1. Objectives

The proposed project is about *rhythm* – an inherent component of speech. The project concentrates on the *Polish* language, whose rhythm has been relatively little investigated comparing to other languages for example, English, German, or Chinese. There are few corpus-based studies on Polish rhythm (and related phenomena) that use state-of-the-art methodologies (e.g. Demenko 1999, Klessa 2006, Gibbon et al. 2007, Malisz 2011, Wagner A. 2013). Many aspects of the rhythmic structure of Polish have not been investigated at all for example, phonotactic properties of Polish rhythm at the phrase level, realization and perception of rhythm in Polish by non-native speakers, or only in a limited scope for example, the effect of speech rate on rhythm (Malisz 2011), production and perception of prominence (Malisz & Wagner P. 2012) or rhythmic classification of Polish (Ramus et al. 1999, Grabe & Low 2002, Mairano et al. 2011). At present it is generally acknowledged that rhythm should not be equated with *timing* since the role of parameters other than duration like tempo and intonation to the creation of rhythm can be substantial (Arvaniti 2009). However, the research on Polish rhythm has hardly ever gone beyond this dimension. It should also be noted that among recent cross-language studies on rhythm, some include Polish only as an example of a rhythmically complex language and investigation of Polish rhythm is *not* their primary objective (Grabe & Low 2002, Ramus et al. 1999, Wagner P. 2008, Mairano et al. 2011). The proposed project addresses issues that so far have been neglected in the investigation of Polish rhythm or haven't received enough attention and should be subjected to a more in-depth and comprehensive analysis.

The first goal of the proposed project is to provide a multidimensional description of the rhythmic structure of utterances in Polish. This description will incorporate knowledge on phonotactic and phrasal properties of Polish rhythm, relation between rhythm and speech rate, stability of rhythm metrics, contribution of duration (timing), pitch (intonation), intensity and tempo to acoustic realization of meter and grouping, and cross-linguistic perception and production of prominence and phrasing. It will be used to *describe quantitatively and qualitatively speech rhythm* realized by native as well as *non-native speakers* of Polish whose first languages (L1s) have perceptually different rhythms. Creation of a multidimensional description of Polish rhythm requires an in-depth analysis of various aspects of rhythmic organization of utterances.

So far, the research on Polish rhythm has been limited to variation in timing at segmental, syllable or foot (or interstress interval) level (e.g., Jassem 1964, Klessa 2006, Gibbon et al. 2007, Malisz & Klessa 2008). There are corpus-based studies that investigated accentual lengthening and phrase-final lengthening in Polish (Jassem 1964, Demenko 1999, Wagner A. 2008, Wagner A. 2010, Malisz & Wagner P. 2012), but not in the context of rhythmic structure of utterances. Phonotactic properties of Polish speech rhythm have not received much attention – the effect of different materials (syllable-timed vs. stress-timed sentences, e.g. Arvaniti 2012a) on the measurements of rhythm has not been yet subjected to any corpus-based analysis. *The proposed project aims at investigating how phrasal structure and phonotactic properties influence the rhythmic organization of Polish utterances.* The results of the analyses will be compared cross-linguistically in order to learn more about rhythmic differences and similarities between Polish (a non-prototypical, rhythmically unclassified language) and other languages (prototypical stress- and syllable-timed e.g., English and Spanish, and unclassified e.g., Catalan, Czech). For the first time in research on Polish rhythm, *relationship between intended and laboratory measured tempo, average speech rate of Polish speakers and the way in which they increase the speech rate will be investigated.* As shown in Dellwo & Wagner (2003) and Dellwo (2008) these aspects are language specific, thus the knowledge of them should constitute part of description of Polish rhythm. The proposed project will complement and extend the scope of the recent tempo differentiated analyses of Polish (Malisz 2011), using specially designed speech corpus representative of five levels of tempo and including various phonotactic, prosodic and rhythmic structures.

The second goal of the proposed project is to provide rhythmic classification of Polish. Traditionally speech and language rhythm have been described with reference to *rhythm class hypothesis* which provides a framework of description and analysis of the rhythmic structure in terms of *stress-timing* (e.g., in English and German), *syllable-timing* (in French and Italian) and *mora-timing* (in Japanese). The rhythmic status of Polish is unclear and different studies provide different classifications. The objective of the proposed project is *to learn about rhythmic characteristics of Polish* – a rhythmically non-prototypical language, and *to compare them cross-linguistically to better understand similarities and differences with other languages.* This comparison will be based partly on the multidimensional description of rhythm created in the proposed project which will serve as a framework for analysis of rhythmic characteristics of other languages. On the basis of the state-of-the-art knowledge it is hypothesized that in order to provide rhythmic classification of Polish it is necessary to go beyond the dimension of timing and to analyze how not only duration, but also pitch, intensity and tempo contribute to production and perception of meter and grouping. The knowledge about rhythmic classification of Polish has important implications for future research in other areas for example, language acquisition, teaching and learning Polish as a foreign language, or dealing with speech impairments (such as

dysphasia or dyslexia) and has many potential applications in speech technology for example, in language identification, speaker characterization, speech recognition and synthesis or computer-assisted pronunciation training. The results of analyses aiming at rhythmic classification of Polish will also provide evidence in the debate between *categorical or gradient (continuous) rhythmic distinction* across languages.

Significant part of recent research on rhythmic organization of utterances in general and rhythmic classification of languages in particular has been carried out using *rhythm metrics* – formulas that measure durational variability in syllables and vocalic and consonantal intervals (e.g., Ramus et al. 1999, Grabe & Low 2002, Dellwo 2006). Results of quantification of speech rhythm in Polish using rhythm metrics are inconclusive and contradictory as regards its rhythmic classification. There exists experimental evidence that rhythm metrics are unstable within language (accent) groups partly due to variation in material (e.g., differences in phonotactic and prosodic structure), measurements, speech styles and tempos (Wiget et al. 2010, Arvaniti 2012a). However, results of a recent study (Prieto et al. 2012) indicate that rhythm metrics can be successfully applied to rhythmic discrimination of languages, because despite variation in phonotactic structure of materials (typical stress- vs. syllable-timed utterances) they showed expected patterns of variation. *The proposed project aims at learning about the effects of possible factors of variation on rhythmic measurements on the basis of Polish speech material.* As a result, the most stable metrics will be selected for use in further analyses planned in the project – it can be hypothesized that vocalic metrics will show greater stability comparing to consonantal metrics. The patterns of variation in metric scores may also bring important information concerning rhythmic characteristics of Polish as compared to other languages which have already been subjected to such analyses. Apart from that, the analyses will verify the hypothesis that even though rhythm metrics can not be regarded as acoustic correlates of perceived rhythm, they still can be useful to describe timing patterns in speech.

