Visualizing Levels of Rhythmic Organization

Petra Wagner

Institut für Kommunikationswissenschaften, Universität Bonn
pwa@ifk.uni-bonn.de

ABSTRACT

The paper presents a method to visualize the timing related levels of prosodic organization that have an influence on the rhythmic shape of an utterance. Timing relations can be characteristic of a language or a speaking style. The method is illustrated on various languages classified as stress timed or syllable timed, on a rhythmically unclassified language and L2 speech.

The visualization method can be used to detect rhythmically relevant levels of organization within the prosodic hierarchy, e.g. whether rhythm manifests itself primarily on the level of prosodic feet, phrasal organization or reduction. Our method helps to identify language and speaking style related rhythmical preferences and can classify languages rhythmically. It is able to visualize subtle and large differences between stress timed and syllable timed languages and timing related performance problems of L2 speech.

Keywords: rhythm, prosody, visualization, L2

1. INTRODUCTION

Much research in the last couple of years concentrated on the development and application of global measures differentiating between stress timed, syllable timed and mora timed languages [1, 2]. The rationale behind these measures lies in the relationships between phonotactic syllable complexity, syllable reduction on the one hand and rhythm class on the other [3] expressed in duration properties of consonantal and vocalic intervals.

However, since rhythm is a hierarchical phenomenon [4,5], it manifests itself on different levels of the prosodic hierarchy. E.g., French rhythm has been explained relative to the stress group [6], while in English [7, 8] and German [9], foot internal timing relations, and in Estonian, timing relations between consecutive feet as well as syllables [10] need to be taken into account for a satisfactory description of rhythmical phenomena. Our approach helps to identify the language, dialect or speaking style specific levels of rhythm related timing relations. Therefore, it provides a useful tool for identifying rhythm classes and the relevant analysis levels in rhythm research.

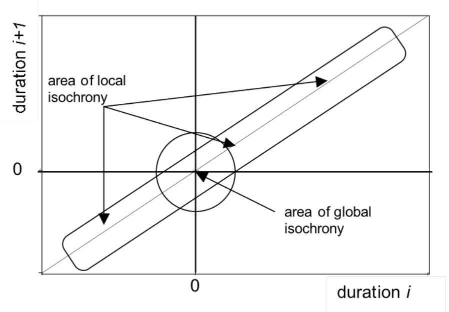
2. VISUALIZATION METHOD

The proposed visualization method is adopted from physical studies of water faucets dripping behavior within the field of applied chaos theory [11]. In order to gain insight into the underlying structures of chaotic behavior, the timing relations between consecutive drips are plotted in a multidimensional graph. We adopt this simple approach in order to find out more about the hidden regularities underlying speech rhythm.

2.1. Visualizing timing relations

The visualization method simply plots duration of consecutive rhythm related events, e.g. syllables or the p-center intervals in a two-dimensional space (cf. Figure 1). The event *i* is plotted on the horizontal axis, the event *i+1* on the vertical axis. In order to compare the timing relations across different speakers, languages etc., the durations are normalized using the z-score method. After normalization, durational averages are “0”. Thus, a data plot around the {0,0}-co-ordinate expresses a sequence of two rhythmical events of average duration. An extension of the visualization into a three-dimensional space is certainly interesting but not investigated here.

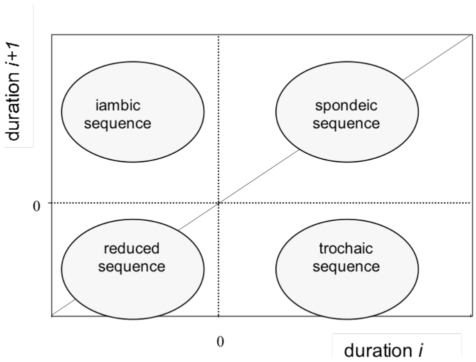
Figure 1: The basic approach towards rhythm visualization: The duration of a rhythm event *i*, is plotted on the x-axis, the duration of a consecutive event *i+1* is plotted on the y-axis.



