

ОТГОВОРНОСТТА ПРЕД ЕЗИКА

Книга 10

СБОРНИК, ПОСВЕТЕН НА
65-ГОДИШНИНАТА НА
ПРОФ. Д-Р ДИМИТЪР ПОПОВ

Университетско издателство
„Епископ Константин Преславски“
Шумен, 2024

SPRACHRRHYTHMUS UND PROSODISCHE STRUKTUR: EINE KRITISCHE BETRACHTUNG¹

Bistra Andreeva
Universität des Saarlandes
andreeva@lst.uni-saarland.de

SPEECH RHYTHM AND PROSODIC STRUCTURE: A CRITICAL PERSPECTIVE

Bistra Andreeva
University Saarlandes
andreeva@lst.uni-saarland.de

РИТЪМ НА РЕЧТА И ПРОЗОДИЧНА СТРУКТУРА: КРИТИЧНА ПЕРСПЕКТИВА

Бистра Андреева
Университет на Саарланд
andreeva@lst.uni-saarland.de

Abstract: This study investigates the role of rhythm in spoken language and the interactions between acoustic parameters that influence prominence and rhythm. It argues that the search for acoustic correlates of rhythm cannot be limited to temporal aspects but must consider all four structuring features of the acoustic speech signal. Additionally, it emphasizes the need to account for individual and language-specific variations in the perception of rhythm.

Key words: rhythm, prominence, acoustic parameters, speech production, language variation, perception

Die vorliegende Arbeit untersucht die Rolle des Rhythmus in der gesprochenen Sprache und die Wechselwirkungen zwischen akustischen Parametern, die die Prominenz und den Rhythmus beeinflussen. Es wird argumentiert, dass die Suche nach akustischen Korrelaten des Rhythmus nicht auf temporale Aspekte beschränkt werden kann, sondern alle vier Strukturierungsmerkmale des akustischen Sprachsignals berücksichtigen muss. Zudem wird die Notwendigkeit hervorgehoben, individuelle und sprachspezifische Variationen in der Wahrnehmung von Rhythmus zu berücksichtigen.

¹ Die vorliegende Arbeit basiert auf einer früheren Studie, die gemeinsam mit Bill Barry veröffentlicht wurde.

Schlüsselwörter: Rhythmus, Prominenz, akustische Parameter, Sprachproduktion, Sprachvariation, Wahrnehmung

1. Einführung

Ein faszinierender, aber zugleich frustrierender Aspekt der instrumentellen Analyse des Sprachrhythmus besteht darin, dass dieses Konzept, trotz fehlender empirischer Begründung, scheinbar so lange überlebt hat, wie ein menschliches Leben. Historisch betrachtet könnte man erwarten, dass sich das Konzept der dritten Generation nähert – einer Phase, die traditionell mit dem allmählichen Auslaufen einer Idee verbunden ist. Dennoch hielten frühe Studien, die sich stark auf das Konzept der Isochronie – ob auf Silben- oder Fußebene – stützten, über ein halbes Jahrhundert lang an, auch wenn die Skepsis gegenüber diesem Ansatz zunahm [Classe 1939, Pointon 1995]. Zahlreiche Versuche, artikulatorische, akustische oder perzeptuelle Erklärungen zu finden, blieben erfolglos (siehe die ausführliche Diskussion bei Bertinetto 1989), bis ein methodischer Paradigmenwechsel eine zweite Generation von Untersuchungen einleitete.

Der Begriff der Isochronie verlor seine Bedeutung als schwer erreichbares Ziel in der Rhythmusforschung und wurde durch das Konzept der Variabilität ersetzt. Die traditionellen rhythmischen Basiseinheiten – Silbe und Takt (später auch die Mora) – wurden durch die universellen Kern- (sonore) und Nicht-Kern- (konsonantische) Segmente innerhalb der Silbenstruktur abgelöst. Gegen Ende der 1990er Jahre etablierten sich zwei wesentliche Maße: (a) Frank Ramus' „Delta“-Werte (ΔC und ΔV) [Ramus et al. 1999], welche die Standardabweichung der vokalischen und konsonantischen Intervalle innerhalb einer Äußerung messen, ergänzt durch den vokalischen Anteil (%V); und (b) der „Pairwise Variability Index“ (PVI) von Ee Low, entwickelt in Zusammenarbeit mit Esther Grabe und Francis Nolan [Low et al. 2000]. Der PVI berechnet den durchschnittlichen Unterschied zwischen aufeinanderfolgenden vokalischen oder intervokalischen Intervallen in einer Äußerung, wobei für die Vokalintervalle eine lokal normalisierte Differenz verwendet wird, um mögliche Schwankungen der Sprechgeschwindigkeit zu berücksichtigen. Beide Maße ermöglichen eine zweidimensionale Zuordnung der Sprachwerte in einen Rhythmusraum (siehe Abbildung 1).

