communication, Tampere University, Finland
Email: joonas.vakkilainen@tuni.fi

(Received 23 June 2022; revised 27 January 2024; accepted 26 June 2024)

Abstract

Rhythm typology seeks to classify languages according to the units on which they base their rhythm. The most used categories are stress-timing, syllable-timing, and mora-timing. Traditional rhythm typology has been questioned by instrumental research, because the claimed isochronous units have not been found. Speech cycling is a method that endeavours to find regular rhythmic units by utilising repeated speech that the speaker produces by accommodating to regularly paced stimuli. This is intended to eliminate the irregularities that would otherwise hide the linguistic rhythm. The present study is a speech cycling experiment on Finnish. The subjects read sentences that were varied in the mora and syllable count. The locations of the stressed syllables were analysed on a relative scale to see if they appear at regular phases. The results show that in Finnish the mora, the syllable, and the quantity pattern all affect the phases.

Keywords: rhythm; Finnish; speech cycling; timing; mora

1. Introduction

Rhythm is typical to human actions. Walking is an example of rhythm in our everyday life. Walking exhibits self-entrainment, in which one oscillator entrains another so that their movements are kept in the same relation to each other regardless of the speed of the movement (Port, Tajima & Cummins 1999).

Rhythm is generally considered a temporal phenomenon based on repetition. Rhythm is seen as intervals of motion, which means that rhythm is quantitative rather than qualitative (Allen 1973:96–97). On the other hand, rhythm units do not need to be based on strict repetition. Rhythm can also be understood as a chain of patterns, in which the rhythmical impression is created by patterns within patterns that differ from each other in accent, which forms a hierarchy, in which the units are perceived as existing on different levels (Sademiemi 1949:10).

Language is also considered to have rhythm. Rhythm typology has traditionally classified languages according to their rhythm types into stress-timed, syllable-timed, and mora-timed languages. Rhythm typology has been criticised for not finding measurable correlates corresponding to the phonetic reality of language

(Dauer 1983, Pamies Bertrán 1999). A method called speech cycling was developed in studies that utilise repetitive speech adjusted to a metronome to seek the regularity of rhythmical elements in speech. Speech cycling creates a cyclical stimulus that forces speech into regular cycles. This is intended to find the oscillators that in the case of speech are cognitive (Port, Cummins & Gasser 1996, Cummins & Port 1998, Tajima 1998, Port et al. 1999).

This paper presents a speech cycling experiment in Finnish. The research questions are as follows.

1. What rhythmic features does Finnish exhibit?
2. What is the role of the mora in Finnish in relation to the syllable?
3. Is a speech cycling task manageable in Finnish at all and what kinds of unit does it indicate?

In Section 2, the previous literature of linguistic rhythm and speech cycling is presented. In Section 3, the method is explained, in Section 4 the results of the experiment are presented, and in Section 5 the conclusions of the results are presented.

2. Background

2.1 Rhythm typology

Rhythm typology seeks to assign languages to certain rhythm types. The traditional view uses two types, stress-timed and syllable-timed. Rhythm typology aims to find the phonological units that are isochronous in speech. In this view, in a stress-timed language stressed syllables are placed at regular intervals, which means that unstressed syllables are compressed or expanded depending on their number in the interval, while in a syllable-timed language syllables are kept equally long without stresses affecting them (Abercrombie 1965:17–18, 1967:97). A third type used in rhythm typology is mora-timed, in which morae instead of syllables are seen as having regular duration and which has most often been used to describe Japanese rhythm (Bloch 1950, Warner & Arai 2001a).

Rhythm typology has been questioned by various studies. Dauer (1983) did not find significant differences in the duration of stress intervals between stress-timed and syllable-timed languages. Pointon (1980) did not find any kind of isochrony either in stress intervals or syllable durations in Spanish. Wenk & Wioland (1982) denounce syllable isochrony in French, and suggest that French stress intervals are actually more similar to English, with the difference that in French stresses are placed at the end of the foot (*trailer-timing*), contrary to English which places them at the beginning of the foot (*leader-timing*). They also criticise rhythm typology for being too Germanic-centred. Beckman (1982) found no regular durations in Japanese morae, and Roach (1982) did not find stress or syllable isochrony either, but states that languages may have different factors that give rise to the rhythmic impression, which means that linguistic rhythm is predominantly perceptual. Pamies Bertrán (1999) even claims that linguistic rhythm is not a real phenomenon

but a metaphor based on classical poetry, and thus does not describe the phonology of languages spoken in reality.

On the other hand, there are also studies that do not completely denounce phonetic rhythmic differences. Hoequist (1983a,b) found that English, Spanish, and Japanese, a putative stress-timed, syllable-timed, and mora-timed language, respectively, have similar kinds of effects but to different magnitudes, e.g. lengthening when stressed, and effects that set one language apart, e.g. following syllables shortening the preceding syllables. Regarding mora-timing, in Japanese adding a mora affected syllable length the most, as expected, whereas the other effects were smaller than in English and Spanish. Hoequist (1983b:229) concludes that instead of the traditional rhythm descriptions Japanese could be classified as a duration-controlling language, English a duration-compensating language, and Spanish as neither, though Japanese best fits its putative rhythm type.

Out of the three most commonly used rhythm classifications, mora-timing has got the most support. Port, Al-Ani & Maeda (1980) found that in Japanese, morae tend to keep a regular duration by compensating the inherent durations of the sounds both in a mora and over mora boundaries, which they did not observe in Arabic and English. Port, Dalby & O'Dell (1987) found that the word and phrase duration in Japanese depends on the number of morae, not of syllables, although Warner & Arai (2001b) found that neat mora compensation does not happen in spontaneous speech, but the mora count still predicts the word duration better than other segments. Port et al. (1980, 1987) do not consider mora isochrony to mean that each individual mora has a similar duration (cf. Bloch 1950, Beckman 1982) but that the morae compensate each other over larger strings of morae.

Dauer (1987) concludes that mere instrumental phonetic research fails to bring any evidence for rhythm types, and mere phonological grouping of units looks identical in every language. Therefore Dauer suggests that rhythm is based on certain rhythmic elements (duration, f_0 , vowel quality, and the function of stress) whose relationship to stresses determines the rhythm type. Based on these features, languages form a continuum from a more stress-timed language to a less stress-timed language rather than a dichotomy. If stresses are related to syllable complexity, intonation, and duration, the rhythm of the language is based on stresses, and if not, the rhythm units lie elsewhere.

Rhythm is also sought in syllable variability by comparing vocalic and consonantal sequences. Ramus, Nespors & Mehler (1999) found that the vocalic percentage and the standard deviation of consonantal intervals indicate the rhythm type, because the more different syllable types a language has, the more consonantal variation it has and its vocalic percentage decreases, which leads to stress-timing. They did not find a clear connection between the standard deviation of vocalic intervals, but it may explain why some languages fall between the types. Grabe & Low (2002) used the *pairwise variability index* (PVI). They found that the putative syllable-timed languages had low vocalic durational variation, whereas the stress-timed languages showed more variation in vowel duration. Some languages fell between these types, because they had high vocalic variation but low consonantal variation (Catalan), or vice versa (Polish). They suggest that the rhythmical impression may be based on languages' having great variation in either of the intervals, and if both had high variation, rhythm could not be formed.

Nolan & Asu (2009) measured the PVI of feet in addition to syllables and found that languages that differ in syllable structure can still be foot-timed (stress-timed). They found that Estonian and English both have isochronous feet, although English has lower vocalic percentage due to vowel reduction and more complex syllables, and Estonian has phonemic quantity. This means that languages can base their rhythm on more than one type of unit, and that syllable-timing and stress-timing can coincide.

Rathcke & Smith (2015) note that the differences in timing can stem from different phonetic realisations of phonological features. For example, reduction and quantity contrasts can be manifested as differences in duration or quality, which give different results in timing and therefore may not be good indicators of phonological rhythm. Katz-Brown & Mandal (2013) note that PVI correlates with a language's phonotactics and thus tells us little about timing. Arvaniti (2009) criticises rhythm typology for focusing almost solely on duration, and says that instead of observing timing it would be better to concentrate on grouping and patterns of prominence. Different speaking styles and even different kinds of sentences can give different rhythmic results (Krull & Engstrand 2003, Arvaniti 2009).

2.2 *Speech cycling*

Speech cycling is intended to respond to the difficulties of previous studies on linguistic rhythm. In a speech cycling task, the speaker repeats a given phrase accommodating to an external sound stimulus. The instructions can vary, and their purpose is to extract certain rhythm units. Tajima (1998) used increasing stimulus intervals in one of his speech cycling series. This was intended to reveal if the speaker has a stable rhythmical structure that is not interrupted by the tempo. Tajima (1998) and Tajima & Port (2003) also used a task in which the speakers were instructed to produce the phrases in a waltz rhythm, which was supposed to reveal whether prominent syllables tend to be located at a fixed time.

The objective of speech cycling is not to prove or disprove phonological rhythmic categories. A speech cycling task is intended to create rhythmic speech so that temporal units that can constitute linguistic rhythm can be observed. Its focus is on the temporal domain and phonetic segments and durations rather than on phonological rhythm, which has been the focus of many past studies (e.g. Wenk & Wioland 1982, Arvaniti 2009, Rathcke & Smith 2015). On the other hand, because of the temporal domain, the focus is not only on phonetic durations either, which has been the main criticism of many past studies on rhythm (e.g. Roach 1982, Dauer 1983, 1987, Arvaniti 2009).