2.2. Interpreting the visualization

The visualization reveals several rhythm related durational properties of speech. A tendency for global isochrony would manifest itself in a cluster of data plots around the {0,0}-co-ordinate. Local isochrony, as measured as an indicator of rhythm class by the PVI-method [1], would result in data plots around the bisecting line. A sequence of reduced, thus comparatively short syllables, as we would expect it in stress timed languages with a tendency for compensatory shortening, would result in plots in the lower left quadrant of the diagram. Asynchrony, or a maximization of durational difference, which has been suggested as a more important feature of rhythm than isochrony by [12] would result in plots above AND below the bisecting line. More specifically, a strictly alternating sequence would result in one group of data plots in the lower right quadrant (trochaic – long short) and another group of data plots in the upper left quadrant of the diagram (iambic – short long). The identification of local and global isochrony is illustrated in Figure 1, the identification of rhythm related timing relations is illustrated in Figure 2.

Figure 2: Rhythm related timing relations expressed in two-dimensional plots.



3. APPLICATION OF VISUALIZATION METHOD

In this section, the visualization method is applied to several languages, the speech material is taken from a multilingual public domain corpus [13]. The database contains L1 and L2 speech from languages classified as stress timed (German, English), syllable timed (French, Italian) and has been supplemented by additional recordings in Polish, which has been difficult to classify in terms of rhythm class [e.g. 14]. The corpus consists of speech material read at different articulation rates and has been segmented into consonantal and vocalic intervals and syllables. The original database has been augmented with labels for feet, anacrusis and intonation phrase boundaries. Feet were annotated based on lexical stress location in content words. The lexical stress annotation was carried out based on pronunciation dictionaries or based on language specific rules for lexical stress assignment (penultimate syllable in Polish, final syllable in French – though we are aware that the status of French lexical stress is controversial) .

In this paper, the visualization for each language is based on realizations of three speakers reading a text passage in five articulation rates. Exploratory investigations have shown that the patterns emerging during the visualization stabilize after two or at most three speakers (roughly 80 intonation phrases). Of course, the tendencies reported here have been calculated and confirmed using the entire corpus.

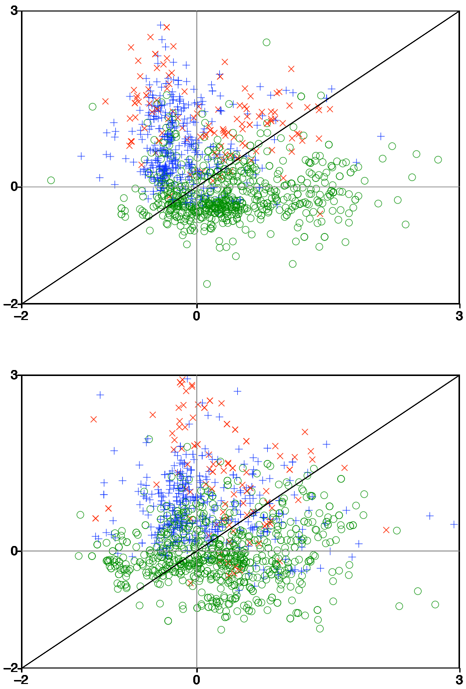
As the basic rhythmical visualization unit, the durations of consecutive syllables were calculated and their durations corrected using P-centers. P-centers were determined according to the simple method proposed in [15], taking into account onset and coda duration. This approach is only slightly less successful than more sophisticated phsychophonetic algorithms [16]. Anacrusis syllables are ignored (for now!) in the analysis, thus only timing relations within “Narrow Rhythm Units” [7] are investigated. In order to detect language specific timing differences of rhythm related units, three different types of event sequences are plotted in different colors and styles:

- *Syllables followed by lexically stressed syllables: Blue/+*
- *Syllables followed by phrase final syllables: Red/x*
- *Syllables followed by unstressed syllables: Green/o*

3.1. Stress timed languages

Two languages which have been previously characterized as stress timed are contained in our database: English and German. English has often been regarded as *the* prototypical stress timed language with a strong tendency to reduce unstressed syllables within a foot. Also, both languages prefer alternating sequences of stressed and unstressed syllables. For both languages, we expect a strict separation between syllables that are followed by a stressed syllable and vice versa which should materialize as a separation by the bisecting line. Furthermore, the trend towards syllable reduction should lead to a relatively large amount of data plots concentrating around the lower left quadrant. Our data plots confirm some expectations, but clear alternation is obeyed more strictly in English than it is in German. Strong reduction effects in sequences of several unstressed syllables, are difficult to trace. Final lengthening effects are evident in both languages, often falling on the final, but sometimes additionally on the penultimate syllable, thus leading to spondeic sequences.