The third goal of the proposed project is to learn about production and perception of prominence and phrasing. Prominence and phrasing are aspects of spoken utterances closely related to the notions of *meter* and *grouping* which are considered as basis of perceived rhythm (Arvaniti 2009, Wagner P. 2008). Corpus-based analyses of perceptually annotated data representative of *different accents* (native and non-native speakers with different L1s), phonotactic, prosodic and rhythmic structures, and speech tempos, will aim at *identification of the contribution of duration, pitch, intensity and tempo to realization of different levels of prominence* (four levels will be distinguished) and *phrasing* (two levels will be taken into account). The proposed project will also explore perception of prominence and phrasing by native and *non-native speakers* of Polish and provide an in-depth knowledge about universal and language-specific aspects of this percept. It can be hypothesized that in the production study differences in the phonetic implementation of variation in pitch, duration and intensity related to various prominence and prosodic boundary levels will be observed between Polish native and non-native speakers and between different non-native accents. In the perception study the effect of bottom-up and top-down factors on the identification of prominence and prosodic phrasing structure is expected (Ericksson et al. 2002, Wagner P. 2005). The results will be used to characterize Polish rhythm cross-linguistically and to develop a multidimensional description of Polish rhythm.

1. Significance

2.1 State of the art

Rhythmic classification of languages

Speech rhythm can be defined as a *systematic temporal organization of prominent and less prominent speech units*. In linguistics and phonetics, the notion of speech rhythm is traditionally related to the notion of *isochrony* – the recurrence of some type of speech unit, such as syllable or foot, in equal or near-equal time intervals that establishes temporal organization of languages (Pike 1946, Abercrombie 1967). On the psychological level rhythm is defined as *grouping of elements into larger units* which need to have some similarity and have to be marked off from each other (Woodrow 1951, Fraisse 1964 after Dauer 1987). The concept of isochrony underlies *rhythm class hypothesis* according to which every language represents a specific rhythm class on the basis of its temporal organization of syllables. Two most prominent rhythm classes include syllable-timed (Spanish, Italian or French) and stress-timed rhythm class (English, German or Dutch). There is also a class of mora-timed languages exemplified by Japanese. However, empirical studies failed to show evidence for *isochrony*, the equal duration of feet, syllables and morae in stress-, syllable-and mora-timed languages respectively. The evidence supporting rhythm class hypothesis came from perception studies which showed that humans (both neonates and adults) and some animals can perceptually distinguish between languages on the basis of their rhythm (Nazzi et al. 1998, Ramus et al. 2003, for an overview see Dellwo 2010). The fact that instrumental studies failed to bring evidence for stress- and syllable-based isochrony caused that in rhythm research the focus moved from duration measurements of syllables and feet to investigation of *phonetic* and *phonological* factors which affect the timing of syllables and feet, most importantly degree of vowel reduction in unstressed syllables and variety of syllable structures and complexity of onsets and codas allowable in a language. In syllable-timed languages there are generally no reduced vowels and simple syllables (CV or

CVC). On the contrary, in stress-timed languages syllables can have quite a complex structure with more than one segment in the onset and coda at a time (mostly CVC, but also CCVC, CVCC, etc.) and vowels undergo reduction processes (Dauer 1983, Dellwo 2010). These observations, supported by results of empirical studies, led to a redefinition of rhythm as the perceptual effect of interaction of a number of components: phonetic and phonological on the one hand, and segmental and prosodic on the other (Dauer 1987).

Rhythm metrics

The differences in vowel reduction, stress-based lengthening and syllable complexity between stress- and syllable-timed languages became the basis of rhythm metrics – formulas that measure durational variability in consonantal (C) and vocalic (V) intervals (or syllables). The intervals are derived from phonetic segmentation of speech signals and no reference is made to their phonological constituency. Basically, a vocalic interval is a stretch of a speech signal consisting of one or more vowels (or vocalic segments like diphthongs, triphthongs, etc.) preceded and followed by a consonant (or by a pause). A consonantal interval (C-interval) is a stretch of a speech signal consisting of one or more consonants preceded and followed by a vowel (or by a pause). The main application of rhythm metrics has been to rhythmic classification of languages. As demonstrated in various studies (Ramus et al. 1999, Grabe & Low 2002, Mok & Dellwo 2008, Mairano & Romano 2011) on the acoustic level purely phonetic measures of rhythm give the best discrimination between speech rhythm classes. The rationale behind using them was that in speech perception and language discrimination infants do not rely on abstract phonological concepts of syllable and stress for which there are no general phonetic definitions. Moreover, such measures correlate the best with listeners' perception of rhythm class (Ramus et al. 1999, White et al. 2007). The most widely used metrics include %V-ΔC (Ramus et al. 1999), PVIs (Grabe & Lee 2002) and Varcos (Dellwo 2006). Languages with the stress-timed rhythm are expected to exhibit higher nPVI and lower %V and higher consonantal rPVI and ΔC that result from vowel reduction and quite complex syllable structure respectively. On the contrary, syllable-timed languages generally do not have spectrally reduced and shortened vowels and most syllables have a simple CV structure. Consequently, these languages are expected to exhibit higher %V and lower values of nPVI and the consonantal metrics. Generally, Polish exhibits phonological properties that are associated with both types of rhythm class (Jassem 1962, Richter 1987, Demenko 1999, Klessa 2006, Wagner 2008, Malisz & Klessa 2010, Malisz 2013). For example, increase in duration on prominent syllables and some compensatory shortening effects together with phonotactic complexity of syllables indicate a tendency towards stress timing. On the contrary, the lack of vowel reduction and a fixed lexical stress on the penultimate syllables indicate a tendency towards syllable timing. Therefore, having some features distinctive of stress-timed languages and others distinctive of syllable-timed languages, Polish would be predicted to behave rhythmically as an *intermediate* or *mixed* language. Yet the acoustic evidence in this respect is somewhat contradictory: According to PVIs (Grabe & Low 2002), Polish is close to languages classified as syllable-timed, but, according to %V – ΔC, it is grouped with stress-timed English and Dutch (Ramus et al. 1999).

There is a growing body of evidence from experimental studies showing that rhythm metrics are not perfect. First of all, they are sensitive to differences in text materials, speech elicitation methods, measurements and changes in speech rate (Arvaniti 2012a, Wiget et al. 2010). This sensitivity explains why rhythm metrics take different values in different studies. Second major problem is that metrics reduce rhythm to a single dimension, namely that of timing, whereas apart from duration, perceived speech rhythm is also the product of properties such as melodic (F0) change and, to a lesser extent, intensity and vowel quality. As a result, different measures do not necessarily reflect perceptually different rhythms (Barry et al. 2009). Even though rhythm metrics have been criticized for equating rhythm with timing and reducing rhythm to the effect of phonotactic and phonological properties of the language, they constitute the methodological foundation of many recent studies concerning topics as diverse as discrimination between rhythm classes (e.g. Mairano & Romano 2011), acquisition of timing patterns in native (Payne et al. 2012) and foreign language (White & Mattys 2007, Mok & Dellwo 2008, Tortel & Hirst 2010, Ordin et al. 2011, Kinoshita & Sheppard 2011, Li & Post 2012, Wagner A. 2013), detection of speech impairments (Liss et al. 2009) or dialect discrimination (Leeman et al. 2012).

Relation between speech rhythm and speech rate

The relationship between rhythmic structure and speech rate has been the subject of extensive research for many languages (Garding 1975, den Os 1985, Garding & Zhang 1997, Dellwo 2010 among many others), but not for Polish. The exception is recent study by Malisz (2011) who investigated the effect of tempo differentiation on rhythmic measurements (based on PVIs and coupled oscillator model). The results indicated that with increasing tempo rhythmic characteristics of Polish tempo-differentiated utterances stay in the syllable-timed range, but at the same time change slowly from syllable-timed to more stress-timed (though the effect is generally weak). The author concludes that "Polish remains quite inflexible in the adjustments of its structure to speech rate".

