

# **The Influence of Speech Rate on Rhythm Patterns**

*Maartje Schreuder and Dicky Gilbers*

## **1. Introduction<sup>1</sup>**

The topic of this paper is how rhythmic variability in speech can be accounted for both phonologically and phonetically. The question is whether a higher speech rate leads to adjustment of the phonological structure, or just to 'phonetic compression', i.e. shortening and merging of vowels and consonants, with preservation of the phonological structure. We claim that the melodic content of a phonological domain is indeed adjusted optionally when the speech rate increases. In other words, every speech rate has its own preferred register, in terms of Optimality Theory (Prince and Smolensky, 1993) its own ranking of constraints.

We will investigate prosodic variability as part of our main research project, which involves a comparison of the analyses of music and language. Our ultimate aim is to provide evidence for the assumption that every temporal behavior is structured similarly (cf. Liberman, 1975). Gilbers and Schreuder (to appear) show that Optimality Theory owes a lot to the constraint-based music theory of Lerdahl and Jackendoff (1983). Based on the great similarities between language and music we claim that musical knowledge can help in solving linguistic issues.

In this paper, we will show that clashes are avoided in allegro tempo. In both language and music distances between beats are enlarged, i.e. there appears to be more melodic content between beats. To illustrate this, we ran a pilot experiment in which we elicited fast speech. As expected, speech rate plays a role in rhythmic variability.

The paper is organized as follows. In section 2 the data of the experiment is introduced. Section 3 is addressed to the phonological framework of Optimality Theory and the different rankings of andante and allegro speech. The method of the experiment is discussed in section 4 and the auditive and acoustic analyses plus the results follow in section 5. The perspectives of our analysis will be discussed in the final section.

## 2. Data

We will discuss three types of rhythmic variability in Dutch. The first we will call “stress shifts to the right”; the second “stress shifts to the left” and the third “beat reduction”. In the first type as exemplified in *stúdiètòelàge* (*s w s w w*) ‘study grant’, we assume that this compound can be realized as *stúdiètòelàge* (*s w w s w*) in allegro speech. *Perfèctioníst* (*w s w s*) is an example of “stress shift to the left” and we expect a realization *pèrfectioníst* (*s w w s*) in allegro speech. The last type does not concern a stress shift, but a stress reduction. In *zùidàfrikáans* (*s s w s*) ‘South African’ compounding of *zuid* and *afrikaans* results in a stress clash. In fast speech this clash is avoided by means of reducing the second beat: *zùidafrikáans* (*s w w s*). Table 1 shows a selection of our data.

Table 1. Data

Type 1: stress shift to the right (andante: *s w s w w*; allegro: *s w w s w*)

<i>stu die toe la ge</i>	‘study grant’
<i>weg werp aan ste ker</i>	‘disposable lighter’
<i>ka mer voor zit ter</i>	‘chairman of the House of Parliament’

Type 2: stress shift to the left (andante: *w s w s*; allegro: *s w w s*)

<i>per fec tio nist</i>	‘perfectionist’
<i>a me ri kaan</i>	‘American’
<i>vi ri li teit</i>	‘virility’

Type 3: beat reduction (andante: *s s w s*; allegro: *s w w s*)

<i>zuid a fri kaans</i>	‘South African’
<i>schier mon nik oog</i>	‘name of an island’
<i>gre go ri aans</i>	‘Gregorian’

The different rhythmic patterns are accounted for phonologically within the framework of OT.

## 3. Framework and phonological analysis

The mechanism of constraint interaction, the essential characteristic of OT, is also used in the generative theory of tonal music (Lerdahl and Jackendoff, 1983). In both frameworks, constraint satisfaction determines grammaticality and in both frameworks the constraints are potentially

conflicting and soft, which means violable. Violation, however, is only allowed if it leads to satisfaction of a more important, higher ranked constraint. The great similarities between these theoretical frameworks make comparison and interdisciplinary research possible.

For example, restructuring rhythm patterns as a consequence of a higher playing rate is a very common phenomenon in music. In Figure 1 we give an example of re-/misinterpretation of rhythm in accelerated or sloppy playing.

Dotted notes rhythm → triplet rhythm

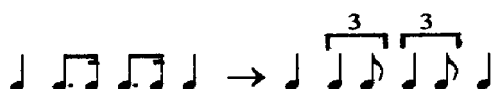


Figure 1. Rhythmic restructuring in music

In Figure 1, the “dotted notes rhythm” (left of the arrow) is played as a triplet rhythm (right of the arrow). In the dotted notes rhythm the second note has a duration which is three times as long as the third, and in the triplet rhythm the second note is twice as long as the third. In fast playing it is easier to have equal durations between note onsets. Clashes are thus avoided and one tries to distribute the notes, the melodic content, over the measures as evenly as possible, even if this implies a restructuring of the rhythmic pattern. To ensure that the beats do not come too close to each other in fast playing, the distances are enlarged, thus avoiding a staccato-like rhythm. In short, in fast tempos the musical equivalents of the Obligatory Contour Principle (OCP), a prohibition on adjacency of identical elements in language (McCarthy, 1986), become more important.