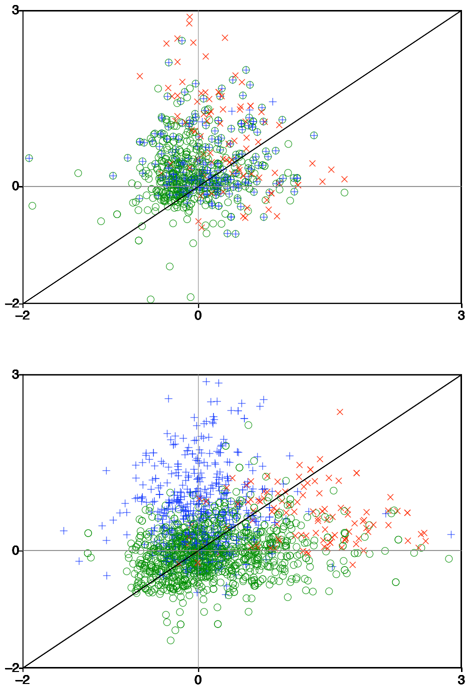
Figure 3: Visualizations of timing relations in English (top) and German (bottom).



3.2. Syllable timed languages

Two language which have been previously characterized as syllable timed are contained in our database: French and Italian. French has often been regarded as *the* prototypical syllable timed language showing a tendency to place stresses at the ends of "stress groups" [6] rather than forming a classical foot. Our visualization method confirms this. For French, we hardly find timing differences between consecutive stressed and unstressed rhythmical events based on the analysis level chosen here. In Italian, there are clear tendencies to differentiate between stressed and unstressed syllables, but a pattern to make unstressed syllables less variable in duration is evident as well. Final lengthening effects are extremely dominant in Italian and clearly affect the penultimate syllable of an intonation phrase. The penultimate often carries lexical and – probably – sentence stress. Final lengthening also plays a role in French, but is less pronounced. The "machine gun" effect attributed to French is expressed by a clustering of data plots around the {0,0}-co-ordinate. In Italian, a similar effect can be traced in sequences of unstressed syllables. Thus, we cannot find strong local isochrony as searched for by the PVI [1], but rather a loose tendency towards global isochrony in both languages, but on different levels of prosodic organization.

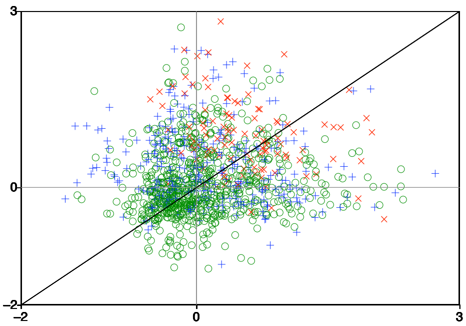
Figure 4: Visualizations of timing relations in French (top) and Italian (bottom).



3.3. Classifying Polish Rhythm

Polish has been difficult to classify in terms of rhythm class. It possesses a high phonotactic complexity, thus qualifying as stress timed, but also shows few reduction, which is more typical for a syllable timed language [3]. According to our visualization, Polish is clearly classified as syllable timed, making even less differences between stressed and unstressed syllables than Italian (cf. Figure 5). In terms of rhythm, it looks organized similarly to French, with a pronounced final lengthening. However, it shows more variation and less isochrony, probably due to production constraints caused by phonotactic complexity.