Dellwo & Wagner (2003) found that the number of syllables that speakers are able to produce per second on an average basis is language-specific, and that this value will be highly influenced by the language individual phonetic, phonological and phonotactic syllable structures. Language-specific is also the way in which speakers increase the syllable rate. In this respect syllable-timed French seems to provide the greatest freedom, speakers of stress-timed English seem to be most restricted in syllable rate increase, while German (also stress-timed language) lies somewhere in between these two extremes.

Recently, influences of speech rate on rhythm metrics were extensively analyzed by Dellwo (2008). He used BonnTempo corpus (Dellwo et al. 2004) that includes tempo differentiated speech data from five languages: Czech, English, French, German, and Italian. The analyses revealed that for the languages under investigation, %V and the normalized rPVI are least affected by speech rate variability in all languages and provide the best discrimination between rhythm classes. Another finding of his analyses was that different languages react differently to changes in speech rate and attributed it to different amounts of consonantal complexity e.g., it was demonstrated for the reading material for BonnTempo that French and Italian (syllable-timed languages) possess a lower number of complex intervals than English and German (both stress-timed) with Czech somewhere in the middle between these two groups. Languages containing a lower complexity have been shown to be lower in the consonantal rhythm measures and higher in CV-rate. This observation gives reason for an assumption that languages with shorter, less complex intervals also show less variability of C-intervals over different speech rates. Dellwo (2008) also demonstrated that languages vary in the range of speech rate differences they allow. In this context it was found that English allows less speech rate differences than the other languages under investigation. If languages allow less overall changes in speech rate, the variability of C-intervals can be expected to be less affected.

Phonotactic and phrasal properties of rhythm

Investigations on linguistic rhythm have commonly taken the view that phonological and phonetic properties of utterances form the basis of perceived rhythmic differences between languages. However, it is also well known that higher levels of prosodic structure such as prominence marking and prosodic phrasing strongly influence the organization of timing across languages. Analyses based on various languages (including Polish) showed that stressed and pitch accented syllables are produced with additional lengthening compared with unstressed syllables (e.g., Jassem 1962, Turk and White 1999, Demenko 1999, Wagner A. 2008, among many others). Final lengthening at the edges of phrasal prosodic constituents is also very widespread (Wightman et al. 1992, Demenko 1999, Wagner A. 2008, among many others). Therefore, it has been acknowledged by many authors (e.g., Beckman 1992, Fant et al. 1991, Arvaniti 2009, Prieto et al. 2012) that the perception of rhythm classes should be examined in relation to prosodic timing phenomena – the durational marking of prominences and phrase boundaries, because they are correlated with the potential rhythm class distinctions and constitute an important ingredient in the rhythm percept across languages.

The results of a recent study (Prieto et al. 2012) showed that differences in the rhythm metrics emerge even when syllable structure is controlled for in the experimental materials, at least between English on the one hand and Spanish/Catalan on the other, suggesting that important differences in durational patterns exist between these languages that cannot simply be attributed to differences in phonotactic properties. Further analyses of the data indicate that the rhythmic class distinctions correlate with differences between the languages in the phonetic implementation of variation in timing related to prominences and phrasing. The results imply that even though phonological properties of the target language such as syllable structure and vowel reduction can predict rhythmic behavior to some extent, the percept of rhythm is relatively independent of them and consistent and language-specific timing patterns occur despite phonotactic differences between languages. As regards Polish, these relationship and effects have not been subject to investigation yet.

Production and perception of prominence and phrasing

At present it is generally acknowledged that *rhythm* should not be equated with *timing*. Listeners extract rhythmic pattern of an utterance by classifying various prosodic units (syllables, feet, phrases) as more or less prominent, on the basis of durational (and other) clues – languages use different parameters to make some syllables more prominent than others, and the same parameter may be allocated different degrees of importance in different languages. Focusing exclusively on durational measurements is misguided (Arvaniti 2009) and the contribution of parameters other than duration to the creation of rhythm should be examined, since their role could be substantial. A perceptual experiment with Bulgarian, English and German subjects (Barry et al. 2009) indicated that the perceived strength of rhythmicity in a line of verse is determined not only by its temporal structure, but also by other acoustic properties, most clearly by F0 change within the metrical foot. Within the production study, nearly all of the variation in 'rhythmic strength' was reflected in the durationally based rhythm measure, but there was evidence from some of the variation that duration and other properties – F0 at least – were complementing one another. These results indicate that perceived rhythm is the effect of *meter*

and *grouping* and therefore research on rhythm should incorporate investigation of prominence (which functions as a *meter*) and phrasing (which is related to *grouping*).

Duration, pitch, intensity and spectral balance are considered the main acoustic correlates of prominence (Terken 1991, Hirschberg 1993, Portele 1998, Tamburini 2003, Sridhar et al 2008, among many others). The realization of prominence is language specific and languages differ in the phonetic implementation of variation in the acoustic parameters. As regards stress-based prominence in Polish, the first accounts indicated the main correlation with overall intensity (Dłuska 1950). On the contrary, Jassem (1962) found that pitch movement, not intensity, is the most salient correlate of lexical stress. Recent studies (Demenko 1999, Wagner A. 2009, Malisz & Wagner P. 2012) indicate that increased duration is just *one of the acoustic correlates* of prominence (next to intensity and pitch) and that durational variation is much more subtle in case of *minor prominences* (associated with stress whose pattern is mostly linguistically determined) than *major prominences* (pitch accents and nuclear accents). The effect of stress on vowel duration is generally weaker than in Germanic languages. Phrase boundaries in Polish are signaled most of all by an increased duration of the phrase-final syllable and its nucleus, and nucleus of the previous syllable (usually a prominent one). F0 features such as tilt, slope (describing shape and variability in pitch on the vowel of the word final and penultimate syllables) and F0 mean play also a significant role (Demenko 1999, Wagner A. 2010). These observations are in line with results from other languages (Carlson & Swerts 2003, Aguilar et al. 2009, among many others) which showed that both duration and F0 variation constitute important cues to phrasing, although boundary presence had greater effect on duration than on F0.

As regards perception of prominence and phrasing, bottom-up and top-down factors were found. On the one hand it was shown that listeners attach different weights to specific acoustic cues (e.g., Ericksson et al. 2002, Rosenberg & Hirschberg 2010), but on the other hand, listeners perception is influenced by their expectancies (mostly shaped by linguistic structures) on the position of prominent syllable (Wagner P. 2005).

2.2 Justification for tackling specific scientific problems by the proposed project

Research on speech rhythm is highly significant, because rhythmic patterning simplifies cognitive and linguistic processes – *rhythmic structure helps us parse, interpret and memorize the content of an utterance*. Rhythmical structuring can extend the capacities of short term memory to a certain extent (Wagner P. 2008, p. 4): "rhythm provides us with an archiving structure minimizing cognitive memory load. Since it is clear that it helps to convey a linguistic message if listeners remember what has been said, a clear rhythmic structure automatically aids human communication." The most important linguistic functions of rhythm include:

- *Grouping* or *structuring*: "Speech rhythm functions mainly to organize the information bearing elements of the utterance into a coherent package, thus permitting speech communication to proceed efficiently." (Allen 1975, after Wagner P., p. 5)
- *interpretation* – it helps segmental and lexical perception as it provides cues concerning the prominent and important stretches of time in course of an utterance
- *parsing* – it provides us with information concerning important linguistic boundaries (e.g. word segmentation in fixed-stress languages like Polish)
- it signals *semantic-pragmatic structure* (e.g., position of focus, "spoken parentheses")

On the other hand, an impairment of rhythm related processes negatively affects linguistic performance e.g., stuttering, dysphasia, dyslexia or difficulties in the language acquisition process. Considering the number of functions that rhythm has in a language and its communicative and cognitive significance it is evident that there is need of more in-depth knowledge on rhythm, so that real progress in areas as diverse as linguistics, phonetics, speech technology, language acquisition, speech and language pathology, is possible.