We claim that - just as in music - the allegro patterns in all the different types of data in Table 1 are caused by clash avoidance. There is a preference for beats that are more evenly distributed over the phrase. The different structures can be described phonologically as a conflict between markedness constraints, such as FOOT REPULSION (\*ΣΣ) (Kager, 1994), and OUTPUT - OUTPUT CORRESPONDENCE constraints (cf. Burzio, 1998) within the framework of OT. FOOT REPULSION prohibits adjacent feet and consequently prefers a structure in which feet are separated from each other by an unparsed syllable. This constraint is in conflict with PARSE-σ, which

demands that every syllable is part of a foot. OUTPUT - OUTPUT CORRESPONDENCE compares the structure of a phonological word with the structure of its individual parts. For example, in a word such as *fototoestel* 'photo camera', OUTPUT - OUTPUT CORRESPONDENCE demands that the rhythmic structure of its part *tóestel* 'camera' with a stressed first syllable is reflected in the rhythmic structure of the output. In other words, OUTPUT - OUTPUT CORRESPONDENCE prefers *fótotòestel*, with secondary stress on *toe*, to *fótotoestèl*, with secondary stress on *stel*.

Whereas the normal patterns in andante speech satisfy OUTPUT - OUTPUT CORRESPONDENCE, the preference for triplet patterns in fast speech is accounted for by means of dominance of the markedness constraint, FOOT REPULSION, as illustrated in Table 2.<sup>2</sup>

Table 2. Rhythmic restructuring in language

a. ranking in andante speech:

constraints → <i>fototoestel</i> candidates ↓	OUTPUT - OUTPUT CORRESPONDENCE	*ΣΣ	PARSE-σ
☞ (fóto)(tòestel)		*	
(fóto)toe(stèl)	*!		*

b. ranking in allegro speech:

constraints → <i>fototoestel</i> candidates ↓	*ΣΣ	OUTPUT - OUTPUT CORRESPONDENCE	PARSE-σ
(fóto)(tòestel)	*!		
☞ (fóto)toe(stèl)		*	*

Dutch is described as a trochaic language (Neijt and Zonneveld, 1982). Table 2a shows a preference for an alternating rhythm. The dactyl pattern as preferred in Table 2b, however, is a very common rhythmic pattern of prosodic words in languages such as Estonian and Cayuvava: every strong syllable alternates with two weak syllables (cf. Kager, 1994). We assume that the rhythm grammar, i.e. constraint ranking, of Dutch allegro speech

resembles the grammar of these languages. In the next section we will explore whether we can find empirical evidence for our hypothesis.

#### 4. Method

To find out whether people indeed prefer triplet patterns in allegro speech, we ran a pilot experiment in which we tried to elicit fast speech. Six subjects participated in a multiple-choice quiz in which they competed each other in answering twenty simple questions as quickly as possible. In this way, we expected them to speak fast without concentrating too much on their own speech. In Table 3 one of the quiz items is depicted.

Table 3. Quiz item

Q4	<i>President Bush is een typische</i>	‘President Bush is a typical ’
A1	<i>intellectueel</i>	‘intellectual’
A2	<i>amerikaan</i>	‘American’
A3	<i>taalkundige</i>	‘linguist’

We categorized the obtained data as allegro speech. As a second task the subjects were asked to read out the answers at a normal speaking rate embedded in the sentence *ik spreek nu het woord ... uit* 'now I pronounce the word ... '. This normal speaking rate generally means that the subjects will produce the words at a rate of approximately 180 words per minute, which we categorize as andante speech. All data were recorded on minidisk in a soundproof studio and normalized in CoolEdit in order to improve the signal-noise (S/N) ratio. Normalizing to 100% yields an S/N ratio approaching 0 dB.

Six trained listeners judged the data auditively and indicated where they perceived secondary stress. After this auditive analysis the data were phonetically analyzed in PRAAT (Boersma and Weenink, 1992). We compared the andante and allegro data by measuring duration, pitch, intensity, spectral balance and rhythmic timing (Sluijter, 1995; Couper-Kuhlen, 1993; Cummins & Port, 1998; Quené & Port, 2002; a.o.). Sluijter claims that, respectively, duration and spectral balance are the main correlates of primary stress. In our experiment, we are concerned with secondary stress.

For the duration measurements, the rhymes of the relevant syllables were observed. For example, in the allegro style answer A2 *amerikaan* in Table 3, we measured the first two rhymes and compared the values in Msec. with the values for the same rhymes at the andante rate. In order to make this comparison valid, we equalized the total durations of both realizations by multiplying the duration of the allegro with a so-called 'acceleration factor', i.e. the duration of the andante version divided by the duration of the allegro version. According to Eefting and Rietveld (1989) and Rietveld and Van Heuven (1997), the just noticeable difference for duration is 4,5%. If the difference in duration between the andante and the allegro realization did not exceed this threshold, we considered the realizations as examples of the same speech rate and neglected them for further analysis.