Port et al. (1996, 1999) criticise traditional rhythm concepts for overlooking physical time in the search for rhythm. Speech cycling aims to find temporal binding of events to certain phases in a cycle (Cummins & Port 1998:147). Speech cycling forces the speaker to adapt to periodic speech production. This eliminates the irregularities of free or read-aloud speech that obscure the hierarchical structure. Producing speech by adapting to a stimulus enables the observation of rhythm units appearing regularly without interference. Speech cycling aims to reveal the cognitive oscillators behind the rhythm (Cummins & Port 1998, Tajima 1998, Port et al. 1999, Cummins 2002). Speech cycling is an artificial speech situation but it is not unnatural: a similar kind of adaptation is used in poetry, joint speech such as protest

rallies, and other synchronised speech (Cummins 2013). If the speaker manages to complete the speech cycling task, their speech has a hierarchical rhythm structure. Speech cycling tasks are not always as manageable in every language. Speakers of Spanish and Italian found it difficult to target the stresses to given stimuli, because stresses do not form regular feet in Spanish and Italian (Cummins 2002). On the other hand, lexical stress is not a necessary requirement to succeed in speech cycling, as speakers of Korean placed the prominent syllables on the stimuli, even though Korean lacks lexical stress (Chung & Arvaniti 2013).

Speech cycling produces evidence for rhythm typology. It both confirms some existing concepts and brings in new details. Tajima (1998) found that English and Japanese behave differently in speech cycling tasks, and these differences can be explained by rhythm typology. For instance, English speakers kept their phrases stable in relation to the cycle despite the increasing rate of repetitions, and ended halfway through the cycle, whereas Japanese speakers ended later and later in the cycle as the repetition rate increased. English stresses and Japanese morae were stable within the phrase. Japanese speakers also had a bigger gap between the repetitions if the phrase had odd-numbered morae than when it had even-numbered morae. The places of English stressed syllables and the duration of the phrase were affected less by increasing unstressed syllables in the phrase than in Japanese. These findings fit into the traditional views of stress and syllable or mora timing.

A common finding in speech cycling tasks on English is that the last stressed syllable is most commonly placed on either the 1/2 point or the 2/3 point of the cycle, according to its syllable count. These regular points are called simple harmonic phases (Port et al. 1999). Port et al. (1996) found that it is difficult to make speakers produce a timing different from these simple harmonic phases, even though the speech cycling stimuli tried to force it. Malisz (2013:24) reports that in her speech cycling task on Polish the last stressed syllable of a short phrase was placed more often on the 1/2 point and a longer phrase on the 3/4 point of the cycle, which shows that Polish behaves similarly to English and Arabic, putative stress-timed languages.

Speech cycling also brings out details that may be overlooked by traditional rhythm typology. For instance, Tajima, Zawaydeh & Kitahara (1999) and Zawaydeh, Tajima & Kitahara (2002) found that English and Arabic both behave in a stress-timed manner but Arabic less so than English, because less reduction and bigger variation in vowel lengths make Arabic stressed syllables deviate more from fixed locations. This means that rhythm types form a continuum. According to Tajima (1998), rhythm types are real, although not based on strict isochrony but on prominent syllables being placed at harmonic phases that are stable in time.

2.3 Finnish rhythm

Finnish has phonemic quantity which is independent of stress. Both consonants and vowels can occur as short and long. There is no consensus as to the rhythm type of Finnish. Lehtonen (1970:34) suggests that the mora best explains the Finnish quantity structure. According to O'Connor (1973:240), a quantity language such as Finnish cannot base its rhythm either on stress groups that alternate the syllable

length or equally long syllables. Karlsson (1983:176) states that Finnish is syllable-timed, because stressed and unstressed syllables alternate regularly, and Finnish does not have vowel reduction. Iivonen (1998:316) calls Finnish an accent language as opposed to an intonation language, but sees it as a subset of syllable-timing. Aoyama (2001:117) found that children could produce [ana] and [an:a] with identical durations but making the quantity difference between them. This could mean that both mora- and syllable-timing might apply to Finnish.

Sadeniemi (1949:55) uses the traditional Finnish concepts of *foot* (Finnish *tahti*), which is a stressed syllable and the following unstressed syllables, and *speech foot* (Finnish *puhetahti*), which is a foot with a primary stress along with any following feet with no primary stress. Lehiste (1990) found that in read-aloud Finnish poetry, there was no tendency for the feet to be stabilised into equal length, because the unstressed long syllables were not affected by the lack of stress. Wiik (1991) suggests foot-timing that is based on groups of two or three syllables. In this model, the syllable nucleus (which in the case of a long vowel means here only the first part of the vowel) is one mora and all the following parts of the syllable constitute the second mora. The following syllable until the end of its nucleus is the latter part of the foot. According to Wiik, the feet are isochronous, which is shown in that the syllable lengths compensate each other so that the second syllable of a foot is relatively long when the first syllable is short (monomoraic), and relatively short when the first syllable is long (bimoraic). O'Dell & Nieminen (2006) found in their experiment that some subjects produced foot-timing, and some did not. In O'Dell, Lennes & Nieminen (2008), rhythm was found to be based on phrase stress in which bimoraic sequences may have an effect.

Because the mora is an abstract concept, which is however systematic within a language, it requires research separately in every language (Ogden 1995). Typically a short open syllable is considered to equal one mora and a syllable with a long vowel or a consonant coda is two morae (Fox 2002:46). Finnish fits the first criterion for morae presented by Trubetzkoy (1969:173): a morpheme boundary can occur within a long monophthong, e.g. *omena* ('apple', nominative singular), *omenaa* ('apple', partitive singular) (the same is true for long (geminate) consonants as well, e.g. *olut* ('beer', nom. sing.), *olutta* ('beer', part. sing.)). Morae can be counted in different ways. It is agreed that a short syllable (i.e. (C)V) counts as a single mora. Karlsson (1983:134) states that in Finnish a (C)VV or (C)VC syllable has two morae, and in syllables with more segments (e.g. (C)VVC or (C)VCC) every following consonant in the same syllable contributes an additional mora. This can be called the multimoraic hypothesis. Alternatively, the bimoraic hypothesis states that any long syllable (with segments following (C)V) counts as a two morae (Wiik 1991). O'Dell et al. (2008) found that the bimoraic hypothesis seems more plausible in Finnish. Speech cycling is a good method to explore this as well. In the present study, one of the research questions is whether the quantity shape CVVCCV has three morae like CVVCV and CVCCV or whether it has greater moraic weight. In the present study, the mora is observed from a phonetic point of view, i.e. how the additional morae affect the timing.

Bimoraic feet are thought to have relevance for Finnish prosody and segment durations. Stressed bimoraic feet are lengthened, and this lengthening also makes room for a rise–fall tonal movement, when the foot is accentuated (emphasised)

Table 1. Number of units contained in the phrases

	<i>tämä</i>	test word					<i>sama</i>
		<i>papa</i>	<i>paapa</i>	<i>pappa</i>	<i>paappa</i>	<i>papapa</i>	
phonemes	4	4	5	5	6	6	4
moraes	2	2	3	3	3 (4)	3	2
syllables	2	2	2	2	2	3	2
feet	1	1	1	1	1	1	1

(Suomi, Toivanen & Ylitalo 2003, Suomi 2005). Suomi & Ylitalo (2004) state that bisyllabic feet are isochronous, as their segments exhibit durational compensation, but this does not occur in trisyllabic feet. Polysyllabic shortening is not found in Finnish according to Suomi (2007). Arnhold (2014:186) states that Finnish has a three-step durational hierarchy, in which CV syllables are the shortest, CVV and CVC are longer than CV syllables, and CVVC are the longest.

The aim of the present study is to conduct a speech cycling experiment on Finnish. It is hypothesised that a change in quantity affects the placing of the phases in the cycle.

3. Method

3.1 The speech cycling task

Five mora/syllable variants were embedded in identical carrier sentences. The varied word was used to test the effect of quantity, i.e. the mora count, separately from the syllable count, as opposed to the previous speech cycling studies which did not vary mora and syllable count independently (Tajima 1998, Tajima et al. 1999, Tajima & Port 2003). The test phrases were nonsensical phrases, whose first word was always *tämä* ('this') and the last word *sama* ('same'). Of the varied words, *pappa* and *paappa* are real words used in Finnish ('grandfather'), whereas *paapa* and *papapa* are nonsensical, and *papa* is recognised as a pan-European word, though not usually used in Finnish. As Finnish has only a handful of monomoraic function words and open-class words always have at least two voiced morae (Suomi 2004:94), the test phrases were constructed according to this. The test word is varied in its quantity shape and syllable count. The number of units in the phrase are shown in Table 1. In the number of phonemes, long sounds count as two individual phonemes, as is the standard in Finnish phonology (Karlsson 1983:71). The same vowel height was used for all words in the phrase so that it would not cause compensation in segmental timing (cf. Lehtonen 1970:86). The phrases used were

1. *Tämä papa sama* ['tæmæ'papa'sama]
2. *Tämä paapa sama* ['tæmæ'pa:pa'sama]
3. *Tämä pappa sama* ['tæmæ'pap:ɑ'sama]
4. *Tämä paappa sama* ['tæmæ'pa:p:ɑ'sama]
5. *Tämä papapa sama* ['tæmæ'papapa'sama]

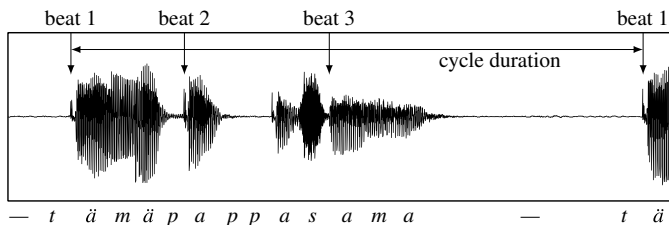


Figure 1. Example showing measurement points of the phases for one cycle.