Dieser Rhythmusraum erleichtert die Visualisierung des Kontinuums zwischen akzent- und silbenzählenden Sprachen und bietet einen diagnostischen Rahmen, um die Dimensionen (vokalisches oder konsonantisches) zu identifizieren, in denen jede Sprache vom Prototyp abweicht. Allerdings gibt es keine theoretisch vorhergesagte Position in diesem Raum für die morazählenden Sprachen.

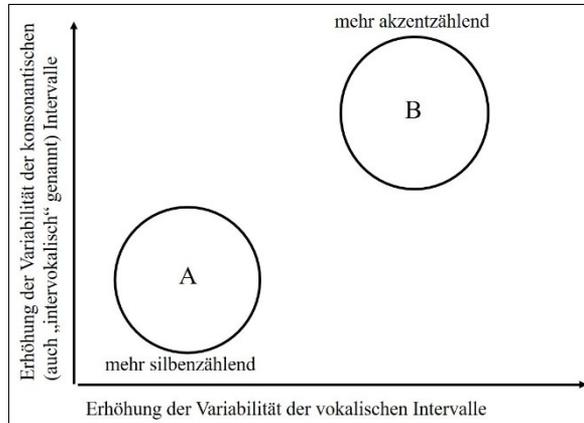


Abbildung 1: Rhythmusraum, definiert durch variabilitätsbasierte Rhythmusmaße.

Anfangs schienen diese Maße der zweiten Generation erfolgreich darin, Sprachen zu gruppieren, die laut wissenschaftlichem Konsens zusammengehörten, und gleichzeitig diese Gruppen voneinander abzugrenzen. Doch mit der zunehmenden Verbreitung abgeleiteter Maße wuchs auch die kritische Aufmerksamkeit [Barry et al. 2003], die bald zunehmend präziser und fundamentaler wurde [Russo, Barry 2008; Arvaniti 2009; Kohler 2009b; Nolan, Asu 2009].

Ob diese Kritik zu einem frühen Ende der Forschung der zweiten Generation führen und den Weg für ein neues, bislang undefiniertes Paradigma der dritten Generation ebnen wird [Kohler 2009b], oder ob neue Maße entstehen, die die spezifischen Kritikpunkte einzelner Forscher berücksichtigen, bleibt ungewiss. Klar ist jedoch, dass Rhythmus weithin als Eigenschaft gesprochener Sprache anerkannt ist und allgemein als instrumentell messbar gilt.

Diese Arbeit untersucht, was Rhythmus tatsächlich ist und ob er sich als von anderen prosodischen Merkmalen unabhängige Eigenschaft darstellen lässt. Beim Nachdenken über das Wesen des Rhythmus erörtern wir auch seine Messbarkeit anhand möglicher Parameter, die ihn formen. Im Wesentlichen fragen wir, ob die Herausforderung der Rhythmusmessung durch eine unkritische Akzeptanz des Phänomens erschwert wurde, die möglicherweise den Blick auf seine Einzelemente trübt.

2. Rhythmus in der gesprochenen Sprache

Instrumentelle Studien zum Rhythmus in der gesprochenen Sprache nehmen häufig die Existenz dieses Phänomens als gegeben an, ohne es in Bezug auf seine auditiven Merkmale klar zu definieren. Obwohl die

Rhythmusforschung oft auf Lloyd James' Beitrag von 1940 zurückgeführt wird, in dem er auditive Unterschiede zwischen Sprachen mit einem „Morse-Code-Rhythmus“ (akzentzählend) und solchen mit einem „Maschinengewehr-Rhythmus“ (silbenzählend) beschreibt, behandelt die instrumentelle Analyse typischerweise diese Unterschiede als kategorisch, was zu einer Rhythmustypologie führt. Dies führt dazu, dass die Korrelation zwischen gemessenen Werten und den systematischen auditiven Eindrücken der analysierten Texte oft übersehen wird.