For the pitch measurements, we took the value in Hz in the middle of the vowel. The just noticeable difference for pitch is 2,5% ('t Hart et al, 1990). For the intensity measurements, we registered the mean value in dB of the whole syllable.

The next parameter we considered concerns spectral balance. Sluijter (1995) claims that the spectral balance of the vowel of a stressed syllable is characterized by more perceived loudness in the higher frequency region, because of the changes in the source spectrum due to a more pulse-like shape of the glottal waveform. The vocal effort, which is used for stress, generates a strongly asymmetrical glottal pulse. As a result of the shortened closing phase, there is an increase of intensity around the four formants in the frequency region above 500 Hz. Following Sluijter (1995) we compared the differences in intensity of the higher and lower frequencies of the relevant syllables in both tempos.

Finally, we considered rhythmic timing. The idea is that the beats in speech are separated from each other at an approximately equal distance independent of the speech rate. In other words, a speaker more or less follows an imaginary metronome. If he/she speaks faster, more melodic content will be placed between beats, which results in a shift of secondary stress. This hypothesis will be confirmed if the distance between the stressed syllables in the andante realization of an item, e.g. *stu* and *toe* in *studietoelage*, approximates the distance between the stressed syllables in the allegro realization of the same item, e.g. *stu* and *la*. If the quotient of the andante beat interval duration divided by the allegro beat interval duration approximates 1, we expect perceived restructuring.

## 5. Results

### 5.1. Auditive analysis

Before we can present an auditive analysis of the data, we have to find out whether or not the quiz design was successful. The results show that the quiz indeed triggers faster speech by all subjects. Figure 2 shows their acceleration factors. Subjects 1, 2 and 4 turned out to be the best accelerating speakers, whereas subjects 3, 5 and 6 showed less difference in duration between andante and allegro realizations. The mean acceleration factor for the three fast speakers is 1.31, whereas the mean acceleration factor for the three slow speakers is 1.13.

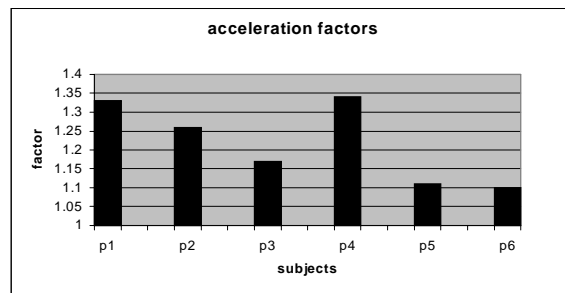


Figure 2. Acceleration factors of all subjects

Figure 3 shows the mean durations of the items at both speech rates. It shows that the best accelerating speakers are also the fastest speakers. We expect to find more restructured patterns for these speakers, mainly subjects 1 and 4, in comparison to the slower speakers, such as subjects 3 and 6.

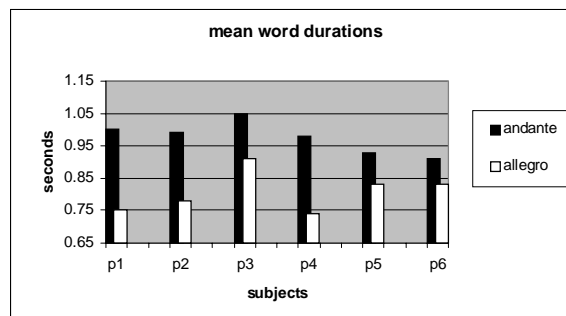


Figure 3. Mean word durations

Figure 4 shows that most subjects prefer patterns in which from a phonological point of view markedness constraints dominate the correspondence constraints at both rates for right and left shift data, but not for beat reduction data. There are slightly more restructured patterns in allegro tempo, although the differences are quite small.

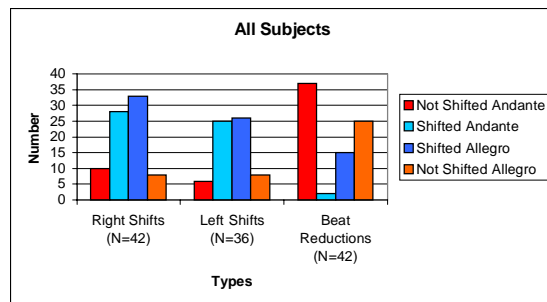


Figure 4. All subjects: Number of restructured items per type

When we take the results of two fast subjects apart, subjects 1 and 4, we observe a stronger preference for restructuring in allegro speech and no restructuring in andante speech, as shown in Figure 5. In other words, the fast subjects display both a greater difference in word durations in andante and allegro speech, and more variability in their speech patterns due to tempo than the slow subjects do.