Generally, the issues addressed by the proposed project are specific to the Polish language for example, rhythmic characterization/classification of Polish, perception and production of prominence and phrasing by native and non-native listeners/speakers of Polish, temporal patterning in tempo differentiated utterances, but the project tackles also universal and language-independent problems for example, the existence of rhythm classes or stability, the framework of speech rhythm description that goes beyond the dimensions of timing, or quantification of rhythm using rhythm metrics. As regards research on speech rhythm in Polish, the proposed project focuses on issues and problems that so far have not been investigated at all, or have received only limited attention. Comparing to other languages, the state of the art knowledge of speech rhythm in Polish needs to be complemented by results from corpus-based analyses using up-to-date methodology. The issues and problems addressed by the proposed project are of a key significance for general understanding of rhythm and for methodological assumptions of future studies on rhythm.

2.3 Pioneering nature of the project

In research on speech rhythm in Polish corpus-based studies using up-to-date methodology are scarce and most of the time limited to the level of segments, syllables, feet or interstress intervals (Klessa 2006, Gibbon et al.

2007, Malisz 2013). The pioneering nature of the proposed project consists in applying a *large and specially designed speech corpus and up-to-date methodology to comprehensive, multidimensional analyses of rhythmic structure of utterances in the Polish language.* The analyses will be carried out *at the phrase level.*

Until now, the analyses of Polish rhythm have been based on measurements carried out on a very restricted speech material for example, including few speakers, with no differentiation in phonotactic/prosodic/rhythmic features, speech rate and speech style. In order to characterize Polish rhythmically extensive analyses will be carried out on *differentiated* speech data (in terms of materials, elicitation methods, prosody, and rhythmic and phonotactic structures, sec. 3.1.1) coming from *20 native and 15 non-native speakers* (ca. 5.5 hours of speech). Moreover, unlike previous studies, the analyses in the proposed project will *go beyond the dimension of timing* and will investigate also the role of pitch (intonation), intensity and tempo in the realization *meter* (related to prominence) and *grouping* (related to phrasing) which constitute basis of perceived rhythm.

One of the results of the proposed project will be a *multidimensional description of speech rhythm in Polish*. It will be a *phrase-level* description and it will incorporate knowledge about relation between timing and phonotactic and phrasal structure of utterances, relation between rhythm and speech rate, contribution of duration, pitch (intonation), intensity and tempo to acoustic realization of meter and grouping, and cross-linguistic production and perception of prominence and phrasing. The multidimensional description will serve as a framework of comparison between rhythm in Polish and in other languages (e.g., Dellwo 2008, Beňuš & Šimko 2012) for needs of rhythmic classification of Polish. For the first time in research on Polish rhythm, apart from material from native Polish speakers, speech data from *non-native speakers of Polish* (with different L1s) will also be used. By including differentiated speech material, the proposed project will provide a more in-depth analysis of rhythm transfer from speakers' L1 to the target language Polish (cf. Wagner A. 2013) and will identify sources of differences in the rhythmic structuring between accent groups. The results of this research will have serious implications for learners of Polish as a foreign language.

The proposed project will be the first attempt to explore the role of *syllable structure complexity and phrasal timing phenomena* in the production of rhythmic structure in Polish native and non-native speech, and their effect on rhythmic measures (PVIs, delta C, delta V, %V, Varco C, Varco V). For the first time the stability of rhythm metrics which constitute the methodological basis of many recent studies on rhythm will be investigated using speech data in the Polish language.

The knowledge on the *relation between speech rate and rhythm in Polish* is very scarce (e.g., Malisz 2011). Pioneering study based on Polish speech data varying in speech rate *and* rhythmic and phonotactic structure will answer the question whether under tempo differentiation Polish behaves more like stress-timed German or more like syllable-timed Italian. For the first time in research on Polish rhythm the relationship between intended and laboratory measured tempo, average speech rate and patterns of tempo differentiation that are typical of Polish utterances will be investigated. The pioneering character of the analyses will also consist in including phonotactically, prosodically and rhythmically differentiated speech material (see also sec. 3.1.1).

Another issue addressed by the project that is novel in the research on Polish rhythm (and prosody in general) concerns *cross-linguistic production and perception of prominence and phrasing*. The analyses will be based on different types of speech material and variable speech tempos. For the first time *non-native speakers* with different accents (L1s) will be included in the study.

2.4 The impact of the project results on the development of the research field and scientific discipline

The results of rhythmic classification of Polish and cross-linguistic characterization of Polish speech rhythm will have important implications for the general understanding of speech rhythm which still predominantly relies on the notion of rhythmic categories and on the *equation of rhythm with timing*. There exists experimental evidence indicating that the notion of syllable-timing is psychologically implausible, because "subjects presented with series of identical stimuli tend to impose rhythmic structure on them, by hearing some stimuli as more prominent than others" (Arvaniti 2009, p. 58). Consequently, one difference between languages called stress-timed and those called syllable-timed may have to do with the *spacing of prominences*, not in terms of duration but in terms of *number of syllables*; in this respect, prominences may be sparser in syllable-timed languages. If so, then it can be assumed that *stress* is the basis of rhythm in all languages, and therefore the rhythm of language and speech should be represented on a *rhythmic continuum* that stretches from least to most stress-based (Dauer 1983). Recently, the idea of "coexisting rhythms in language" was formulated (Nolan & Asu 2009) which claims that syllable-timing and stress-timing can be independent dimensions of temporal organization in languages rather than opposite ends of a continuum. A corpus based, multidimensional analysis of Polish rhythm is expected to answer the question whether Polish groups closer with stress-timed or syllable-timed languages, or whether it shows a behavior "in between" the endpoints of the continuum. If no rhythmic classification can be achieved (in terms of stress-timing, syllable-timing or intermediate rhythm), it will indicate the need of considering the idea of *rhythmic continuum* and/or *coexisting rhythms*. The outcome of the debate is highly significant, because "given the importance attributed to rhythm, it is essential to

accumulate as much evidence as possible about the accuracy of the rhythm typology to which the function of rhythm in processing and acquisition is crucially linked" (Arvaniti 2012b, p. 76).

The issue of *rhythm metrics stability* is of a key significance for future research on rhythm. Despite all criticism rhythm metrics still constitute methodological basis of studies (White & Mattys 2007, Mok & Dellwo 2008, Tortel & Hirst 2010, Kinoshita & Sheppard 2011, Li & Post 2012, Leeman et al. 2012). Since application of rhythm metrics in future research seems likely, it is very important to learn more about factors that affect their values. The proposed project will address this issue and will aim at identification of new factors as well as combined effects of factors whose influence on rhythm metrics is already known. This resulting knowledge will help researchers design experiments and interpret rhythmic measurements. Moreover, the analyses will also test the hypothesis that rhythm metrics can be successfully applied to measure variability in timing patterns.