Zu der Zeit, als Isochronie im Mittelpunkt stand, schien eine plausible Verbindung zwischen den gemessenen Einheiten und Lloyd James' Beschreibungen gegeben zu sein. Allerdings können Veränderungen in der Textstruktur drastische Änderungen der rhythmischen Eindrücke bewirken. Selbst innerhalb eines Dialekts variieren individuelle Sprechstile stark in ihrer wahrgenommenen Rhythmizität [Kohler 2009a]. Auditive Bewertungen, wie verschiedene Texte zu Lloyd James' Beschreibungen passen, wären daher weiterhin wertvoll, idealerweise gefolgt von Vergleichen mit instrumentellen Messungen.

Mit dem Übergang von Isochronie zu Variabilität als Paradigma ging die direkte Verbindung zu den ursprünglichen auditiven Eindrücken verloren, ebenso wie die Relevanz der Silbe als Rhythmuseinheit. Intervokalische Intervalle (d. h. die Abstände zwischen Vokalen) erstrecken sich häufig über Silbengrenzen hinweg, was eine perfekte silbische Regelmäßigkeit unwahrscheinlich macht und Variabilität als differenzierendes Merkmal zwischen Sprachen hervorhebt. Ein Konsens über die akustische Grundlage für die auditive Unterscheidung von Sprachrhythmen wurde jedoch über die inhärenten Dauerverhältnisse sub-silbischer Maße hinaus nicht erreicht, auch wenn angenommen wird, dass die Sensibilität hierfür sich bereits im Mutterleib entwickelt und auch bei nicht-menschlichen Arten beobachtet werden kann [Nazzi et al. 1998; Ramus 2002; Ramus et al. 2000; Ramus, Mehler 1999].

Diese Überlegung ist jedoch zweitrangig gegenüber dem Verständnis der Wahrnehmung von Rhythmus in der gesprochenen Sprache. Die Suche nach einer akustischen Grundlage dieser Unterscheidung bleibt weitgehend unbeantwortet. Dellwo [2008] untersuchte die Beziehung zwischen Aufnahmen einer vermeintlich silbenzählenden Sprache, die geringere Variabilität aufweist, sowie den Urteilen über deren Rhythmus. Er verglich diese mit Aufnahmen, die eine höhere Variabilität aufweisen, die mit einem akzentzählenden rhythmischen Muster verbunden sind. Auffällig war, dass diese Urteile mehr vom Abspieltempo als von der Herkunft der Aufnahme beeinflusst wurden: verlangsamte Aufnahmen wurden als unregelmäßig (=akzentzählend) wahrgenommen, während beschleunigte Versionen als regelmäßig (=silbenzählend) empfunden wurden.

Ein weiterer Aspekt, der die Definition von Rhythmus in der gesprochenen Sprache erschwert, ist die Erwartung, dass sich Rhythmus als schlagähnliches Phänomen manifestiert. Es ist jedoch nicht nur unangebracht, den Begriff Rhythmus auf regelmäßige Abfolgen innerhalb kurzer Zeitintervalle zu beschränken; dies würde auch die Auswirkungen der informativen Gewichtung von Wörtern im spezifischen kommunikativen Kontext des Sprechers vernachlässigen. Selbst in Sprachen ohne offensichtlichen lexikalischen Akzent, wie dem Französischen, gibt es deutliche supralexikalische, informationsstrukturelle und affektbasierte Hervorhebungen, die das angenommene Muster zeitlicher Regelmäßigkeit (Isochronie) in den Silbenfolgen vor der abschließenden Betonung einer Phrase stören. In Sprachen mit lexikalischem Akzent scheitert das Konzept der Fußregelmäßigkeit zudem zwangsläufig, wenn prenukleare, nukleare und möglicherweise emphatische Akzente innerhalb einer Äußerung berücksichtigt werden.