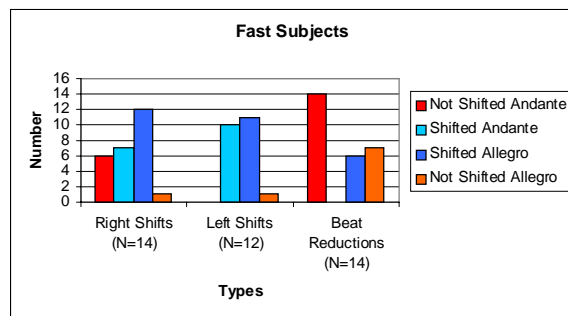


Figure 5. Fast subjects: Number of restructured items per type

Obviously, the preference for restructuring the rhythmic pattern in allegro speech is not an absolute preference. Sometimes restructuring does not take place in allegro speech, but on the other hand restructured patterns also show up in andante speech.<sup>3</sup> Some items were realized with the same rhythmic pattern irrespective of the tempo. Therefore, we also looked at the word pairs with a different rhythmic pattern in both tempos for each subject. We observe that the relatively fast speakers p1, p2 and p4, show the expected pattern according to our hypothesis, which means that they show a restructured pattern in allegro tempo, as shown in Figure 6 for the right shifts.

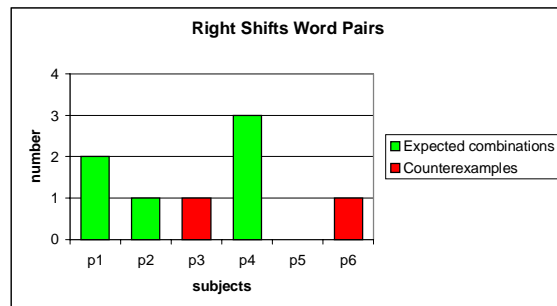


Figure 6. Right Shifts: Expected combinations

Two of the relatively slow speakers, p3 and p6, show one counterexample each, where the subject prefers the restructured patterns in andante tempo. The other slow speaker, P5, displays no different patterns in andante and

allegro at all. Clearly, we have two different groups of speakers and this observation strengthens our claim that restructuring relates to speech rate.

Some items, such as *hobbywerkruinte* (Type 1) 'hobby room', never show a stress shift and other items, such as *viriliteit* (Type 2) 'virility', prefer the shifted pattern in both tempos for all subjects. Possibly, the syllable structure plays an important role; open syllables seem to lose stress more easily than closed ones.

## 5.2. Acoustic analysis

In the current state of phonological research, embodied in e.g. laboratory phonology, much value is set on acoustic evidence for phonological analyses. Studies such as Sluijter (1995) and Sluijter and Van Heuven (1996) provide acoustic correlates for primary stress. In our study we are concerned with beat reduction and secondary stress shifts and we wonder whether or not the same acoustic correlates hold for secondary stress. Shattuck Hufnagel et al (1994) and Cooper and Eady (1986) do not find acoustic correlates of rhythmic stress at all. They claim that it is not entirely clear which acoustic correlates are appropriate to measure, since these correlates are dependent on the relative strength of the syllables of an utterance. The absolute values of a single syllable can hardly be compared without reference to their context and the intonation pattern of the complete phrase. Huss (1978) claims that some cases of perceived rhythmic stress shift may be perceptual rather than acoustic in nature. Grabe and Warren (1995) also suggest that stress shifts can only be perceived in rhythmic contexts. In isolation, the prominence patterns are unlikely to be judged reliably. In the remainder of this paper we try to find out if we can support one of these lines of reasoning. In other words, are we able to support our perceived rhythmic variability with a phonetic analysis? Therefore, we measured the duration, pitch, intensity, spectral balance and rhythmic timing of the relevant syllables as realized by subject P1.

Because Dutch is a quantity-sensitive language, the duration of the relevant syllable rhymes was considered. Onsets do not contribute to the weight of a syllable. In Figure 7, the duration analysis is shown for Type 2 data (left shifts). The four columns indicate, respectively, the duration of the rhyme of the first and second syllable in andante speech, and the duration of the first and second one in allegro speech. According to Sluijter (1995), duration is the main correlate of primary stress. As a starting point,

we adopt her claim for our analysis of secondary stress. Our measurements would confirm our hypothesis and our auditive analysis, if the second column were higher than the first one and if the fourth column were lower than the third one. In that case, the subject would realize a word such as *perfectionist* as *perfectionist* in andante tempo and as *perfectionist* in allegro tempo.

In the andante tempo, three out of six items show the dominant correspondence pattern and in the allegro tempo, four out of six items show the dominant markedness pattern. That is hardly a preference and it does not confirm our auditive analysis of the same data. Furthermore, if we consider the word pairs with different patterns, there is only one pair that has the ideal ratio: the patterns of *amerikaan*.