The research on the *relation between rhythm and speech rate* is highly important for the results of the ongoing debate on the existence of speech rhythm classes. In a cross-linguistic study Dellwo (2008) found that languages have characteristic speech rates and characteristic patterns of tempo differentiation (see sec. 2.1). He hypothesized that perceived differences between languages, which hitherto have been attributed to different rhythms (like stress- or syllable-timing), are the effect of differences in language-specific speech rates and language-specific patterns of tempo differentiation. The proposed project will contribute to the debate about the hypothesis by providing experimental evidence from Polish.

2.5 Economic and societal impact

Nowadays it is generally acknowledged that having control of "the music of a language" is extremely important to being understood and socially accepted, however stress, rhythm and intonation have rarely been taught in any conscious or systematic way. Since rhythmic structure helps us parse, interpret and memorize the content of an utterance, rhythmic classification/characterization of Polish has important implications for language teachers and learners. It can be expected that linguistic and cognitive processing will be simplified if the source (L1) and target (L2) language share features that affect rhythmic organization of utterances e.g., native speakers of Polish have less difficulty in learning English than Korean which is rhythmically very different from both Polish and English. Incorporation of activities oriented at increasing rhythm awareness results in improving learners' abilities in language acquisition (Holliman et al. 2012) and has become a very popular practice in the language classroom. Training in this field may also be beneficial for persons who aim at improving communicative skills in their L1 like professional speakers, lecturers or actors. Apart from the foreign language teaching and learning context, the knowledge about speech rhythm has important implications for other research areas such as language acquisition or dealing with speech impairments (e.g., dysphasia or dyslexia) and has many potential applications in speech technology (e.g., language identification, speech recognition and synthesis and computer-assisted pronunciation training).

3. Methodology

3.1 Underlying scientific methodology

3.1.1 Speech corpus

20 native speakers of Polish (10 man and 10 female) will be recruited from the student population of the Institute of Linguistics at the Faculty of Modern Languages and Literature (AMU). As shown in a recent study (Klessa 2012) despite small regional differentiation of Polish some variability in the timing patterns between speakers from Western (->Cracow-Poznan pronunciation) and Eastern Poland (Warsaw pronunciation, see e.g., Ostaszewska & Tambor 2000) could be observed. In order to keep dialectal variation to a minimum, only speakers from the region of the Greater Poland (presenting Cracow-Poznan pronunciation) will be recorded. The non-native speakers will be recruited from among participants of Polish Phonetics and Experimental Phonetics courses in the Institute of Linguistics (AMU) and participants of The School of Polish Language and Culture for Foreign Students (AMU). Native languages of non-native speakers should represent perceptually different rhythms: stress-timed (e.g., German, English), syllable-timed (e.g., Spanish, Italian) and rhythmically unclassified (e.g., Korean).

Speech material will represent different elicitation methods and phonotactic, prosodic and rhythmic structures. It will consist of five parts, whereof four will include read speech and one will include spontaneous speech.

1) *Running speech*: reading of a short literary story for use in analyses related to tasks 7 and 9-13.
2) *Set of sentences* representing different phonotactic structures: syllable-timed, stress-timed and uncontrolled sentences. This part of the corpus will be used in analyses related to tasks 7, 8, 9 and 12.
3) *Mini-dialogues* containing different intonation and rhythmic patterns for use in analyses related to tasks 10-12.
4) *Poetic verses* presenting a different rhythmic structure. The main application of this part of the corpus is related to goal 10, but also goals 7, 8, 9 and 13.

5) *Semi-spontaneous speech* for use in analyses related to tasks 9, 12 and 13.

For parts 1, 2 and 3 of the corpus both native and non-native speakers will be recorded. Parts 1, 2 and 4 realized by native speakers will be recorded at five different tempos (normal, slow, very slow, fast and very fast).

It can be expected that approximately 15 minutes of speech per native Polish speaker will be obtained and approximately 2 minutes of speech per non-native Polish speaker. Total duration of the corpus will be approximately 5 hours and 30 minutes. The recordings will be carried out in a professional recording studio of the Department of Psycholinguistics at the Faculty of Modern Languages and Literature (AMU). The resulting speech corpus will be stored on server and will be accessed via specially designed speech database.

3.1.2 Annotation

All annotations will be carried out using two tools: *Praat* (Boersma & Weenink 2013) and *Annotation Pro* (Klessa et al. 2013) that will be integrated with the speech database. At first, all the speech material will be automatically segmented at phoneme, syllable and word level and transcribed in a broad phonetic transcription and position of lexical stress will be automatically marked (Demenko et al. 2003). The results will be manually verified by two expert phoneticians. It can be expected that the non-native speech material will require more corrections. Secondly, on the basis of segmentations and transcriptions boundaries of vocalic and consonantal intervals will be determined (partly automatically) using criteria presented by Arvaniti (2012a). Next, suprasegmental annotation will be carried out. It will consists in marking four levels of prominence (unstressed, stressed unaccented, pitch accented and nuclear accented, e.g., Selkirk 1995) and position of minor and major phrase boundary (Wagner A. 2008). The suprasegmental annotations will be performed at syllable level (each syllable will be marked using prominence or/and phrasing labels) by five phoneticians as well as non-native labelers. Polish labelers will analyze perceptually speech material realized by 20 native Polish speakers (parts 1, 3 and 4 of the corpus) and by 15 non-native speakers (parts 1 and 3). The same 15 persons that provided recordings of non-native Polish will be engaged in the perceptual annotation of prominence and phrasing in the utterances realized by five selected native Polish speakers. Contrary to Polish labelers who will perform the annotation at all speech rates, non-native labelers will analyze only the material realized at normal pace. The results of segmentation and annotation will be automatically saved in text or database files.

3.1.3 Measurements

Measurements will be carried out in *Praat* and *Annotation Pro* using scripts and specially developed plug-ins integrated with the speech database. The results will be saved in text or database files.

a) *Rhythm metrics:* for each utterance and speaker six metrics scores will be calculated: %V – proportion of vocalic intervals, ΔC – standard deviation of consonantal intervals, nPVI-rPVI – vocalic rate-normalized Pairwise Variability Index and consonantal raw Variability Index and VarcoV and VarcoC – coefficients of variation of V- and C-interval durations (formulas are given in Dellwo 2008).

b) *Speech rate:* it will be calculated in syllables/second (e.g., Dellwo & Wagner 2003) and as the sum of the number of C- and V-intervals per second (e.g., Dellwo 2008, Beňuš & Šimko 2012). Calculations will be carried out on the parts 1, 2 and 4 of the corpus (only material in native Polish).

c) *Prosody:* measurements in duration, pitch and intensity domains will be carried out. In each domain mean, maximum, minimum and standard deviation from the mean will be measured for vowels and syllables. In the pitch and intensity domain parameters describing the shape of the contour will be also extracted (e.g., Wagner A. 2008). The measurements will be normalized using the z-score normalization method. Other potentially useful parameter is spectral emphasis (e.g., Portele 1998, Tamburini 2003). Prosodic measurements will be carried out on the speech material realized at five different tempos by native speakers (parts 1, 3 and 4 of the corpus) and at normal tempo by non-native speakers (parts 1 and 3 of the corpus).