Zusammenfassend lässt sich sagen, dass die Berücksichtigung der mehrdimensionalen Natur der Akzentuierung jede Erwartung an Einheitlichkeit – sei es auf Silben- oder Fußebene – als verfehlt erscheinen lässt. Dies wird durch instrumentelle Ergebnisse zur Isochronie gestützt, wie zahlreiche rhythmusorientierte Publikationen der letzten Jahrzehnte gezeigt haben. Der Einfluss einer solchen durch die Informationsstruktur bedingten „kommunikativen Akzentuierung“ auf Variabilitätsmessungen ist schwer vorherzusagen, insbesondere in Sprachen mit einem lexikalischen Akzentsystem. In Äußerungen von Sprachen ohne lexikalischen Akzent führen „kommunikative“ Akzente zwangsläufig zu einer erhöhten Silben-zu-Silben-Variabilität, während ihr Einfluss in Sprachen mit lexikalischem Akzent unterschiedlich ist. Barry et al. [2009] zeigen, dass eine kontrastive Akzentuierung die silbische Variabilität im Deutschen im Vergleich zu einem weiten Fokus um etwa 13 % erhöht – sowohl in frühen als auch in späten Satzpositionen. Im Bulgarischen führt ein früher kontrastiver Akzent zu einem deutlichen Anstieg des Variabilitätswerts um 25 %, während ein später Akzent kaum einen Effekt zeigt (4 %). Im Englischen hingegen wurde kaum eine Veränderung der Variabilität festgestellt (0,4 % Anstieg für einen frühen und 3,2 % für einen späten kontrastiven Akzent).

Es wäre illusorisch zu behaupten, dass der Rhythmus englischer Äußerungen mit unterschiedlichen Akzentuierungsmustern trotz ähnlicher Veränderungen im Muster der Hervorhebungen gleich bleibt, während sich der Rhythmus deutscher und bulgarischer Äußerungen mit wechselnder Akzentuierung verändert. Es scheint also, dass Rhythmusmaße, die auf Variabilität basieren, den wahrgenommenen Rhythmus ebenso wenig widerspiegeln wie Isochroniemaße. Auf der anderen Seite berichtete derselbe Artikel [Barry et al. 2009] über systematische Verschiebungen in den

Rhythmusmaßen, wie sie für verschiedene Arten poetischer Versmaße vorhergesagt wurden, einschließlich eines vorhergesagten Rückgangs der Variabilitätswerte in den Fällen, in denen eine Deakzentuierung innerhalb des regelmäßigen Versmaßes wahrgenommen wurde.

Daher müssen wir schlussfolgern, dass die variabilitätsbasierten Rhythmusmaße Dauerunterschiede zwischen aufeinanderfolgenden vokalischen und intervokalischen Intervallen erfassen, von denen viele die Unterschiede in der Silberstruktur widerspiegeln und einige die dauerbezogenen Auswirkungen der Prominenzstufen zeigen. Wenn Prominenzen regelmäßig auftreten, wie im poetischen Versmaß, können wir leicht einen regelmäßigen Rhythmus identifizieren. In der normalen, auf Kommunikation ausgerichteten Sprache wird der Rhythmus einer Äußerung weiterhin von der Akzentuierung getragen, aber das Muster kann selten regelmäßig sein, da es das Produkt des relativen informativen Gewichts ist, das den spezifischen Wörtern gegeben wird, die ausgewählt wurden, um die Nachricht in ihrer syntaktisch bestimmten Position in der Sequenz zu übermitteln. Es scheint, dass wir für Bertráns [1999] Alternative des nicht isochronen Rhythmus in der Sprache plädieren. Sein Status als rhythmisches Muster wird durch die Tatsache unterstützt, dass eine Veränderung im Muster der relativen Prominenz entweder die Gesamtbotschaft der Äußerung verändert oder sie für einen Muttersprachler falsch oder möglicherweise fremd klingen lässt. Ein Teil unserer Sprachkompetenz besteht in der Sensibilität für die Beziehung zwischen dem sich entwickelnden rhythmischen Muster und der Gesamtbotschaft.