The speakers were given written instructions in Finnish: ‘Your task is to repeat the phrases given to you. You will simultaneously hear a series of beeps. First, listen to four beeps to find out their speed. Start repeating the sentence at the fifth beep. Repeat the sentence so that its beginning will always coincide with the beep. Repeat the sentence wholly before the next beep. Do the whole series of repetitions in one breath and repeat the sentence as long as you can without breathing or until the beeps end. There are five sentences altogether, and each of them has its own series of beeps. Don’t be confused by the fact that the sentences are not sensible in meaning.’

Ten Finnish-speaking subjects took part in the experiment. The phrases were given on paper one at a time in random order. The stimulus was a 440 Hz beep of 50 ms repeated every 1500 ms. The subjects repeated the same phrase once for every beep as long as the beeps continued, which totalled 11 repetitions for every phrase if the subject managed to do them all. The first and last repetitions were left out of the measurement (as done by Tajima 1998). In this study, the varied word is called the *test word*. The whole interval from the onset of the phrase until the onset of the next repetitive phrase is called the *cycle*, and the phrase within the cycle the *phrase*.

3.2 Measurement

In this study, the measurements are done from the beginning of the vowel, which means here the explosion in the case of stops (Turk, Nakai & Sugahara 2006). This way of measuring is also in line with common ways of counting morae (Auer 1989). Because the focus is on the timing of these vocalic onsets, this also means that the intervals do not reflect phonological segmentation (Tajima 1998:26). The measurement points, which are the onsets of *tämä*, the test word, and the word *sama*, are called beat 1, beat 2, and beat 3 respectively. Beat 1 is therefore also the onset of the phrase and what the speaker targeted to the stimulus sound. The speaker’s accuracy in matching the onset to the beep is not relevant in this study. Because of that, the stimulus was only used to make the speakers produce rhythmically regular speech and its location is not taken into account. These measurement points are shown in Figure 1.

To find out the relative timing, certain intervals are measured in relation to the cycle. The cycle means the interval between beat 1 and the following beat 1. The measured intervals are called the *external phase* and *internal phase*, as used for example by Tajima (1998) and Port et al. (1999). External phase means the position of a beat in relation to the entire cycle. Therefore, the external phase of beat 2 is $tämä/cycle$ and the external phase of beat 3 is $(tämä + test\ word)/cycle$. Internal

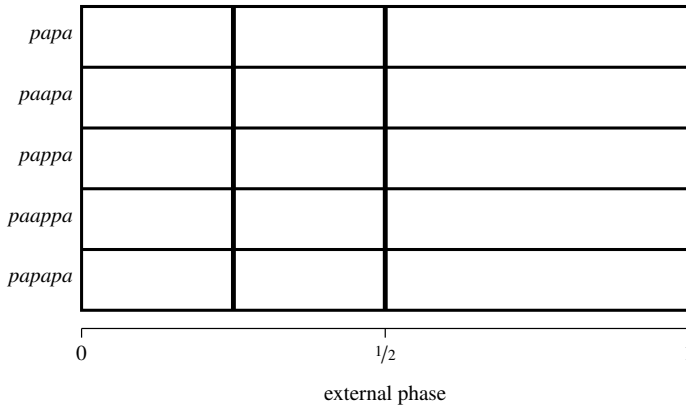


Figure 2. Bar plots for beat 2 and beat 3 in pure stress-timing.

phase means the position of the beat in relation to the interval between beat 1 and beat 3, so the internal phase of beat 2 is $täm\ddot{a}/(täm\ddot{a} + test\ word)$. In addition to the phases, the proportion of the test word is measured. This means the relation of the interval between beat 2 and beat 3 to the cycle and is therefore $testword/cycle$. Because the measurements are taken at the vocalic onset, the proportions and measures of the words in the cycle are in practice $\ddot{a}m\ddot{a}p$, $a(a)p(p)a(pa)s$, and ama .

The calculation of the beats goes as follows. One of the *paapa* phrase repetitions from speaker A is used as an example. The cycle is 1506 ms, of which *tämä* is 381 ms and the test word *paapa* is 408 ms. The external phase of beat 2 is therefore $381 \div 1506 \approx 0.2530$ and the external phase of beat 3 is $(381 + 408) \div 1506 \approx 0.5239$. The internal phase of beat 2 is $381 \div (381 + 408) \approx 0.4829$. The proportion of the test word is $408 \div 1506 \approx 0.2709$. This gives the percentages that are used to denote the placement of the beats within the cycle. The external and internal phases are the main interest in this study, so they will be discussed the most.

3.3 Expectations

Hypothesised results according to different rhythm types are presented here as bar plots in Figures 2, 3, and 4 and also as a scatterplot in Figure 5 according to the example in Tajima (1998). How stable the phases are show how fixed the stressed syllables are and how much the number of syllables or morae affect them. If the phases are always placed nearly on the same points, it can be seen as a sign of stress-timing, and if they are affected by the changing syllable or mora count, it can be seen as a sign of syllable- or mora-timing (Port et al. 1999). The internal phase can be seen in the bar plots as the distance of beat 2 to beat 1 and beat 3. In stress-timing it is expected to be halfway between the other beats, and in syllable- and mora-timing increasing the syllable or mora count in the test word would shift it closer to beat 1 (Tajima et al. 1999, Zawaydeh et al. 2002).

If Finnish is stress-timed, the stressed syllables should be at the same points in every phrase. Based on the earlier speech cycling studies, it can be expected that beats 2 and 3 would be located at the $1/3$ and $1/2$ points on the external phase

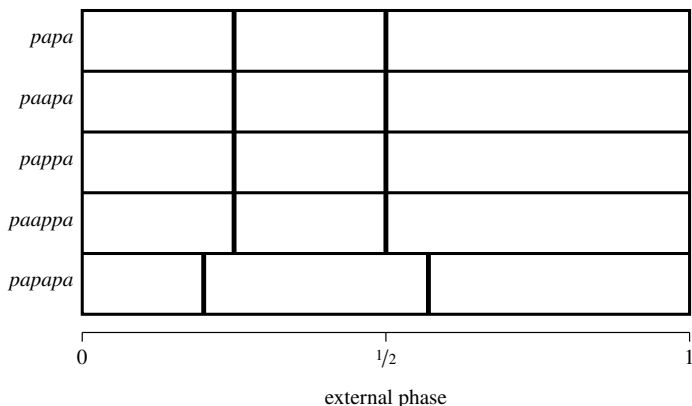


Figure 3. Bar plots for beat 2 and beat 3 in pure syllable-timing.

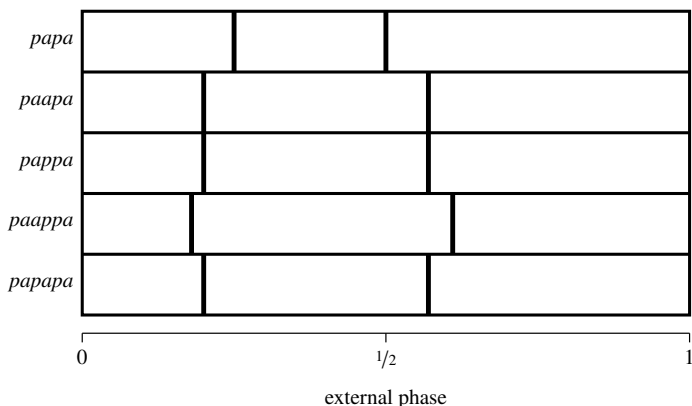


Figure 4. Bar plots for beat 2 and beat 3 in pure mora-timing.

(e.g. Port et al. 1999), as presented in Figure 2. The internal phase of beat 2 would always be located at the 1/2 point and the test word proportion could be expected to be a 1/4 of the cycle. When presented as a scatterplot, it would mean that all phrases would have the proportions A presents in Figure 5.

If Finnish is syllable-timed, phrases 1–4 should have the beats placed at different points compared with phrase 5. In this scenario, test word 5 is 50 % longer than the other test words, and beat 2 is therefore earlier and beat 3 later than in the other test phrases, as is presented in Figure 3. When shown as a scatterplot in Figure 5, syllable-timing would mean that phrases 1–4 would be located at A similarly to the hypothesised stress-timed phrases and the *papāpa* phrase at B.

Mora-timing functions similarly to syllable-timing with the difference that quantity and coda consonants shift the beats as well. If Finnish is mora-timed, phrase 1 should have different timing from phrases 2, 3, 5, which would be equal to each other and again different from phrase 4, at least if the multimoraic hypothesis is true. Possible bar plots for mora-timing are presented in Figure 4. As a scatterplot

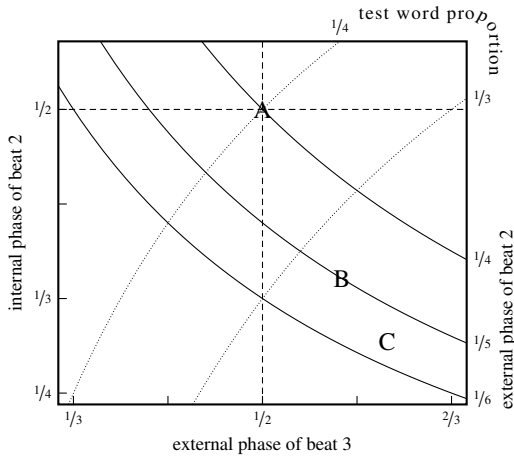


Figure 5. Theoretical Tajima-type scatterplot for different types of timing. Stress-timing predicts a location near A for all test phrases; syllable-timing predicts phrase 5 *papapa* should be located near B, and a location near A for all other test phrases; for mora-timing test phrase 1 *papa* should be located near A, phrase 4 *paappa* near C, and all other phrases near B.

in Figure 5, in mora-timing, the phases of the *papa* phrase would be located at A, phrases 3, 2, and 5 at B, and the *paappa* phrase at C.