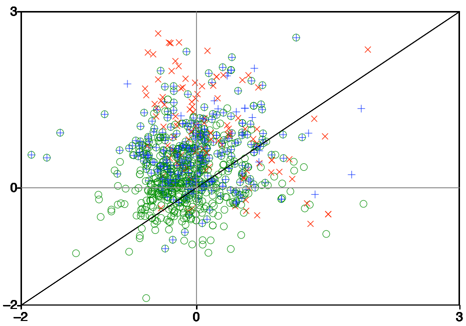
Figure 5: Visualizations of timing relations in Polish.



3.4. Rhythmical L1 interference on L2

Our method can also visualize rhythmical performance problems in L2 rhythm. Figure 6 depicts the timing patterns of three native speakers of German speaking French. Their productions are clearly influenced by their native timing patterns with a tendency to lengthen "lexically stressed" syllables. This results in less overlap between green and blue areas in comparison with native speakers.

Figure 6: Visualizations of L2 French produced by German L1 speakers.



4. CONCLUSIONS

We have presented a novel and straightforward method to visualize rhythm related timing relations in different languages. The method can be used to identify relevant prosodic levels of rhythmic organization and may help tracing rhythm related effects of speaking style, linguistic variation and L2 speech.

5. ACKNOWLEDGEMENTS

Warm thanks go to Maciej Karpinski for providing the translation and transcription of the corpus text into Polish, and to Dafydd Gibbon and Nick Campbell for a lively discussions of my ideas.

6. REFERENCES

[1] Low, E.L, Grabe, E., Nolan, F. 1999. A contrastive study of prosody and lexical stress placement in Singapore English and British English. *Language and Speech* 42, 39–56.
[2] Ramus, F., Nespor, M, Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292.
[3] Dauer, R. 1983. Stress timing and syllable timing reanalyzed. *Journal of Phonetics* 11, 51–62.
[4] Cummins, F., Port, R. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26, 145–171.
[5] Gibbon, D. 2003. Computational modelling of rhythm as alternation, iteration and hierarchy, *Proc. 15th ICPhS*, Barcelona.
[6] di Cristo, A. 1998. Intonation in French. In: Hirst, D., di Cristo, A. (eds.). *Intonation systems: a survey of twenty languages*. Cambridge: CUP, 195–218.
[7] Jassem, W. 1984. Isochrony in English speech: its Statistical Validity and Linguistic Relevance. In: Gibbon, D., Richter, H. (eds.). *Intonation, Accent and Rhythm*, Berlin: Walter de Gruyter, 203–225.
[8] Bouzon, C., Hirst, D. 2004. Isochrony and Prosodic Structure in British English. *Proc. of Speech Prosody 2004*, Nara, Japan.
[9] Gut, U. 2003. Non-native speech rhythm in German. *Proceedings of ICPhS 2003*, Barcelona, Spain.
[10] Asu, E.L., Nolan, F. 2005. Estonian Rhythm and the Pairwise Variability Index. *Proc. FONETIK 2005*, Department of Linguistics, Göteburg University, Sweden, 29–32.
[11] Baier, G. 2001, *Rhythmus – Tanz in Körper und Gehirn*. Reinbek: Rowohlt.
[12] Cauldwell, R. 2002. The functional irrhythmicality of spontaneous speech: A discourse view. *Applied Language Studies*, 2, 1–24.
[13] Dellwo, V., Aschenberner, B., Wagner, P. Dankovičová, J, Steiner, I. 2004, BonnTempo-Corpus and BonnTempo-Tools: A Database for the Study of Speech Rhythm and Rate, *Proc. of INTERSPEECH 2004*, Jeju Island, 777–780.
[14] Ramus, F., Dupoux, E., Mehler, J. 2003. The psychological reality of rhythm classes: Perceptual studies. *Proc. 15th ICPhS 2003*, Barcelona.
[15] Marcus, S.M. 1981. Acoustic determinants of perceptual center (P-center) location. *Perception and Psychophysics*, 30, 247–256.
[16] Janker, P. 1996. *Evidence for the p-center syllable-nucleus-onset correspondence hpothesis*. ZAS Paper in Linguistics, 7, 94–124.