3. Welche Parameter liegen dem Rhythmus der gesprochenen Sprache zugrunde?

Wenn wir erkennen, dass Rhythmus aus dem differenzierten Aufwand resultiert, der mit der Übermittlung einer Botschaft durch die Wörter einer Äußerung verbunden ist, um deren notwendige relative Informationsgewichtung zu erreichen, wird offensichtlich, dass die Suche nach akustischen Korrelaten des Rhythmus nicht auf temporale Aspekte beschränkt werden kann [Kohler 2009b; Nolan, Asu 2009]. Physiologisch manifestiert sich dieser Aufwand in allen Aspekten des Produktionsmechanismus und umfasst alle vier Strukturierungsmerkmale des akustischen Sprachsignals, sofern nicht anders eingegrenzt. Ein Wort hervorzuheben bedeutet, Reduktionsprozesse zu vermeiden, die in vielen Sprachen für akustisch unwichtige Wörter zulässig sind [Barry, Andreeva 2001]. Dies erfordert das Streben nach vollständiger segmentaler Klarheit, was artikulatorisch komplex und zeitaufwändig ist und zu einer erhöhten Dauer führt. Gleichzeitig wird die spektrale Unterscheidbarkeit größer sein als bei einer unbetonten Produktion desselben Wortes. *Ceteris paribus* sollte

das betonte Wort lauter sein als die angrenzenden weniger bedeutenden Wörter, was eine Erhöhung der Ausatemungskraft und des subglottalen Drucks impliziert. Dies erhöht die Signalintensität und ohne glottale Anpassungen auch die Grundfrequenz (f_0).

Diese Wechselwirkung zwischen den Parametern könnte darauf hindeuten, dass es nicht notwendig ist, alle zu messen, da die bisher dominierenden Daten zur Dauer, die in Rhythmusanalysen verwendet wurden, ausreichend sein könnten. Wir wissen jedoch, dass f_0 auch unabhängig von der Intensität modifiziert werden kann, wobei Akzente relativ hoch oder niedrig im Vergleich zu den umgebenden Silben sein oder durch einen ansteigenden oder fallenden f_0 die Prominenz erhöhen können. In ähnlicher Weise ist eine bestimmte Mindestdauer erforderlich, um eine vollständige segmentale Definition zu erreichen, aber darüber hinaus hört die Wechselwirkung auf. Längungen können auch für Emphase oder Affekt verwendet werden. Es wurde jedoch auch festgestellt [Gay 1981], dass ein größerer artikulatorischer Aufwand dazu führen kann, einige Reduktionen artikulatorischer Gesten zu vermeiden, was die Wechselwirkung schwächt.

Da es das Potenzial gibt, die vier akustischen Eigenschaften unabhängig voneinander zu nutzen, ist eine ausschließlich dauerabhängige Beschreibung des Rhythmus theoretisch inakzeptabel. Es gibt jedoch keine Möglichkeit, die gemessenen Werte unterschiedlicher Parameter auf eine einzige Maßzahl, z. B. die silbische Prominenz, zu reduzieren, bis der relative Beitrag jedes Parameters zu einem definierten Anstieg der Prominenz ermittelt wurde. Diese Beziehung könnte sich als eine Tauschbeziehung herausstellen, bei der eine Verringerung des Wertes eines Parameters durch eine Erhöhung eines anderen kompensiert werden kann. Eine solche Wechselwirkung könnte je nach Ausmaß, in dem die Parameter in der Produktion genutzt werden, von Hörer zu Hörer, von Dialekt zu Dialekt oder von Sprache zu Sprache variieren. Jüngere Untersuchungen zur Sprachproduktion in verschiedenen Sprachen zeigen systematische Unterschiede darin, wie stark Sprachen Dauer, f_0 und Intensität zur Steigerung der Prominenz einsetzen [Andreeva et al. 2007; Barry et al. 2007; Koreman et al. 2009]. Hinter diesen Sprachunterschieden gibt es jedoch auch erhebliche individuelle Variationen. Es ist ungewiss, ob die Unterschiede in der Produktion auch in die Wahrnehmungsstrategien übertragen werden; hierzu läuft die Forschung noch, erste Ergebnisse deuten jedoch darauf hin, dass dies in gewissem Maße der Fall ist [Rathcke et al. 2024].