It is not expected that the results would show any of these types alone, but that some effects of syllable- and mora-timing combined with possible effects of feet may emerge.

4. Results

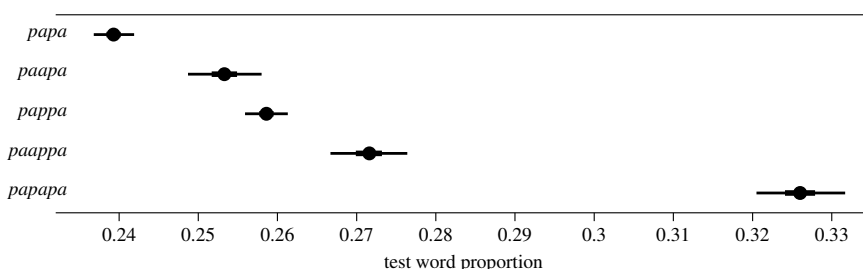
In the results, the average durations and proportions for the test words and beats are presented both averaged over all subjects and for individual subjects. Calculations were performed using an isometric log ratio (ilr) transformation for proportions (a multivariate version of the logit transformation; see e.g. Egozcue & Pawlowsky-Glahn 2019). For presentation purposes all values have been transformed back to percentages. In order to carry out statistical inference for various differences, a Bayesian analysis of the data was performed using a hierarchical model (see Gelman & Hill 2007) with phrase, subject, and phrase \times subject interaction as factors. Posterior distributions and probabilities were computed using R and JAGS (R Core Team 2023, Plummer 2017). The term *posterior* refers to estimated distributions and probabilities updated with the evidence provided by the data (see e.g. Vasisht, Nicenboim, Beckman, Li & Kong 2018).

4.1 Average durations and timing for all speakers

The average durations of the parts of the cycle and the average proportion of the test word are shown in Table 2. The durations of the test words increase as expected when the mora number increases, and when a syllable is added in the fifth test word the duration greatly increases. The proportions of the test words are as expected:

Table 2. Average duration and proportion of the cycle for each test word

test word	duration (ms)	proportion (%)
<i>papa</i>	356.7	23.9
<i>paapa</i>	382.9	25.7
<i>pappa</i>	385.7	25.8
<i>paappa</i>	400.3	26.8
<i>papapa</i>	485.2	32.8

**Figure 6.** Estimated (posterior) average test word proportion (subject effects removed).

papa has the smallest proportion and *papapa* the biggest one. Test word *paappa* has a smaller proportion than *papapa* but a bigger proportion than test words 1–3.

Figure 6 shows the estimated average test word proportion (posterior distribution) for each phrase. The thick line presents the 50 % credible interval (CI), the thin line presents the 95 % CI, and the dot represents the median. As can be seen in the figure, *papapa* clearly has a larger proportion than the other test words. Test word *paappa* is closer to the test words 1–3, but its proportion is clearly larger than their proportions. The proportion of *papa* is also clearly smaller than that of the other test words. Test words *paapa* and *pappa* have almost the same average proportions (Table 2), but the posterior distribution in Figure 6 shows that it is likely that *pappa* has a somewhat larger proportion than *paapa*. In fact, the posterior probability that average test word proportion is greater for *pappa* than for *paapa* is $p = 0.973$; for all other adjacent pairs in Table 2, the posterior probability is $p > 0.999$ that the average test word proportion is larger for the latter phrase.

Average timing, i.e. the alignment of the onsets to the phases of the cycle, is shown as bar plots in Figure 7. The figure shows average external phases (i.e. alignment in relation to the cycle) for beat 2 and beat 3 in black and the means of individual speakers in grey. Figure 8 presents the average timing as scatterplots so that the internal phase of beat 2 is on the y axis, the external phase of beat 3 on the x axis, the external phase of beat 2 on the black curved lines, and the test word proportion on the dotted curved lines.

Figures 9, 10, and 11 show estimated (posterior) distributions of average phases for each phrase. As in Figure 6, the thick line in these figures indicates the 50 % CI, the thin line the 95 % CI, and the dot represents the median. Figures 9 and 10 show the external phases of beat 2 and beat 3, respectively. Figure 11 shows the internal

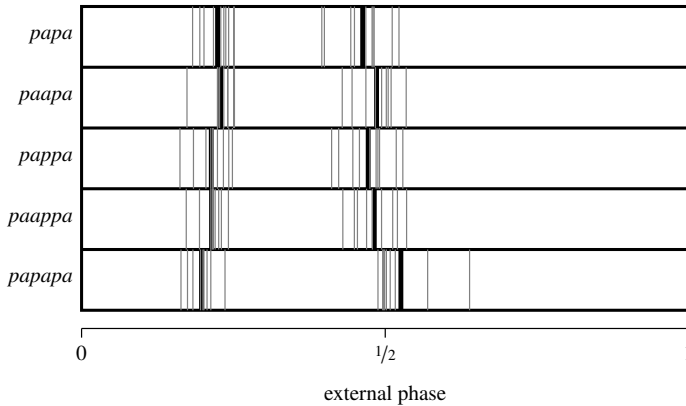


Figure 7. Bar plots for beat 2 and beat 3, averaged over all subjects, shown as thick bars. Individual subject averages in grey.

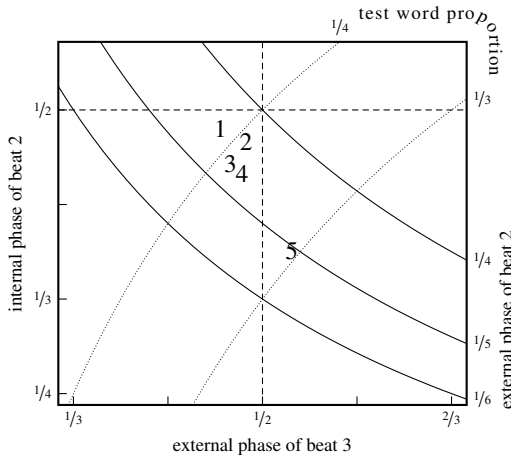


Figure 8. Tajima-type scatterplot, averaged over all subjects (test phrases 1 = *papa*, 2 = *paapa*, 3 = *pappā*, 4 = *paappa*, 5 = *papapa*).

phase of beat 2 and the proportion of the test word within the internal phase (1 – the internal phase of beat 2).

It can be clearly seen that the *papapa* phrase differs from the other phrases. While the other test words have their external phases for beat 2 between the 1/4 and 1/5 points, the posterior distributions in Figure 9 show that the *papapa* phrase has beat 2 placed on average around 1/5, and its posterior distribution is very separate from the other test phrases. The posterior probability is $p > 0.999$ that the beat 2 external phase for *papapa* is earlier than for any other phrase.

The difference is even clearer for beat 3, whose external phase is earlier than 1/2 for phrases 1–4 and later for phrase 5. This means that *papapa* takes time from *tämä* and pushes beat 3 further, as can be seen in Figure 7. The posterior distributions in Figure 10 indicate that beat 3 for the *papapa* phrase is quite different

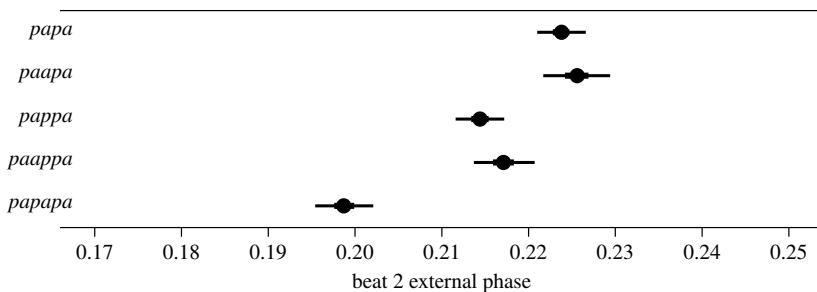


Figure 9. Estimated (posterior) average beat 2 external phase (subject effects removed).

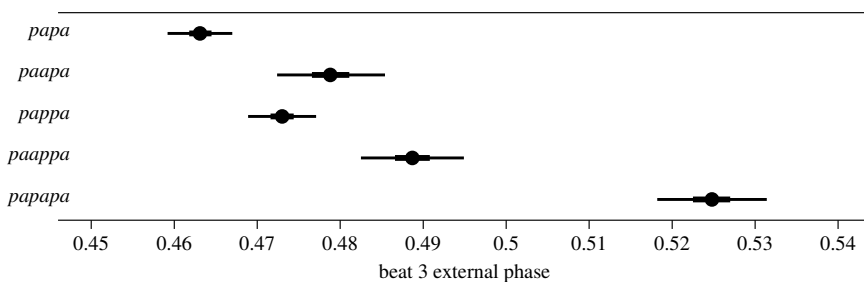


Figure 10. Estimated (posterior) average beat 3 external phase (subject effects removed).