3.1.4 Statistical analyses

Data needed for specific experiments will be exported to spreadsheets. Statistical analyses will be carried out in *Statistica 10*. The most important tests & methods include: one-way and multivariate ANOVAs, regression, post-hoc comparisons, correlations, Euclidean distances and generalized linear mixed models (GLMM).

One-way and multivariate ANOVAs will be applied to test whether observed differences between the means of the parameters (e.g., rhythm metrics, pitch or intensity measures) used in a specific experimental design are significant. After obtaining a statistically significant F test from ANOVA, post-hoc comparisons (Tukey HSD or Fishers LSD tests) will be carried out to examine which means contributed to the effect (see e.g., White & Mattys 2007, Dellwo 2008, Rosenberg & Hirschberg 2010, Beňuš & Šimko 2012). In some of the experimental designs, generalized linear mixed models will be used instead of ANOVA, because they have the ability to account for random subject and item effects in one step of analysis and offer considerable advantages over repeated measures ANOVAs (e.g., Prieto et al. 2012, Wagner & Malisz 2012). Correlation coefficients (Pearson's r) will be used to find consistent relationships between analyzed parameters and to avoid data

redundancy. Euclidean distances will be applied to investigate the actual distances between utterances (representing different elicitation methods, prosodic, rhythmic and phonotactic structures, tempos) and accents (native vs. different non-native accents) in rhythm space described by rhythm metrics (e.g., Arvaniti 2012a).

3.2 Type and degree of access to the equipment to be used in the proposed research

Part of the equipment necessary to accomplish project tasks and goals is available: the recording studio (Department of Psycholinguistics), software for segmentation and annotation of the speech corpus (*Annotation Pro*) and for acoustic measurements (*Praat*), and data mining software *Statistica 10* (Department of Phonetics). The equipment that needs to be purchased includes: server (needed to store the speech database), a notebook (needed to carry out recordings of speech material (task 4) and for storing the recorded material for the time of the recordings), two stationary computers (to be used for annotations of the speech corpus) and computer accessories like LCD monitors, mice and keyboards (part of workstations of annotators).

4. Work plan

4.1 Outline of the work plan

Task no.	Task name and description	Expected duration (in months)	Expected start of the task (month of the project)
1	**Creation of specifications for speech corpus for analyses of rhythmic structure of utterances in Polish.** Specifications will concern: a) *speakers* – number, (standard) accent, no speech or hearing disorders, basic knowledge in phonetics and first language in case of non-natives, b) *materials* – linguistic structure of sentences and mini-dialogues, number of sentences/mini-dialogues/verses, selection of verse and text (for continuous reading), elicitation of semi-spontaneous speech, and c) *recording procedure* – recording conditions and software, audio quality of the recordings, instructions for the speakers.	2	1
2	**Creation of specifications on linguistic annotation of the speech corpus.** Specifications will determine the procedure of segmentation of the speech recordings, the number of segmentation levels (phonemes/syllables/ words), phonetic transcription, stress marking, procedure of verification of segmentation and transcription, segmentation into vocalic and consonantal intervals, and suprasegmental annotation – the number of prominence and phrasing levels. Specifications will also determine which parts of the corpus will be labeled at the suprasegmental level.	2	3
3	**Creation of specifications for acoustic measurements and linguistic feature extraction for needs of speech rhythm analyses in Polish.** On the basis of the state-of-the-art knowledge, acoustic parameters in the duration, pitch and intensity domain (spectral emphasis will also be considered) and linguistic features (e.g., syllable and C- and V-interval structure, syllable and word position in phrase) for use in the analyses (tasks 7-13) will be specified.	2	3
4	**Development of speech corpus for rhythm research.** This task concerns recordings of 20 native and 15 non-native speakers of Polish. The whole recorded speech material will be evaluated in terms of audio quality and accordance with the specifications (e.g., identification of disfluent utterances that will not qualify for acoustic analyses). A database will be developed via which the recordings will be accessed. The database will also integrate tools for segmentation and annotation of the speech material and for acoustic measurements & linguistic feature extraction.	3	5
5	**Creation of linguistic annotations of the speech corpus.** It will involve automatic segmentation at phoneme, syllable and word level and phonetic transcription, and their verification by two expert	3	8

	phoneticians. Suprasegmental annotation will consist in marking four levels of prominence and position and strength of intonation phrase boundaries (parts 1, 3 & 4 of the corpus). Each utterance in the corpus will be segmented into vocalic (V) and consonantal (C) intervals.		
6	**Creation of database of acoustic measurements and linguistic features.** The database will be created on the basis of measurements carried out automatically using scripts and plug-ins developed for Annotation Pro and Praat. For each sentence and speaker six metric scores will be calculated on the basis of segmentation into vocalic and consonantal intervals: %V, ΔC, nPVI-rPVI, VarcoV and VarcoC. Parameters in duration, pitch and intensity domain will be automatically extracted for each syllable and vowel and each prominence and phrasing level. The measurements will be normalized using the z-score normalization method. Speech rate will be calculated in syllables per second and as the sum of the number of C- and V-intervals per second. Linguistic information will be extracted as well for needs of specific analyses. The database will be subject to validation to identify missing and (potentially) faulty annotations and measurements (e.g., unrealistic pitch or duration values).	2	11
7	**Investigation of the relation between Polish rhythm and speech rate.** For this task only recordings of native speakers will be used including running speech (text), sentences and poetic verses produced at five different tempos: normal, very slow, slow, very fast and fast. For each of the three parts, relation between intended and laboratory measured speech rate and ratios between syllable- and CV-rate will be analyzed by means of correlation coefficients. On this basis speech rate typical of Polish utterances will be determined. Next, variation in timing patterns expressed by means of rhythm metrics due to speech rate differences will be investigated within and between various speech styles. Another analysis will concern the effect of speech rate on realization of prominence and phrasing. The effects will be examined using ANOVA or GLMM.	3	13
8	**Investigation of phonotactic and phrasal properties of Polish rhythm.** It will be based on native and non-native speech material (text, sentences and mini-dialogues produced at normal tempo). Effects of syllable structure complexity on the rhythm measures and effects of levels of prominence and phrasing on vowel and syllable durations across different accents (native vs. non-native) will be analyzed using ANOVA or GLMM.	3	13
9	**Investigation of stability of rhythm metrics.** First, the effect of the elicitation method and speech style on the metric scores will be investigated using utterances produced at a normal speech rate (native speakers, parts 1, 2, 4 & 5 of the corpus). Then the effect of speech rate will be examined using tempo differentiated utterances (native speakers, parts 1, 2 & 4). Stability of metrics will be also investigated using recordings representing native Polish and three foreign accents from two parts of the corpus: running speech and sentences. If metrics are stable then there should be no statistically significant differences in metric scores within accent groups (White & Mattys 2007, Wiget et al. 2010). The most stable metrics will be selected (on the basis of statistical analyses results) and used in further analyses. The analysis will also test the hypothesis that rhythm metrics can be a useful measure of timing processes.	2	16
10	**Investigation of contribution of duration, pitch, intensity and tempo to acoustic realization of meter and grouping.** Acoustic realization of different levels of prominences (unstressed, stressed unaccented, pitch accented and nuclear accent, e.g. Selkirk 1985) and phrasing (minor and major boundary) will be examined on the basis of native Polish material (running speech and poetic verses read at all speech rates). The contribution of pitch, duration and intensity features to realization of prominence (meter) and phrasing (grouping) will be assessed	3	18