Die unterschiedlichen Beiträge der Produktionsparameter zur Prominenz und die Rolle der Prominenz in der rhythmischen Struktur sollten uns nicht dazu verleiten, den Beitrag eines Parameters zur Prominenz automatisch mit seinem Beitrag zum Rhythmus gleichzusetzen. Barry et al. (2009) haben experimentell gezeigt, dass die Stärke des rhythmischen

Eindrucks, den wir Rhythmizität nennen, von mehreren Parametern abhängt, wobei der dauerbezogene Kontrast und die f_0 -Änderungen zwischen benachbarten Silben stärker zum Eindruck beitragen als Intensität oder Vokalkürzung. Der Grad der Rhythmizität ist jedoch nicht dasselbe wie rhythmisch versus arhythmisch. Tatsächlich erklärten einige Hörer, dass die Stimuli, bei denen alle alternierenden Eigenschaften neutralisiert wurden, ebenso rhythmisch – wenn auch anders rhythmisch – waren wie die vollständig differenzierten trochäischen Sequenzen.

Die in diesem Abschnitt diskutierten Ergebnisse kratzen nur an der Oberfläche des Problems, wie die akustischen Parameter, die das Sprachsignal strukturieren, miteinander interagieren, um die Prominenz zu bestimmen und wie sie über die Prominenz dazu beitragen, den Rhythmus in der gesprochenen Sprache zu definieren. Es ist jedoch klar, dass die Muster der Prominenz und die daraus resultierenden Rhythmen von mehr als nur den bisher gemessenen dauerbezogenen Parametern beeinflusst werden.

4. Schlussfolgerungen und Diskussion

In dieser Arbeit haben wir uns mit dem komplexen Phänomen des Rhythmus in der gesprochenen Sprache auseinandergesetzt. Ziel war es, die vielschichtigen Elemente zu beleuchten, die zur Entstehung von Rhythmus beitragen, und deren Wechselwirkungen zu untersuchen. Wir haben argumentiert, dass Rhythmus nicht einfach auf zeitliche Aspekte beschränkt werden kann, sondern dass er ein Zusammenspiel verschiedener akustischer Parameter darstellt, einschließlich Dauer, Tonhöhe und Intensität.

Unsere Untersuchungen zeigen, dass die Wahrnehmung von Rhythmus eng mit der kommunikativen Funktion von Sprache verknüpft ist. Der Rhythmus wird durch den differenziellen Aufwand geprägt, der in die Aussprache von Wörtern investiert wird, um deren relative Informationsgewichtung zu erreichen. Diese Erkenntnis fordert bestehende Konzepte heraus, die Rhythmus oft auf regelmäßige Zeitintervalle reduzieren. Stattdessen ist es notwendig, die verschiedenen Dimensionen des Rhythmus zu betrachten, um ein umfassenderes Verständnis zu erlangen.

Ein zentrales Ergebnis unserer Analyse ist, dass die Akzentuierung und die damit verbundenen Prominenzmuster entscheidend für die rhythmische Struktur sind. Die Tatsache, dass unterschiedliche Sprachen unterschiedliche Strategien zur Hervorhebung von Prominenz nutzen, weist darauf hin, dass Rhythmus ein linguistisches Phänomen ist, das stark von den spezifischen Eigenschaften der jeweiligen Sprache abhängt.

Die Diskussion um die Unterscheidung zwischen rhythmischen und arhythmischen Mustern zeigt, dass die Wahrnehmung von Rhythmus auch von individuellen und kulturellen Faktoren beeinflusst wird. Es ist bemerkenswert, dass einige Hörer auch in neutralisierten Stimuli rhythmische

Eigenschaften wahrnehmen können, was darauf hindeutet, dass Rhythmus nicht nur von akustischen Mustern abhängt, sondern auch von den kognitiven Referenzen, die die Hörer entwickeln.

Trotz der Erkenntnisse, die wir gewonnen haben, bleibt viel Raum für weitere Forschung. Es ist notwendig, empirische Studien durchzuführen, um die Wechselwirkungen zwischen den verschiedenen Parametern des Rhythmus besser zu verstehen. Insbesondere sollte untersucht werden, inwieweit diese Parameter nicht nur in der Produktion, sondern auch in der Wahrnehmung von Sprache eine Rolle spielen.

Insgesamt hoffen wir, dass unsere Arbeit nicht nur zur Klärung der Komplexität von Rhythmus in der gesprochenen Sprache beiträgt, sondern auch Anreize für zukünftige Forschungsarbeiten schafft. Wir sind überzeugt, dass die Auseinandersetzung mit diesen Fragen nicht nur unser Verständnis von Sprache vertieft, sondern auch praktischen Nutzen für Bereiche wie Sprachtherapie, Sprachunterricht und die Entwicklung von Sprachtechnologien haben kann.