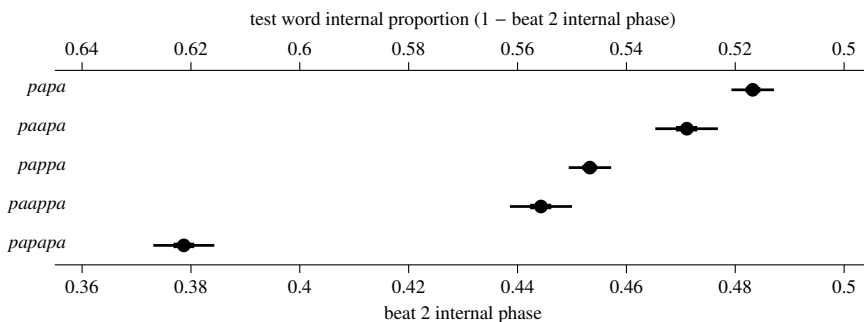


Figure 11. Estimated (posterior) average beat 2 internal phase (subject effects removed).

from the other phrases. The posterior probability is $p > 0.999$ that the beat 3 external phase for *papāpa* is later than for any other phrase. The posterior probability is also $p > 0.999$ that the beat 3 external phase for *papa* is earlier than for any other phrase. The other phrases also differ, although there is more overlap in their posterior distributions. The *paapa* phrase probably has slightly later beat 3 external phase than *pappā* (posterior probability $p = 0.926$) and *paappa* probably has later beat 3 external phase than *paapa* ($p = 0.984$).

Another way to observe the effect of the number of syllables or morae to the beats is the internal phase of beat 2, i.e. the alignment of the medial beat between the phrase onset and the onset of the word following the test word. The average internal phase of beat 2 also separates the different phrases (Figure 11). The posterior

probability that the average beat 2 internal phase is greater for *pappa* than for *paappa* is $p = 0.994$. For all other adjacent pairs (in the order of Table 2), the posterior probability is $p > 0.999$ that the average beat 2 internal phase is earlier for the latter phrase. The internal phase of beat 2 for phrases 1–4 is closer to 1/2, whereas for the *papapa* phrase it is nearer to 1/3 (Figure 8).

The *pappa* and *paappa* phrases are grouped close to each other both for the internal and external phase of beat 2, whereas the *papa* and *paapa* phrases are close for the external phase of beat 2 but more distinct from each other for the internal phase of beat 2, as can be seen in Figures 9 and 11. Interestingly, the beat 2 external phase for *pappa* is earlier than for *paapa* with high probability (posterior probability $p > 0.999$), so that the beat 2 external phase roughly divides the phrases into three groups: $\{papa, paapa\} > \{pappa, paappa\} > \{papapa\}$. This means that the *paapa* and *pappa* phrases do not behave identically, despite their words having the same number of morae and an average proportion with only a 0.1 percentage point difference. Test word *paapa* generally has a later timing than *pappa*, as can be seen in Figure 7.

4.2 Between-speaker differences

The ten speakers show different timing strategies. Their relative timings are shown numerically in Table 3. The external phases for each speaker are shown as bar plots in Figure 12, which shows the average beats in black and the beats in individual repetitions in grey. Figure 13 shows the relative timing for each speaker in a style similar to Tajima (1998) so that the internal phase of beat 2 is on the y axis, the external phase of beat 3 on the x axis, the external phase of beat 2 on the black curved lines, and the test word proportion on dotted curved lines. Phrase 5 (*papapa*) is missing for speaker G, because G had difficulties producing this phrase. Phrase 2 (*paapa*) is missing for speaker H and phrase 4 (*paappa*) is missing for speaker I due to technical failures in recording.

Speakers I and D are the most syllable-timed. Their timings are very similar for the phrases with a two-syllable test word, but their *papapa* phrases are different: for I, *papapa* pushes beat 3 further and keeps the external phase of beat 2 near 1/4, whereas for D *papapa* takes time away from *tämä*, moving the external phase of beat 2 from near 1/4 to earlier than 1/5 (Figures 12 and 13). For both speakers the internal phase of *papapa* is closer to 1/3, whereas for the other phrases it is 1/2 (Figure 13). Speaker C is also very syllable-timed, and he combines the movements caused by *papapa* as shown by speakers I and D, since for him, the test word takes time away from *tämä* and also pushes beat 3 forward. An interesting aspect of C's timing is that his beat 2 in the *paapa* phrase is over 2 percentage points later than in his *papa* and *pappa* phrases, but the timing of beat 3 is not affected, meaning that the proportion of *paapa* is smaller.

Speaker A is the most stressed-timed. Her external phase of beat 3 is stable near 1/2. Her external phase of beat 2 shifts from near 1/4 closer to 1/5 (Figure 12), but the relation between the timing of beat 2 and beat 3 is almost the same for each phrase (Table 3), which suggests that she has just produced the first two phrases in a slower tempo. Only *papapa* causes a little more movement in the phases. However, the changes in the internal phase (Figure 13) do not support stress-timing.

Table 3. Average external phases of beat 2 and beat 3 for each subject (%)

phrase	beat 2	beat 3	beat 2	beat 3
	subject A		subject B	
<i>papa</i>	25.14	51.15	18.28	39.55
<i>paapa</i>	24.08	50.18	17.36	42.92
<i>pappa</i>	21.62	48.79	16.20	42.33
<i>paappa</i>	21.70	47.88	17.24	44.91
<i>parapa</i>	20.15	49.58	16.37	49.90
	subject C		subject D	
<i>papa</i>	20.16	44.32	23.75	47.83
<i>paapa</i>	22.43	44.54	23.42	46.81
<i>pappa</i>	20.48	44.67	22.37	47.46
<i>paappa</i>	21.28	45.42	23.02	46.93
<i>parapa</i>	17.44	50.19	18.33	48.80
	subject E		subject F	
<i>papa</i>	22.92	47.99	24.22	48.17
<i>paapa</i>	22.37	50.49	25.17	50.97
<i>pappa</i>	21.63	49.06	24.24	51.81
<i>paappa</i>	21.99	52.00	24.18	53.52
<i>parapa</i>	19.99	50.82	20.66	51.65
	subject G		subject H	
<i>papa</i>	23.42	46.80	19.47	39.98
<i>paapa</i>	24.07	49.40		
<i>pappa</i>	23.35	48.49	18.42	41.17
<i>paappa</i>	22.58	51.23	19.42	43.02
<i>parapa</i>			19.58	49.74
	subject I		subject J	
<i>papa</i>	25.04	52.26	21.72	44.93
<i>paapa</i>	25.03	53.45	22.72	48.39
<i>pappa</i>	24.84	52.90	21.27	45.75
<i>paappa</i>			21.54	49.39
<i>parapa</i>	23.62	63.89	21.28	56.98

Speaker E can also be seen to have features of stress-timing for similar reasons to A (Figure 12), but he also has more changes due to the morae, which also groups him close to speaker F (Table 3), who is probably the most mora-timed speaker. For speaker F, the *paapa* and *pappa* phrases have later external phases for beat 3

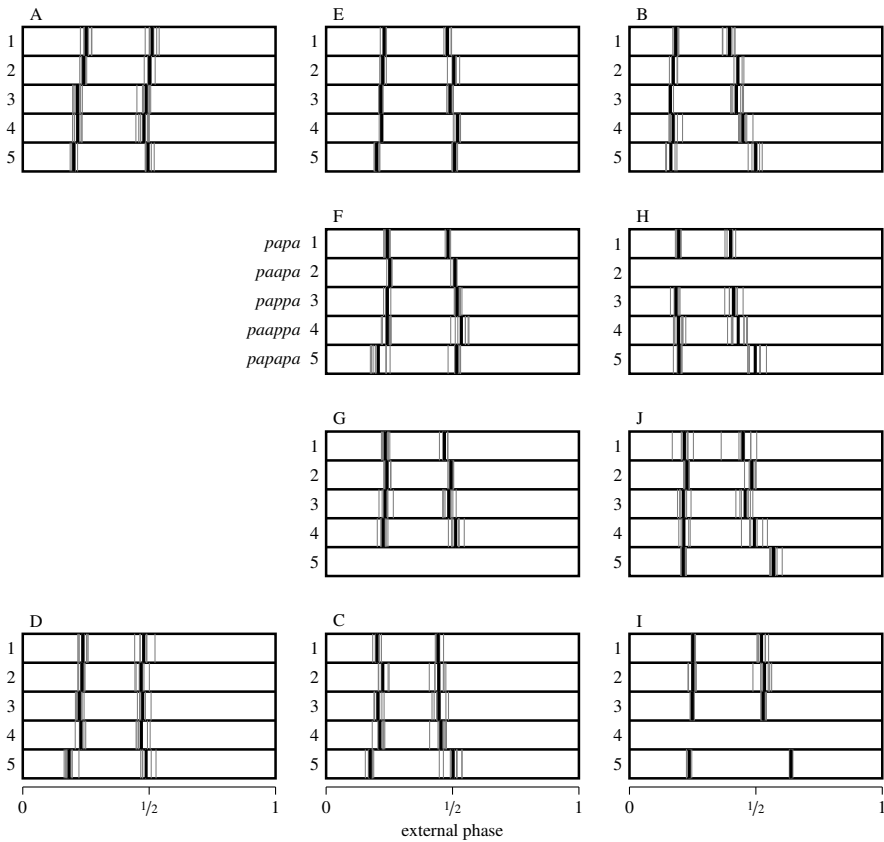


Figure 12. Bar plots for individual subjects (test phrases 1 = *papa*, 2 = *paapa*, 3 = *pappaa*, 4 = *paappaa*, 5 = *papapaa*). Average timing is shown as thick bars and individual repetitions in grey.

than the *papa* phrase does, and the *paappaa* phrase has them even later. In these speakers' timing, the *papapaa* phrase is close to the *paappaa* phrase in having a similar kind of relation between the external phases of beat 2 and beat 3, but the difference between these two phrases is that for both speakers beat 2 of the *papapaa* phrase comes earlier, which enables beat 3 to come earlier as well compared to the *paappaa* phrases (Figure 12 and Table 3). Speaker G also looks mora-timed because for him the *paapa* and *pappaa* phrases have later timing for beat 3 than the *papa* phrase does, and the *paappaa* phrase has it even later. The internal phase of G's *paappaa* phrase also speaks for mora-timing, as it is earlier than that of the other phrases (Figure 13). (G's fifth phrase was left out, as G had difficulties with the test word, producing it as [pa'pap:a].)