	using regression analysis. Differences due to speech rate changes will also be analyzed (both in terms of acoustic features and distribution & number of prominent syllables and phrase boundaries).		
11	**Analysis of cross-linguistic perception and production of prominence and phrasing.** It will be based on prosodic annotations of running speech (text) and mini-dialogues (only at normal tempo) provided by native and non-native speakers. Regression analysis will be applied to reveal to which extent pitch, duration and intensity features are used by native and non-native speakers to produce specific prosodic structures. Next, particular contexts which make it difficult for non-native speakers to recognize prominent syllables in Polish will be identified. Acoustic correlates of subjects' prominence detection will be examined to see which features subjects attend to in making their judgments. A subject score for each utterance will be calculated and a comparison will be made between the acoustic correlates of misses and false alarms and the correlates of subject score.	3	18
12	**Creation of a quantitative description of speech rhythm in native and non-native Polish.** A meta-analysis of the results of the tasks 7-11 will be carried out to provide a multidimensional description of speech rhythm in native and non-native Polish. The description will incorporate knowledge about timing patterns expressed using selected rhythm metrics, acoustic realization of prominence and phrasing, and the effect of phonotactic structure and tempo on rhythmic organization of utterances. It will be used to study rhythm transfer from speaker's L1 to target language Polish.	2	21
13	**Investigation of rhythmic classification of Polish.** It will consist in a cross-linguistic comparison of the results from analyses in the tasks 7-11 with results for other languages presented in the literature. Rhythm of other languages will be described using the framework of the multidimensional description rhythm (proposed in task 12).	2	23
14	**Analysis of experimental results.** This task consists in compilation and processing of experimental results related to tasks 7-13 for publication and dissemination of project results.	12	12

4.2 State of pre-exisitng and initial research indicating feasibility of research objectives

Recently, a quantitative analysis of the realization of speech rhythm in Polish by native Polish speakers (5) and two groups of non-native speakers (7 with L1 German and 5 with L1 Korean) has been carried out (Wagner A. 2013). Two types of rhythm metrics were used: %V-ΔC and nPVI-rPVI and rate-normalized consonantal metrics lnΔC and nPVI-C. ANOVA and Tukey's tests results showed that combination of vocalic metrics %V and nPVI discriminates the best between native and non-native speech, as well as between the two non-native accents and can be used in analysis of rhythm transfer from speaker's L1 to the target language Polish. Smaller amount of significant differences in the rhythm scores between German-accented and native Polish than between Korean-accented and native Polish were in line with the perceptual assessment of the salience of the foreign accent in these two accent groups (strong in Korean-accented Polish and moderate in German-accented Polish). The scores of the rhythm metrics obtained for native Polish utterances indicated "mixed" character of Polish rhythm, with some features of syllable-timing and some of stress-timing. Cross-linguistic analysis of phonotactic structure of syllables (comparison with results for English, German, French, Italian and Czech was carried out) also indicated differences from both stress- and syllable-timed languages ("Quantitative description of rhythm in Polish as a foreign language", paper in review). The greatest similarity in the phonotactic structure was found with Czech which is considered as rhythmically unclassified. The study found also some evidence for instability of rhythm metrics which in certain cases are sensitive to speech rate.

Generally, the preliminary results indicate need for a more in-depth analysis of factors which affect rhythmic measurements and need to expand the scope of Polish rhythm research to encompass other important and so far neglected aspects, such as relation between rhythm and tempo, phonotactic and phrasal properties of speech rhythm, contribution of duration, pitch, intensity and tempo to acoustic realization of meter and grouping, and perception of prominence and phrasing.

5. References

Abercrombie, D. (1967). *Elements of general phonetics* (Vol. 203). Edinburgh: Edinburgh University Press.

Aguilar, L., Bonafonte, A., Campillo, F., & Mancebo, D. E. (2009). Determining intonational boundaries from the acoustic signal. In *INTERSPEECH* (pp. 2447-2450).

Allen, G. D. (1975). Speech rhythms: Its relation to performance universals and articulatory timing. *Journal of Phonetics, 3*:75–86.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66*(1-2), 46-63.

Arvaniti, A. (2012a). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351-373.

Arvaniti, A. (2012b). Rhythm classes and speech perception. *Understanding Prosody: The Role of Context, Function and Communication, 13*, 75-92.

Barry, W., Andreeva, B., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm? *Phonetica, 66*(1-2), 78-94.

Beckman, M. E. (1992). Evidence for speech rhythms across languages. *Speech perception, production and linguistic structure*, 457-463.

Beňuš, Š. & Šimko, J. (2012). Rhythm and tempo in Slovak. In *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China.

Boersma, Paul & Weenink, David (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.59, retrieved 20 November 2013 from http://www.praat.org/

Carlson, R., Hirschberg, J., & Swerts, M. (2005). Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech communication, 46*(3), 326-333.

Dauer, R. (1983). Stress timing and syllable timing reanalyzed. *Journal of Phonetics, 11*:51–62.

Dauer, R. (1987). Phonetic and phonological components of language rhythm. In *Proceedings of the XIth International Congress of the Phonetic Sciences*, volume 5, pages 447–450. Tallinn: Academy of Sciences.

Dellwo, V. & Wagner, P. (2003). Relationships between speech rate and rhythm. In *Proceedings of the 15th International Conference of the Phonetic Sciences*, pages 471–474, Barcelona, Spain.

Dellwo, V. (2008). The influence of speech rate on speech rhythm. Unpublished PhD thesis, Universität Bonn

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient forΔ C. *Language and language processing*, 231-241.

Dellwo, V., Steiner, I., Aschenberner, B., Dankovicova, J., & Wagner, P. (2004). BonnTempo-Corpus & BonnTempo-Tools: A database for the study of speech rhythm and rate. In *Proceedings INTERSPEECH 2004*, pages 777–780, Jeju Island, Korea.

Demenko, G. (1999). Analysis of Polish Suprasegmentals for needs of Speech Technology. UAM: Poznań.

Demenko G., Wypych M. & Baranowska E. (2003.) Implementation of Polish grapheme-tophoneme rules and extended SAMPA alphabet in Polish text-to-speech synthesis. *Speech and Language Technology 7*:79-96

Den Os, E. (1985). Perception of speech rate of Dutch and Italian utterances. *Phonetica, 42*(2-3), 124-134.

Dłuska, M. 1950. Fonetyka polska. PWN: Warszawa.

Eriksson, A., Grabe, E., and Traunmuller, H. (2002). Perception of syllable prominence by listeners with and without competence in the tested language. In Proceedings of Speech Prosody, Aix en Provence, France.

Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French, and English. *Journal of phonetics 19*(3-4), Jul-Oct 1991, 351-365.

Fraisse, P. (1963). The psychology of time. NY: Harper and Row.

Garding, E. (1975). The influence of tempo on rhythmic and tonal patterns in three Swedish dialects. *Lund University Phonetic Laboratory Working Papers, 12*, 71-83.