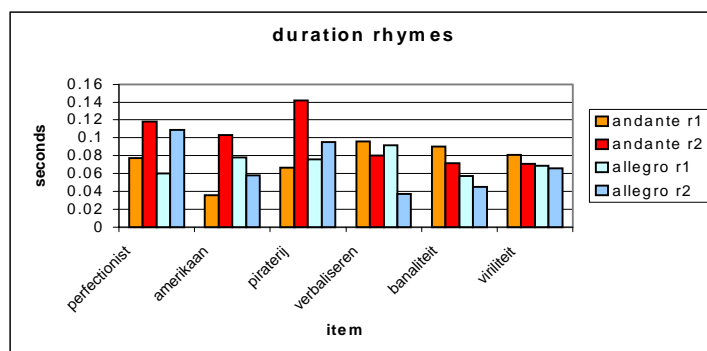


Figure 7. Duration (Left Shifts by Subject P1)

If duration does not enable us to confirm our auditive findings, maybe pitch is the main stress correlate for this speaker. However, pitch measurements reveal the same fuzzy result as the duration measurements. Again, only one pattern confirms the auditive analysis. This time it is not the item *amerikaan*, but the item *perfectionist*. Moreover, the differences in pitch in this item do not exceed the threshold of the 2.5%, which is the just noticeable difference for pitch. We also analyzed the mean intensity value of the relevant vowels without recognizable patterns between allegro and andante style. These results support the analyses of Sluijter (1995) and Sluijter and Van Heuven (1996), who also claim that the intensity parameter does not contribute much to the perception of stress.

Next, we considered the spectral balance. In order to rule out the influence of the other parameters, we monotonized the data for volume and

pitch. Then we selected the relevant vowels and analyzed them as a cochleagram in PRAAT. The cochleagram simulates the way the tympanic membrane functions, in other words the way in which we perceive sounds. In Figure 8 we show two cochleagrams of the vowel [a] in the fourth syllable of, respectively, *stúdiètòelàge* 'study grant' (Type 1) in andante tempo and *stúdiètòelàge* in allegro tempo. This item was taken from a pre-study. The allegro data show the expected increased perceived loudness in the higher frequencies, indicated by means of shades of gray; the darker gray the more perceived loudness.

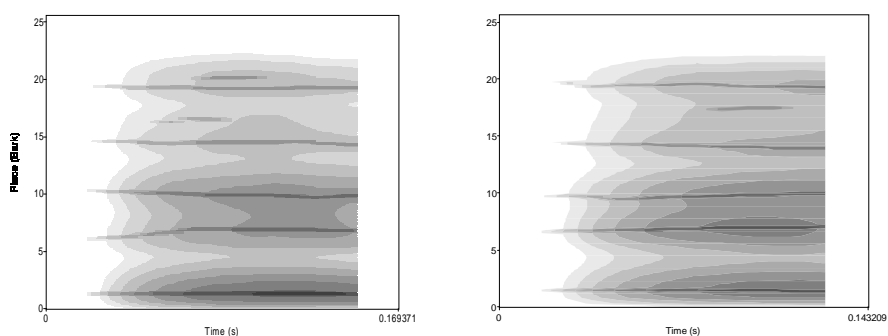


Figure 8. Cochleagrams of [a] in studietoelage

The right cochleagram (stressed [a]) in Figure 8 shows increased perceived loudness in the regions of approximately 5 to 22 Bark in the allegro version of [a] in comparison with the left cochleagram (unstressed [a]). This confirms the results of the study of primary stress in Sluijter (1995). If we convert this perceptive, almost logarithmic, Bark scale into its linear counterpart, the Hertz scale, this area correlates with the frequency region of 3 to 10 kHz.

In order to measure perceived secondary stress, we will measure the relative loudness in the different frequency regions in Phon.<sup>4</sup> According to Sluijter (1995) stressed vowels have increased loudness above 500 Hz compared to the same vowel in an unstressed position. This can be shown if we take a point in time from both cochleagrams in Figure 8 in which the F1 reaches its highest value (following Sluijter, 1995). In Figure 9 the values in Phon are depicted for these points and plotted against the Bark values in 25 steps.

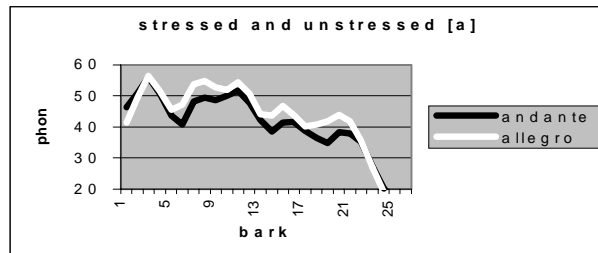


Figure 9. Loudness in Phon

The white line in Figure 9 indicates the pattern of the allegro stressed [a] in *studietoelage* and the black line indicates the pattern of the andante unstressed [a]. We see increased loudness in the region of 13 to 21 Bark, which correlates with the most sensitive region of our ear. The mean Phon value in Figure 9 between 5 and 21 Bark is 43.6 Phon for the andante unstressed [a] and 47.4 Phon for the allegro stressed [a]; a mean difference of 3.8 Phon.