Speakers B, H, and J are between mora-timing and syllable-timing. As can be seen in Figure 13, for them, adding a mora affects the external phase of beat 3, making it later in the *paapa* and *pappaa* phrases than in the *papa* phrase (H's *paapa* phrase is missing due to technical problems), and in their *paappaa* phrases it is even later. However, *papapaa* in turn makes the external phase of beat 3 later. The internal

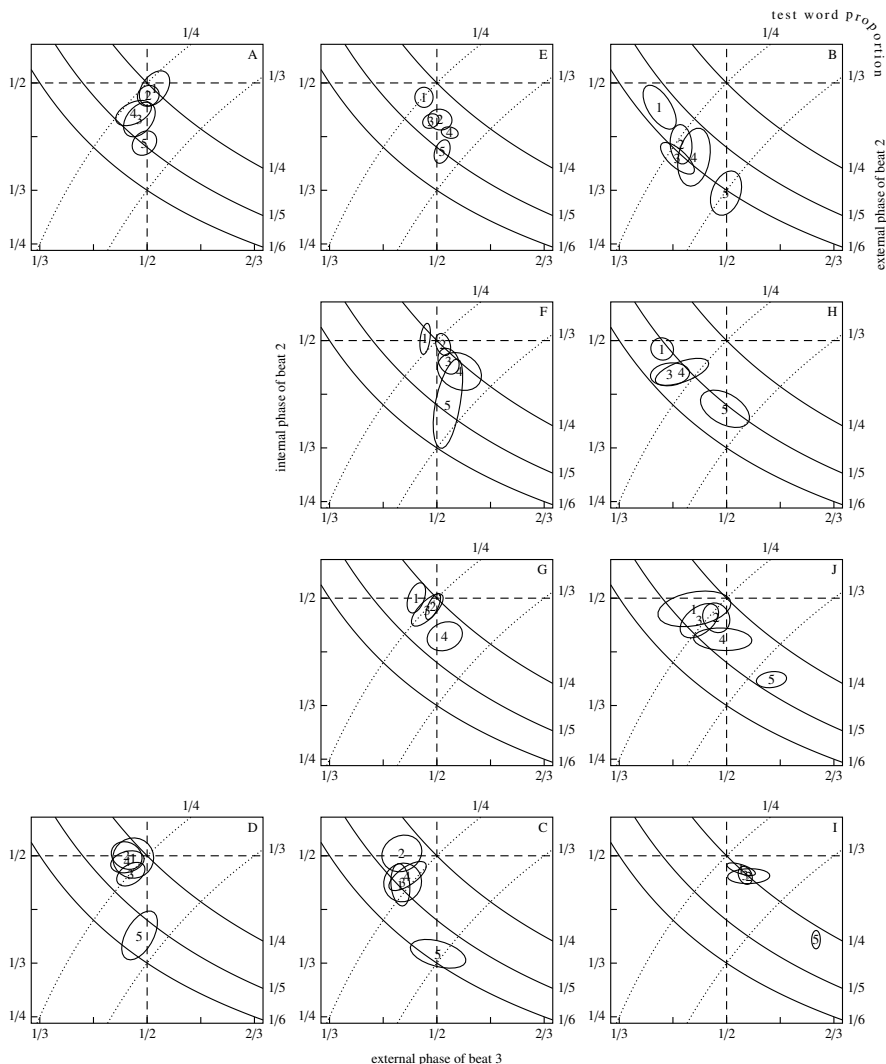


Figure 13. Tajima-type scatterplots for individual subjects (test phrases 1 = *papa*, 2 = *paapa*, 3 = *pappa*, 4 = *paappa*, 5 = *papapa*).

phases of beat 2 of their phrases 1–4 also show evidence of mora-timing. On the other hand, all of them have overlap in the variation of some of the beats: speaker B has overlap in phrases 2–4, speaker H in phrases 3 and 4, and speaker J in phrases 1–3. Their external phases of beat 2 are quite stable.

Since the timing of the internal phase suggests the degree of stress-timing, it is instructive to look at speakers’ posterior probabilities for the internal phase of beat 2, presented in Tables 4 and 5. In these tables, probabilities involving missing cases (*papapa* for G, *paapa* for H, and *paappa* for I) are enclosed in parentheses as a reminder, although in the Bayesian statistical analysis posterior probabilities are computed for these cases as well.

Table 4. Posterior probability for individual speakers that the average beat 2 internal phase is later than the average beat 2 internal phase for *papapa*

speaker:	A	B	C	D	E	F	G	H	I	J
<i>papa</i>	»»	»»	»»	»»	»»	»»	(»»)	»»	»»	»»
<i>paapa</i>	»»	»»	»»	»»	»»	»»	(»»)	(»»)	»»	»»
<i>pappa</i>	»»	»»	»»	»»	»»	»»	(»»)	»»	»»	»»
<i>paappa</i>	»»	»»	»»	»»	»»	»»	(0.99)	»»	(»»)	»»

»»: $p > 0.999$.

Table 5. Posterior probability for individual speakers that the average beat 2 internal phase is earlier than the average beat 2 internal phase for *papa*

speaker:	A	B	C	D	E	F	G	H	I	J
<i>paapa</i>	0.92	»»	«	0.41	»»	0.86	0.97	(0.80)	0.99	0.98
<i>pappa</i>	»»	»»	0.62	»	»»	»»	»	»»	0.99	»
<i>paappa</i>	»»	»»	0.37	0.88	»»	»»	»»	»»	(0.94)	»»
<i>papapa</i>	»»	»»	»»	»»	»»	»»	(»»)	»»	»»	»»

»»: $p > 0.999$, »: $p > 0.99$, «: $p < 0.01$.

Table 4 shows the posterior probability for individual speakers that the average internal phase of beat 2 in phrases 1–4 is later than the average internal phase for *papapa*. As can be seen, the probability for the internal phase of beat 2 of being later than the average internal phase of *papapa* is $p > 0.999$ for every speaker for every phrase, whose data was collected. This means that an additional syllable in the test word affects timing.

Table 5 shows the posterior probability for individual speakers that the average internal phase of beat 2 for phrases 2–5 is earlier than the average internal phase of beat 2 for the *papa* phrase. More variation between the speakers than in Table 4 can be seen here, although for most of the pairs the posterior probability is $p > 0.999$. Speaker C deviates from this pattern the most; for this speaker there is a high posterior probability ($p \approx 0.999$) that the average internal phase of beat 2 for *papa* is actually *later* than for *paapa*. The test word *paapa* exhibits the most variation among speakers and the lowest probability overall of having an earlier internal phase than *papa*. The most unified test word for all speakers is *papapa*, whose probability of having an earlier internal phase than *papa* is $p > 0.999$ for all speakers.

5. Discussion

The speech cycling task was successful, which could not be known beforehand, as speech cycling has not been easy to carry out in all languages, as noted by

Cummins (2002). The task appeared strange to the speakers, and some required several takes, but every speaker managed to produce all phrases, except for speaker G, whose *papapa* phrase was considered to deviate from the target so much that it was left out of the analysis. Although no speaker was able to follow the instructions not to breathe between cycles, this did not appear to hinder the rhythmical production, and the productions fell into the regular form that is sought in speech cycling. It can be concluded that speech cycling fits Finnish, and it also makes units visible that may have relevance for Finnish rhythm and mora count.

5.1 Timing of the CVVC syllable in Finnish

One of the research questions was to observe the effects of the mora in Finnish rhythm. This was done by using test words that differed from each other in both syllable and mora count (e.g. *papa* vs. *pappa* vs. *papapa*) and quantity shape (*paapa* vs. *pappa* vs. *paappa*). The question with the test word *paappa* was to observe whether its quantity shape tends towards the trimoraic two-syllable test words (*paapa* and *pappa*), towards the trimoraic three-syllable test word (*papapa*), or whether it has another kind of timing.

The test word *paappa* was on average 400.3 ms, which is shorter than the three-syllable *papapa* (485.2 ms) but longer than the two-syllable *paapa* (382.9 ms) and *pappa* (385.7 ms). However, it is closer to the bimoraic two-syllable words than to the trimoraic, three-syllable word. The mean proportion of *paappa* is 26.8, which is closer to those of *paapa* (25.7) and *pappa* (25.8) than that of *papapa* (32.8) (the durations and proportions are presented in Table 2). The posterior distributions in Figure 6 show that the proportion of *paappa* is between the test words 1–3 on the one hand and *papapa* on the other. In order to have four morae, *paappa* would be expected to be even longer than *papapa*, which is not the case. When looking at the averages of the individual speakers (presented graphically in Figure 12), the proportions of *paappa* vary. For some speakers, *paappa* is grouped together with the two-syllable test words. For example, for speaker A, all four test words have their proportion within a little over 1 percentage point, and for C, *paappa* is almost the same as *pappa* and *papa*, whereas *papapa* has a bigger proportion. For some speakers, *papapa* is grouped together with *paappa*. For E, for example, they are almost identical. For D, *paappa* has a smaller proportion than *pappa*.