Garding, E., & Zhang, J. (1997). Tempo effects in Chinese prosodic patterns. In *Intonation: Theory, Models and Applications.*

Gibbon, D., Bachan, J., & Demenko, G. (2007). Syllable timing patterns in Polish: results from annotation mining. In *INTERSPEECH* (pp. 994-997).

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*(515-546).

Hirschberg J. (1993). Pitch accent in context: predicting intonational prominence from text. *Aritificial Intelligence 63*(1-2):305-340.

Holliman, A. J., Wood, C., & Sheehy, K. (2012). A cross-sectional study of prosodic sensitivity and reading difficulties. *Journal of Research in Reading, 35*(1), 32-48.

Jassem, W. (1962). Akcent języka polskiego. Wrocław: Ossolineum.

Kinoshita, N., & Sheppard, C. (2011). Validating acoustic measures of speech rhythm for second language acquisition. In *Proceedings of the XIth International Congress of the Phonetic Sciences* (Vol. 17, pp. 1086-1089).

Klessa, K. (2006). Analiza iloczasu głoskowego na potrzeby syntezy mowy polskiej (Analysis of segmental duration for needs of speech synthesis in Polish). Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.

Klessa, K. (2011). Polish segmental duration: selected observations based on corpus data. *Speech and Language Technology* vol. 14/15, pp. 95-103

Klessa, K., Karpiński, M. & Wagner, A. (2013). Annotation Pro – a new software tool for annotation of linguistic and paralinguistic features [in:] Brigitte Bigi and Daniel Hirst (Eds.), *Proceedings of TRASP*. Aix-en-Provence, 30 August, 2013. ISBN: 978-2-7466-6443-2

Leemann, A., Dellwo, V., Kolly, M. J., & Schmid, S. (2012). Rhythmic variability in Swiss German dialects. In *Proceedings of the 6th International Conference on Speech Prosody,* Shanghai, China.

Li, A. and Post, B. (2012). L2 rhythm development by Mandarin Chinese learners of English. In *Proceedings of Perspectives on Rhythm and Timing*, 19-21 July 2012, Glasgow, UK.

Liss, J., White, L., Mattys, S., Lansford, K., Lotto, A., Spitzer, S. & Caviness, J. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research* Vol.52 1334-1352.

Mairano, P. and Romano, A. (2011). Rhythm metrics for 21 languages. In *Proceedings of the XVIIth International Congress of Phonetic Sciences,* 17-21 August 2011, Hong Kong, China. 1318-1321.

Malisz, Z., & Klessa, K. (2008). A preliminary study of temporal adaptation in Polish VC groups. In *Proceedings of the Speech Prosody 2008 Conference*, Campinas, Brazil.

Malisz, Z., & Wagner, P. (2012). Acoustic-phonetic realization of Polish syllable prominence: a corpus study. *Speech and Language Technology* vol. 14/15, pp. 105-114

Malisz, Z. (2011). Tempo differentiated analyses of timing in Polish. In *Proceedings of the XVIIth International Congress of Phonetic Sciences,* 17-21 August 2011, Hong Kong, China. 1322-1225.

Mok, P. P., & Dellwo, V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In *Proceedings of the Speech Prosody 2008 Conference* (pp. 423-426).

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology 24*(3). 756-766.

Nolan, F., & Asu, E. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica, 66*(1-2), 64-77.

Ordin, M., Polyanskaya, L. & Ulbrich, Ch. (2011). Acquisition of Timing Patterns in Second Language. In *Proceedings of INTERSPEECH 2011*, 27-31 August 2011, Florence, Italy. 1129-1132.

Ostaszewska, D., & Tambor, J. (2000). Fonetyka i fonologia współczesnego języka polskiego. Wydawnictwo Naukowe PWN.

Paquette, K. R., & Rieg, S. A. (2008). Using music to support the literacy development of young English language learners. *Early Childhood Education Journal, 36*(3), 227-232.

Payne, E., Post, B., Astruc, L., Prieto, P. & del Mar Vanrell, M. (2012). Measuring child rhythm. *Language and Speech, 55*(2), 203-229.

Pike, K. L. (1945). The Intonation of American English. Ann Arbor: University of Michigan.

Portele, T. (1998). Perceived Prominence and acoustic parameters in American English. In *Proceedings of ICSLP'98*, Sydney, 3, 667–670.

Prieto, P., Vanrell, M. D. M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication, 54*(6), 681-702.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition 73*(3). 1-28.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the XVth International Congress of Phonetic Sciences*, August 3-9, 2003, Barcelona, Spain. 337–342.

Richter, L. (1987). Modelling of the rhythmic structure of utterances in Polish. *Studia Phonetica Posnaniensia 1*, 91–125.

Rosenberg, A., Hirschberg, J. & Manis, K. (2010). Perception of English prominence by native Mandarin Chinese speakers. In *Proceedings of Speech Prosody 2010*.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress and phrasing. In Goldsmith, J., editor, Handbook of Phonological Theory. Oxford.

Sridhar, V. K. R., Nenkova, A., Narayanan, S., & Jurafsky, D. (2008). Detecting prominence in conversational speech: pitch accent, givenness and focus. In *Proceedings of Speech Prosody*, Campinas, Brazil.

Tamburini, F. (2003). Automatic prosodic prominence detection in speech using acoustic features: an unsupervised system. In *INTERSPEECH*, vol.1, pp. 385-388

Terken J. (1991). Fundamental frequency and perceived prominence of accented syllables. *J. Acoust. Soc. Am. 89*:1768-1776

Tortel, A., & Hirst, D. (2010). Rhythm metrics and the production of English L1/L2. In *Proceedings of Speech Prosody 2010*.

Turk, A.E. & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics 27*, 171–206.

Wagner, A. (2008). Comprehensive model of intonation for application in speech synthesis. Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.

Wagner, A. (2009.) Analysis and recognition of accentual patterns. In *Proceedings of INTERSPEECH 2009*, 6-10 Sept., Brighton, UK

Wagner, A. (2010). Acoustic cues for automatic determination of phrasing. In *Proceedings of Speech Prosody 2010*, 11-14 May 2010, Chicago, USA.

Wagner, A. (2013). Struktura rytmiczna wypowiedzi w polskiej mowie natywnej i nienatywnej (Rhythmic structure of utterances in Polish native and non-native speech). Scripta manent - res novae, Stanisław Puppel, Teresa Tomaszkiewicz (eds.), pp. 499-510

Wagner, P. (2005). Great expectations - introspective vs. perceptual prominence ratings and their acoustic correlates. In *Proceedings of Interspeech*, pages 2381–2384, Lisbon, Portugal.

Wagner, P. (2008). The rhythm of language and speech: Constraining factors, models, metrics and applications. Habilitationsschrift, University of Bonn.

White, L. & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics 3*(5). 501-522

White L., Mattys S. L., Series L. & Gage S. (2007). Rhythm metrics predict rhythmic discrimination. In *Proceedings of the Sixteenth International Congress of Phonetic Sciences*, Saarbrücken, Germany, pp. 1009-1012.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O. & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *J. Acoust. Soc. Am 127*(3). 1559-1569.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am. 91*, 1707–1717.

Woodrow, H. (1951). Time perception. In Stevens, S. S., editor, Handbook of Experimental Psychology, chapter 32, pages 1224–36.Wiley and Sons, New York.