Now, let us see whether or not we can find similar results for our subject P1. Figure 10 shows that the spectral balance confirms the leftward stress shift we perceived in the allegro realization of *amerikaan*. The first syllable vowel in allegro tempo is characterized by more loudness in the higher frequency regions than its andante counterpart. In the second syllable vowel it is just the other way around.

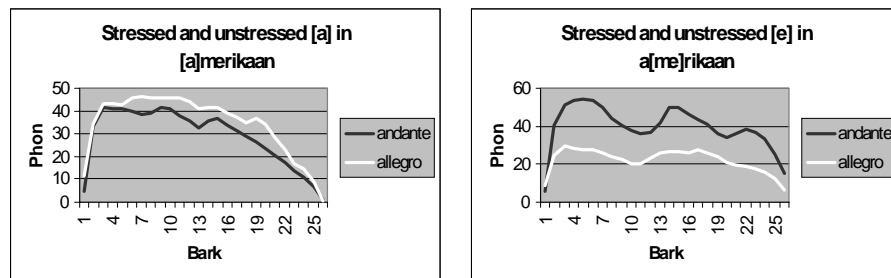


Figure 10. Spectral balance comparison of the first two vowels of *amerikaan*

Unfortunately, not all spectral balance data confirm our auditory analysis. For example, we claimed that the pitch analysis of the stress shift in

*perfectionist* did confirm our auditive analysis. Therefore, we expected more loudness in the allegro realization of the first vowel and less loudness in the allegro realization of the second vowel, but it appeared that there is relatively more loudness in the andante realization of *per*. This result contradicts our auditive and our pitch analysis.

We have to conclude that the different phonetic analyses contradict each other. Sometimes the perceived stress shift is characterized by a longer duration of the stressed syllable; sometimes a relatively higher pitch characterizes it. The results of our spectral balance analysis show that the differences in loudness pattern with differences in duration. In our perceived stress shift in allegro *perfectionist*, pitch turned out to be the decisive correlate, whereas duration and spectral balance measurements indicated no shift at all. On the other hand, the perceived shift in allegro *amerikaan* was confirmed by the duration and spectral balance analyses together, whereas pitch measurements indicated the opposite pattern. For most perceived stress shifts, however, the acoustic correlates did not give any clue.

Finally, we will consider whether the perception of restructuring depends on rhythmic timing. Just as in music, speech can be divided into a melodic string and a rhythmic string as partly independent entities. With respect to speech, the melodic string seems to be more flexible than the rhythmic one. Imagine that the rhythm constitutes a kind of metronome pulse to which the melodic content has to be aligned. The listener expects prominent syllables to occur with beats. This behavior is formulated as the *Equal Spacing Constraint*: prominent vowel onsets are attracted to periodically spaced temporal locations (Couper-Kuhlen, 1993; Cummins & Port, 1998; Quené & Port, 2002; a.o.). Dependent on speech rate the number of intervening syllables between beats may differ. Suppose the beat interval is constant at 300 msec., there will be more linguistic material in between in allegro speech, e.g. the two syllables *die* and *toe* in *stúdietoelàge*, than in andante speech, e.g. only one syllable *die* in *stúdiètòelàge*.

If indeed the perception of secondary stress shifts depends on rhythmic timing, i.e. the beat interval between prominent syllables in andante and allegro speech is approximately equal, than we expect that the duration quotient of the interval between, for example, *stu* and *toe* in the andante realization of *studietoelàge* and *stu* and *la* in the allegro realization approximates 1.

In our pre-study, the interval between the vowel onsets of the first and third syllable in *studietoelage* (andante) is 0.358 sec, whereas the interval between the first and the fourth syllable in the allegro realization of the same word is 0.328 sec. This means that the duration quotient is 1.091, which indeed approximates 1. In other words, this example supports the idea of the *Equal Spacing Constraint*.

Does the same result hold for our present data? We measured the beat intervals between all possible stress placement sites for all six subjects. Figure 11 depicts the duration quotients for subject 1. Figure 12 shows the beat intervals of the same data. It depicts as well the duration interval between the first and the third, as the first and fourth syllable for both speech rates. We expect restructuring for those data in which the line of the first to third syllable interval (andante (black line)) coincides with the line of the first to fourth syllable interval (allegro (white line)).