The durational differences between these three words are also not big enough to support the multimoraic hypothesis in Finnish. This means that similarly to the fact that words can have the same quantity but a different duration, words that have a different quantity can have the same mora count. These findings can be interpreted to suggest that a coda consonant is a mora in Finnish if it follows a short vowel but not if it follows a long vowel. Thus, in CVVCCV the VC sequence shares one mora. Sharing a mora between two segments is common in languages (Hubbard 1995, Nam 2004). It also supports the bimoraic hypothesis, according to which Finnish syllables can be divided into groups having one mora or two morae, and all syllables having more than one mora are seen as bimoraic regardless of the number of their segments, as opposed to the multimoraic hypothesis, in which each segment increases the mora count one by one. This conclusion is the same as that of O'Dell et al. (2008).

5.2 Re-visiting the Finnish rhythm type

The speech cycling task exhibited units that can be seen as the basis of linguistic rhythm. The results were not unexpected for speech cycling: conclusive evidence for or against a single rhythm type did not emerge but there was an indication of components on which the rhythm can be considered to be based.

The test word *papapa* has an earlier external phase for beat 2 than the other test words, which is compatible with syllable-timing. Within the phrase, syllable-timing is not seen in the two-syllable test words, as different test words shift beat 3 differently. This is understandable, because due to quantity differences pure syllable-timing is not expected without heavy durational compensation, which does not happen here. Some individual speakers, however, are very close to syllable-timing (D, I).

As expected, in general Finnish did not prove to be stress-timed, although in regard to the external phase, speaker A comes close to it. According to Tajima et al. (1999), Zawaydeh et al. (2002), and Tajima & Port (2003), in order to be interpreted as stress-timed, the internal phase of beat 2 should always be at 1/2, and neither speaker A nor any other speaker exhibited that. The internal phase decreased when the mora or syllable count increased, which means that the phase of the stressed syllable is not fixed, which in turn means that the proportion of the word that increases with the mora or syllable count also increases in duration. In pure stress-timing, the additional segments should be compressed, which did not happen in the present study.

Mora-timing can be seen in the results. Increasing quantity, i.e. adding a mora, increases the proportion of the test word and shifts the external phase for beat 3 further. Beat 2 is also timed earlier, except for *paapa*, for which it is timed later. One observation which speaks against mora-timing is the difference between the phrases with *paapa* and *pappa*. In pure mora-timing, they would be expected to be the same, which is not the case here. No rhythm theory proposes this kind of timing that is based on the quantity shape. These two test words have the same segments on every level (mora and syllable count, phoneme count, and phoneme identity) and almost the same proportion (although *pappa* has a greater probability of having a slightly greater proportion), so they should be expected to have the same timing as well. Although according to Lehtonen (1970) segment durations vary between different quantity shapes, this should not affect timing when the vocalic onset is considered to serve as the beat that the speaker is assumed to place at some particular phase (Cummins & Port 1998). Since the vocalic onset is what is placed at simple harmonic phases, the duration of the preceding consonant should not have an effect on that, but it should be expected to make room for the vocalic onset instead. This means that a possible difference in the duration of the initial [p] should not have any relevance.

The differences between speakers can also be a sign of mora-timing, as Tajima & Port (2003) found that Japanese speakers had different kinds of rhythm strategies, whereas English speakers all used the same strategy. This indicates that greater variation in timing may be due to mora-timing. On the other hand, Japanese has been the only putative mora-timed language investigated with speech cycling experiments so far. The relevance of the mora in Finnish rhythm may also be

supported by the findings of O'Dell & Nieminen (2019) that in addition to the number of syllables, the number of morae affects perceived speech rate in Finnish.

Warner & Arai (2001b) note that in Japanese the number of segments were comparable with morae in determining a word's duration. In the present study the phonemes were kept the same, so the grouping of the segments into quantity shapes can predict the duration, which means that the rhythm unit is the quantity shape. Because segmental timing differences in Finnish arose in speech cycling, they can be considered more important than the segments' differences in Japanese: in Japanese those were observable in spontaneous speech (Warner & Arai 2001b) but not in controlled carrier sentences (Port et al. 1987), which suggests that in Finnish they are a more prominent part of the timing, as they are not erased in controlled speech. The effect of segments or phonemes on timing has also been noted by Yun (1998), who suggests that Korean could be described as segment- or phoneme-timed.

The number of feet was not varied in this study, but it is possible to draw some conclusions about feet based on the results. The timing of the phrases that contain a two-syllable test word is more similar to each other and differs from that of phrase 5, which has a three-syllable test word. For average *papapa* phrase, the test word has an earlier external and internal phase than the other four phrases but it also pushes the external phase of beat 3 further forward (see Figures 7 and 8 and Table 3). When looking at the individual subjects (Figures 12 and 13), it can be seen that *papapa* has a different external phase for many of them. For some of the subjects, beat 2 in the *papapa* phrase is close to that of the other phrases, while beat 3 moves further, whereas for some subjects beat 2 also moves back towards beat 1. There is no unified way to treat the timing in the *papapa* phrase.

If quantity alone determined the external phase of beat 2, we would not expect a big difference in it between *papapa* and the other test words. One explanation is that *papapa* breaks the two foot timing, which makes it take time from *tämä* as well as push beat 3 more than the two-syllable test words do. This suggests that the timing is based on feet of two syllables and not of two morae, as the difference in the mora count does not cause a similar change. This also fits the finding by Suomi & Ylitalo (2004) that three-syllable feet function differently from two-syllable feet.

This also fits into the final foot (beat 3) moving more in the *papapa* phrase than in the other phrases, because in the case of the *papapa* phrase, they are pushed by an additional syllable (or half of a foot), not by an additional mora. Based on this interpretation, an important unit in Finnish rhythm is a syllable-based foot, whose increasing quantity moves the beats of the other feet. Therefore, Finnish rhythm can be considered a combination of foot-timing and syllable-timing, as suggested for Estonian by Nolan & Asu (2009).

Tajima (1998:102–103) considers isochrony as certain units being attracted to regular parts of the cycle. In his study, stressed syllables in English were placed on these attractors. In the present study, a fixed attractor for the beats in every phrase was not found, but the test words were grouped according to their syllable count both on average and for individual speakers, that is, *papapa* differed in its phases from the other test words. As the quantity shape affects every word differently, it is reasonable to say that true attractors were not found, which again speaks against stress-timing. On the other hand, in the present study the speakers were not asked to imitate waltz rhythm, as in Tajima (1998) and Tajima & Port (2003); they were not

given instructions on how to organise their utterance, which may affect the emergence of attractors.

6. Conclusions

The results do not put Finnish into any individual rhythm type, which is an expected outcome in a speech cycling study. The traditional view of Finnish being syllable-timed is not completely true, but it cannot be denied either. The syllable has an effect but there are other factors as well.

As was expected, speech cycling did not provide clear proof of isochrony or that Finnish behaves according to a certain rhythm type: instead it exhibited different units that form the timing of Finnish. Both the mora and the syllable were shown to have relevance in timing. Morae and syllables affect phases of the beats differently. Increasing the syllable count moves beat 3 more than increasing the mora count, but it also moves beat 2 back, as a bigger proportion of the three-syllable word is reflected in the phases. Segments also have an effect, because they produce differences within the same mora count, making the phases of the *paapa* and *pappa* phrases differ from each other, which was not expected.

Based on these results, Finnish rhythm is a combination of syllable-timing and mora-timing, which favours two-syllable feet. The fact that the speakers managed the speech cycling task also tells something about Finnish rhythm on a broad level. Speech cycling clearly brought forth possible rhythm units in Finnish.

Acknowledgements. I present thanks to the four anonymous peer reviewers for their constructive feedback and to Michael L. O'Dell for his help with the figures and statistical analysis. This study was funded by the Emil Aaltonen Foundation.

References

- Abercrombie, David. 1965. *Studies in phonetics and linguistics*. Oxford: Oxford University Press.
- Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Allen, W. Sidney. 1973. *Accent and rhythm: Prosodic features of Latin and Greek: A study in theory and reconstruction*. Cambridge University Press.
- Aoyama, Katsura. 2001. *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions*. Boston: Kluwer Academic.
- Arnhold, Anja. 2014. *Finnish prosody: Studies in intonation and phrasing*. Johann Wolfgang Goethe-Universität PhD dissertation.
- Arvaniti, Amalia. 2009. Rhythm, timing and the timing of rhythm. *Phonetica* 66(1–2). 46–63.
- Auer, Peter. 1989. Some ways to count morae: Prokosch's law, Streitberg's law, Pfalz's law, and other rhythmic regularities. *Linguistics* 27. 1071–1102.
- Beckman, Mary. 1982. Segment duration and the 'mora' in Japanese. *Phonetica* 39(2–3). 113–135.
- Bloch, Bernard. 1950. Studies in colloquial Japanese IV: Phonemics. *Language* 26(1). 86–125.
- Chung, Younah & Amalia Arvaniti. 2013. Speech rhythm in Korean: Experiments in speech cycling. *Proceedings of Meetings on Acoustics ICA2013* 19(1). 060216.
- Connell, Bruce & Amalia Arvaniti (eds.). 1995. *Phonological and phonetic evidence: Papers in laboratory phonology IV*. Cambridge: Cambridge University Press.
- Cummins, Fred. 2002. Speech rhythm and rhythmic taxonomy. In Bernard Bel & Isabelle Marlien (eds.), *Speech Prosody 2002: Proceedings of the First International Conference on Speech Prosody*, 121–126. Aix-en-Provence, France: Laboratoire Parole et Langage.
- Cummins, Fred. 2013. Joint speech: The missing link between speech and music? *Percepta: Revista de Cognição Musical* 1(1). 17–32.