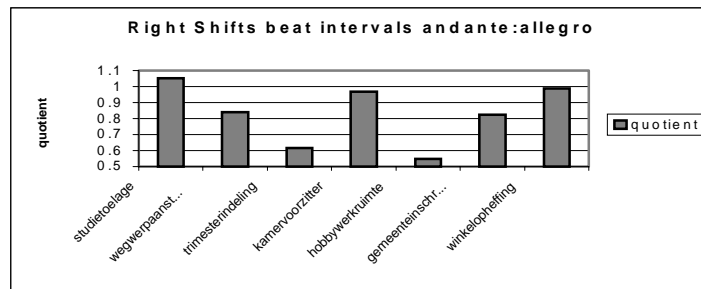


Figure 11. Quotient beat intervals of Subject P1

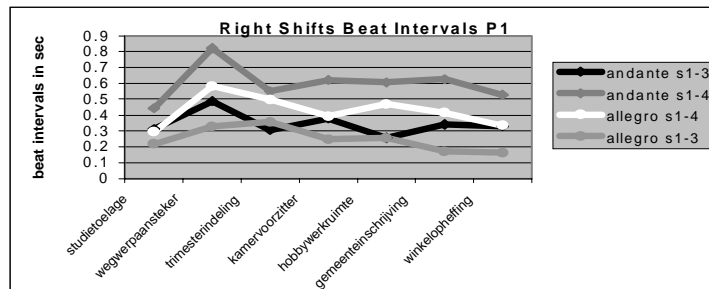


Figure 12. Beat intervals of Subject P1

The Figures 11 and 12 indicate that the relevant beat intervals of the items 1, 4 and 7, *studietoelage* 'study grant', *kamervoorzitter* 'chairman of the House of Parliament' and *winkelopheffing* 'closing down of a shop', respectively, coincide. In other words, we expect to hear restructuring in exactly these three items.

Unfortunately, our auditive analysis indicates only attested combinations of restructuring in items 2 and 6: *wegwerpaansteker* 'disposable lighter' and *gemeente-inschrijving* 'municipal registration', respectively. Obviously, rhythmic timing is not the decisive characteristic of perceived restructuring in allegro speech either.

## 6. Discussion and Conclusion

In section 4, we presented our phonological account of the restructuring within the framework of OT. Our main conclusion is that phonetic compression cannot be the sole explanation of the different rhythm patterns. Although the results cannot really confirm our hypothesis that there are different grammars, i.e. constraint rankings for different rates of speaking, there seems to be something that relates to speech rate. The fast speakers display different grammars, i.e. constraint rankings, for different rates of speaking. In their andante tempo, correspondence constraints prevail, whereas in allegro tempo markedness constraints dominate the correspondence ones. These preferences resemble the preferences of andante and allegro music. In both disciplines clashes are avoided in allegro tempo by means of enlarging the distances between beats.

In section 5, we attempted to confirm our phonological account with a phonetic analysis. Unfortunately, the phonetic correlates of stress - duration, pitch, intensity and spectral balance - do not show the expected and perceived differences in rhythm patterns in all pairs. Sluiter (1995) found out that duration is the main correlate of primary stress with spectral balance as an important second characteristic. In our analysis, however, neither differences in duration nor differences in spectral balance could identify secondary stress. Therefore, we have to conclude that our analysis supports earlier work by Shattuck Hufnagel et al (1994), Cooper and Eady (1986), Huss (1978) and Grabe and Warren (1995), who all claim that acoustic evidence for secondary stress cannot be found unambiguously. Although we did find some differences in duration, spectral balance or pitch, these differences were not systematically found in all pairs in which

we perceived rhythmic variability. Finally, we discussed rhythmic timing as a cue for variable patterns. However, the hypothesis that the duration between prominent syllables is approximately equal in both *andante* and *allegro* speech was not confirmed by the auditive analysis of the data. It seems that rhythmic restructuring is more a matter of perception than of production. At this point, the question remains: are we fooled by our brains and is there no phonetic correlate of the perceived phonological stress shifts in the acoustic signal or do we have to conclude that the real phonetic correlate of secondary stress has yet to be found?

## Notes

- 
- <sup>1</sup> This paper is an extension of our paper "Restructuring the melodic content of feet", which is submitted to the proceedings of the 9<sup>th</sup> International Phonology Meeting: Structure and melody, Vienna 2002. We wish to thank Grzegorz Dogil, Hidetoshi Shiraishi plus the participants of the 9<sup>th</sup> International Phonology Meeting, Vienna 2002 and the participants of the 11<sup>th</sup> Manchester Phonology Meeting, Manchester 2003 for their useful comments. We are also grateful to Sible Andringa, Nynke van den Bergh, Gerlof Bouma, John Hoeks, Jack Hoeksema, Wander Lowie, Dirk-Bart den Ouden, Joanneke Prenger, Ingeborg Prinsen, Femke Wester for participating in our experiment. We especially thank Wilbert Heeringa and Hugo Quené for supplying us with the PRAAT scripts that we could use for our spectral balance and rhythmic timing analyses.
  - <sup>2</sup> For reasons of clarity, we abstract from constraints such as FOOTBINARITY (FTBIN) and WEIGHT-TO-STRESS PRINCIPLE in Table 2. Although these constraints play an important role in the Dutch stress system (cf. Gilbers & Jansen, 1996), the conflict between OUTPUT-OUTPUT CORRESPONDENCE and FOOT REPULSION is essential for our present analysis.
  - <sup>3</sup> With respect to the phonological analysis of the data, we suggest a random ranking of weighed correspondence and markedness constraints. By means of weighing constraints we adopt an OT variant that more or less resembles the analyses in OT's predecessor Harmonic Grammar (cf. Legendre, G., Y. Miyata & P. Smolensky, 1990). Note that we do not opt for a co-phonology for *allegro*-style speech in our analysis. In a co-phonology, the output of the *andante*-style ranking is input or base for the *allegro*-style ranking. We opt for a random ranking with different preferences for *allegro* and *andante* speech, because our data show variable rhythmic structures at both rates. Both rankings evaluate the same input form.