- Cummins, Fred & Robert Port. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26(2). 145–171.
- Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11(1). 51–62.
- Dauer, Rebecca M. 1987. Phonetic and phonological components of language rhythm. In Tamaz Gamkrelidze (ed.), *Proceedings of the 11th International Congress of Phonetic Sciences*, vol. 5, 447–450. Tallinn: Academy of Sciences of the Estonian SSR.
- Egozcue, Juan José & Vera Pawlowsky-Glahn. 2019. Compositional data: The sample space and its structure. *TEST* 28. 599–638.
- Fox, Anthony. 2002. *Prosodic features and prosodic structure: The phonology of 'suprasegmentals'*. Oxford: Oxford University Press.
- Gelman, Andrew & Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Grabe, Esther & Ee Ling Low. 2002. Durational variability in speech and the rhythm class hypothesis. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory phonology 7*, 515–546. Berlin: De Gruyter Mouton.
- Hoequist, Charles, Jr. 1983a. Durational correlates of linguistic rhythm categories. *Phonetica* 40(1). 19–31.
- Hoequist, Charles, Jr. 1983b. Syllable duration in stress-, syllable- and mora-timed languages. *Phonetica* 40(3). 203–237.
- Hubbard, Kathleen. 1995. Toward a theory of phonological and phonetic timing: Evidence from Bantu. In Connell & Arvaniti (eds.), 168–187.
- Iivonen, Antti. 1998. Intonation in Finnish. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 311–327. Cambridge: Cambridge University Press.
- Karlsson, Fred. 1983. *Suomen kielen äänne- ja muotorakenne* [the phonology and morphology of Finnish]. Helsinki: WSOY.
- Katz-Brown, Jason & Sam Mandal. 2013. Rhythm, metrics, and the link to phonology. In Karimi Sarvnaz & Karin Verspoor (eds.), *Proceedings of the Australasian Language Technology Association Workshop 2013 (ALTA 2013)*, 112–117.
- Krull, Diana & Olle Engstrand. 2003. Speech rhythm: Intention or consequence? Cross-language observations on the hyper/hypo dimension. *Phonum* 9. 133–136.
- Lehiste, Ilse. 1990. Phonetic investigation of metrical structure in orally produced poetry. *Journal of Phonetics* 18(2). 123–133.
- Lehtonen, Jaakko. 1970. *Aspects of quantity in standard Finnish* (Studia Philologica Jyväskyläensia VI). Jyväskylä: University of Jyväskylä.
- Malisz, Zofia. 2013. *Speech rhythm variability in Polish and English: A study of interaction between rhythmic levels*. Adam Mickiewicz University PhD dissertation.
- Nam, Hosung. 2004. Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory phonology 9*, 483–506. Berlin: Mouton De Gruyter.
- Nolan, Francis & Eva Liina Asu. 2009. The pairwise variability index and coexisting rhythms in language. *Phonetica* 66(1–2). 64–77.
- O'Connor, J. D. 1973. *Phonetics: A simple and practical introduction to the nature and use of sound in language*. London: Penguin.
- O'Dell, Michael, Mieta Lennes & Tommi Nieminen. 2008. Hierarchical levels of rhythm in conversational speech. In Plínio A. Barbosa, Sandra Madureira & César Reis (eds.), *Speech Prosody 2008: Fourth Conference on Speech Prosody*, 355–358. Campinas, Brazil: Luso-Brazilian Association of Speech Sciences.
- O'Dell, Michael & Tommi Nieminen. 2006. Tahdin ajoitus suomessa oskillaattorimallin näkökulmasta [Timing of the foot in the coupled oscillator model]. In Reijo Aulanko, Leena Wahlberg & Martti Vainio (eds.), *Fonetiikan päivät 2006/The phonetics symposium 2006*, 134–143. University of Helsinki.
- O'Dell, Michael & Tommi Nieminen. 2019. Syllable rate, syllable complexity and speech tempo perception in Finnish. In Sasha Calhoun et al. (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 622–626. Melbourne, Australia: Australasian Speech Science and Technology Association.
- Ogden, Richard. 1995. 'Where' is timing? Comments on Smith. In Connell & Arvaniti (eds.), 223–234.
- Pamies Bertrán, Antonio. 1999. Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. *Language Design: Journal of Theoretical and Experimental Linguistics* 2. 103–130.
- Plummer, Martyn. 2017. *JAGS version 4.3.0 user manual*.

- Pointon, Graham E. 1980. Is Spanish really syllable-timed? *Journal of Phonetics* 8(3). 293–304.
- Port, Robert, Fred Cummins & Michael Gasser. 1996. A dynamic approach to rhythm in language: Toward a temporal phonology. In B. Luka & B. Need (eds.), *Papers from the 31st Regional Meeting of the Chicago Linguistic Society*, vol. 1, 375–397.
- Port, Robert, Jonathan Dalby & Michael O'Dell. 1987. Evidence for mora timing in Japanese. *The Journal of the Acoustical Society of America* 81(5). 1574–1585.
- Port, Robert, Keiichi Tajima & Fred Cummins. 1999. Speech and rhythmic behavior. In G. J. P. Savelsbergh, H. van der Maas & P. C. L. van Geert (eds.), *The non-linear analysis of developmental processes*, 53–78. Amsterdam: Elsevier.
- Port, Robert F., Salman Al-Ani & Shosaku Maeda. 1980. Temporal compensation and universal phonetics. *Phonetica* 37(4). 235–252.
- R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Ramus, Franck, Marina Nespor & Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73(3). 265–292.
- Rathcke, Tamara V. & Rachel H. Smith. 2015. Speech timing and linguistic rhythm: On the acoustic bases of rhythm typologies. *The Journal of the Acoustical Society of America* 137(5). 2834–2845.
- Roach, Peter. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In David Crystal (ed.), *Linguistic controversies: Essays in linguistic theory and practice in honour of F. R. Palmer*, 73–79. London: Arnold.
- Sadeniemi, Matti. 1949. *Metrikkamme perusteet* [The fundamentals of Finnish metrics]. Helsinki: Finnish Literature Society (SKS).
- Suomi, Kari. 2004. Moraic patterns in Finnish prosody and lexicon. In Maija S. Peltola & Jyrki Tuomainen (eds.), *Studies in speech communication: Dedicated to Ilkka Raimo on his 65th birthday* (Publications of the Department of Finnish and General Linguistics of the University of Turku 72), 87–97. Turku: University of Turku.
- Suomi, Kari. 2005. Temporal conspiracies for a tonal end: Segmental durations and accentual f_0 movement in a quantity language. *Journal of Phonetics* 33(3). 291–309.
- Suomi, Kari. 2007. On the tonal and temporal domains of accent in Finnish. *Journal of Phonetics* 35(1). 40–55.
- Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2003. Durational and tonal correlates of accent in Finnish. *Journal of Phonetics* 31(1). 113–138.
- Suomi, Kari & Riikka Ylitalo. 2004. On durational correlates of word stress in Finnish. *Journal of Phonetics* 32(1). 35–63.
- Tajima, Keiichi. 1998. *Speech rhythm in English and Japanese: Experiments in speech cycling*. Indiana University PhD dissertation.
- Tajima, Keiichi & Robert F. Port. 2003. Speech rhythm in English and Japanese. In John Local, Richard Ogden & Rosalind Temple (eds.), *Phonetic interpretation: Papers in laboratory phonology VI*, 317–334. Cambridge: Cambridge University Press.
- Tajima, Keiichi, Bushra Zawaydeh & Mafuyu Kitahara. 1999. A comparative study of speech rhythm in Arabic, English and Japanese. In John J. Ohala et al. (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California.
- Trubetzkoy, Nikolay. 1969. *Principles of phonology*. Berkeley & Los Angeles: Vandenhoeck & Ruprecht, University of California Press. (Translated by Christiane A. M. Baltaxe from N. S. Trubetzkoy 1958, *Grundzüge der Phonologie*, Göttingen.)
- Turk, Alice, Satsuki Nakai & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff et al. (eds.), *Methods in empirical prosody research*, 1–28. Berlin: Walter de Gruyter.
- Vasisith, Shravan, Bruno Nicenboim, Mary E. Beckman, Fangfang Li & Eun Jong Kong. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71. 147–161.
- Warner, Natasha & Takayuki Arai. 2001a. Japanese mora-timing: A review. *Phonetica* 58(1–2). 1–25.
- Warner, Natasha & Takayuki Arai. 2001b. The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America* 109(3). 1144–1156.
- Wenk, Brian J. & François Wioland. 1982. Is French really syllable-timed? *Journal of Phonetics* 10(2). 193–216.

- Wiik, Kalevi. 1991. On a third type of speech rhythm: Foot timing. In *Proceedings of the Twelfth International Congress of Phonetic Sciences*, vol. 3, 298–301. Aix-en-Provence, France: University of Provence.
- Yun, Ilsung. 1998. *A study of timing in Korean speech*. University of Reading PhD dissertation.
- Zawaydeh, Bushra Adnan, Keiichi Tajima & Mafuyu Kitahara. 2002. Discovering Arabic rhythm through a speech cycling task. In Dilworth P. Parkinson & Elabbas Benmamoun (eds.), *Perspectives on Arabic linguistics XIII-XIV: Papers from the Thirteenth and Fourteenth Annual Symposia on Arabic Linguistics* (Current Issues in Linguistic Theory 230), 39–58. Amsterdam: John Benjamins.