- 
- <sup>4</sup> The perceived loudness depends on the frequency of the tone. The Phon entity is defined using the 1kHz tone and the decibel scale. A pure sinus tone at any frequency with 100 Phon is as loud as a pure tone with 100 dB at 1kHz (Rietveld and Van Heuven, 1997: 199). We are most sensitive to frequencies around 3kHz. The hearing threshold rapidly rises around the lower and upper frequency limits, which are respectively about 20Hz and 16kHz.

## References

- Boersma, Paul, and David Weenink (1992-2002). PRAAT, phonetics by computer. Available at <http://www.praat.org>, University of Amsterdam.
- Burzio, Luigi (1998). Multiple Correspondence. *Lingua*, 104: 79-109.
- Cooper, W., and J. Eady (1986). Metrical phonology in speech production. *Journal of Memory and Language*, 25: 369-384.
- Couper-Kuhlen, Elizabeth (1993). *English speech rhythm: form and function in everyday verbal interaction*. Benjamins, Amsterdam.
- Cummins, Fred, and Robert Port (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26(2): 145-171.
- Eefting, Wieke, and Toni Rietveld (1989). Just noticeable differences of articulation rate at sentence level. *Speech Communication*, 8: 355-351.
- Gilbers, Dicky, and Wouter Jansen (1996). Klemtoon en ritme in Optimality Theory, deel 1: hoofd-, neven-, samenstellings- en woordgroeps-klemtoon in het Nederlands [Stress and rhythm in Optimality Theory, part 1: primary stress, secondary stress, compound stress and phrasal stress in Dutch]. *TABU*, 26(2): 53-101.
- Gilbers, Dicky, and Maartje Schreuder (to appear). Language and Music in Optimality Theory. *Proceedings of the 7th International Congress on Musical Signification 2001, Imatra, Finland*. Extended manuscript available as ROA-571.
- Grabe, Esther, and Paul Warren (1995). Stress shift: do speakers do it or do listeners hear it? In: Connell, Bruce and Amalia Arvaniti (eds.). *Phonology and phonetic evidence*. Papers in Laboratory Phonology IV.
- Hart, Johan, René Collier, and Antonie Cohen (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge University Press, Cambridge.

- Huss, V. (1978). English word stress in the postnuclear position. *Phonetica*, 35: 86-105.
- Kager, René (1994). *Ternary rhythm in alignment theory*. ROA-35.
- Legendre, Geraldine, Yoshiro Miyata, and Paul Smolensky (1990). Harmonic Grammar - A formal multi-level connectionist theory of linguistic well- formedness: An application. In: *Proceedings of the Twelfth Annual Meeting of the Cognitive Science Society*, 884-891.
- Lerdahl, Fred, and Ray Jackendoff (1983). *A Generative Theory of Tonal Music*. The MIT Press, Cambridge, Massachusetts, London, England.
- Liberman, Mark (1975). *The Intonational System of English*. Garland, New York and London.
- McCarthy, John J. (1986). OCP Effects: Gemination and antigemination. *Linguistic Inquiry*, 17: 207-263.
- Neijt, Anneke, and Wim Zonneveld (1982). Metrische fonologie - De representatie van klemtoon in Nederlandse monomorfematische woorden. [Metrical phonology – The representation of stress in Dutch monomorphemic words] *De nieuwe Taalgids*, 75: 527-547.
- Prince, Alan, and Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms., ROA-537.
- Quené, Hugo, and Robert F. Port (2002). *Rhythmical factors in stress shift*. Paper presented at the 38th Meeting of the Chicago Linguistic Society, Chicago.
- Rietveld, Toni, and Vincent van Heuven (1997). *Algemene Fonetiek. [General Phonetics]*. Dick Coutinho, Bussum.
- Schreuder, Maartje, and Dicky Gilbers (submitted). Restructuring the melodic content of feet. In: *Proceedings of the 9th International Phonology Meeting 2002, Vienna, Austria*.
- Shattuck Hufnagel, Stephanie, Mari Ostendorf, and Ken Ross (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics*, 22: 357-388.
- Sluijter, Agaath (1995). *Phonetic Correlates of Stress and Accent*. HIL dissertations 15, Leiden University.
- Sluijter, Agaath, and Vincent van Heuven (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4): 2471-2485.