Благодарности:

Изследването е разработено в рамките на научен проект „Прозодични аспекти на българския език в съпоставителен план с други езици с лексикално акцентуване“, финансиран от ФНИ на МОН с договор № Кп-06-Н40/11 от 12.12.2019.

Literature

Andreeva et al. 2007: Andreeva, B., W. Barry, I. Steiner. Producing Phrasal Prominence in German. – In: Proceedings of ICPhS XVI, Saarbrücken, 2007, 1209–1212.

Arvaniti 2009: Arvaniti, A. Rhythm Timing and the Timing of Rhythm. – *Phonetica* 66, 2009, 46–63.

Barry, Andreeva 2001: Barry, W., B. Andreeva. Cross-language similarities and differences in spontaneous speech patterns. – *Journal of the IPA* 31,1, 2001, 51–66.

Barry et al. 2003: Barry, W., B. Andreeva, M. Russo, S. Dimitrova, T. Kostadinova. Do rhythm measures tell us anything about language type? – In: Proceedings of XVth ICPhS, Barcelona, 2003, 2693–2696.

Barry et al. 2007: Barry, W., B. Andreeva, I. Steiner. The Phonetic Exponency of Phrasal Accentuation in French and German. In: Proceedings of Interspeech 2007, Antwerp, 1010–1013.

Barry et al. 2009: Barry, W., B. Andreeva, J. Koreman. Do Rhythm Measures Reflect Perceived Rhythm? – *Phonetica* 66, 78–94.

Bertrán 1999: Bertrán, A. P. Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. – *Language Design*, 2, 1999, 103–130.

Bertinetto 1989: Bertinetto, P. M. Reflections on the dichotomy ‘stress’ vs. ‘syllable-timing’. – *Revue de phonétique appliquée* 91.93, 1989, 99–130.

Classe 1939: Classe, A. *The Rhythm of English Prose*. Oxford: Basil Blackwell, 1939.

Dellwo 2008: Dellwo, V. The role of speech rate in perceiving speech rhythm. – In: Proceedings of Speech Prosody, Campinas, Brazil, 2008, 375–378.

Gay 1981: Gay, T. Mechanisms in the control of speech rate. – *Phonetica* 38, 1981, 148–158

Kohler 2009a: Kohler, K. J. Whither Speech Rhythm Research? Editorial to Rhythm in Speech and Language. From Theory to Data. – *Phonetica* 66, 2009, 5–14.

Kohler 2009b: Kohler, K. J. Rhythm in Speech and Language. A New Research Paradigm. – *Phonetica* 66, 2009, 29–45.

Koreman et al. 2009: Koreman, J., W. van Dommelen, R.-O. Sikveland, B. Andreeva, W. J. Barry. Cross-language differences in the production of phrasal prominence in Norwegian and German. – In: Martti Vainio, Reijo Aulanko, and Olli Aaltonen (eds.), *Nordic Prosody, Proceedings of the Xth Conference, Helsinki 2008*, Frankfurt: Peter Lang, 139–150.

Lloyd James 1940: Lloyd James, A. *Speech Signals in Telephony*. London: Pitman & Sons, 1940.

Low et al. 2000: Low, E. L., E. Grabe, F. Nolan. Quantitative characterisations of speech rhythm: ‘Syllable-timing’ in Singapore English. – *Language and Speech*, 43, 2000, 377–401.

Nazzi et al. 1998: Nazzi, T., J. Bertoincini, J. Mehler. Language discrimination by newborns: Toward an understanding of the role of rhythm. – *Experimental Psychology* 24(3), 1998, 756–766.

Nolan, Asu 2009: Nolan, F., E. L. Asu. The Pairwise Variability Index and Coexisting Rhythms in Languages. – *Phonetica* 66, 2009, 64–77.

Pointon 1995: Pointon, G. E. Rhythm and duration in Spanish. – In Lewis, J. W. (ed.), *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O’Connor*. New York: Routledge, 1995, 266–269.

Ramus 2002: Ramus, F. Language discrimination by newborns. – *Annual Review of Language Acquisition* 2, 2002, 85–115.

Ramus et al. 2000: Ramus, F., M. D. Hauser, C. Miller, D. Morris, J. Mehler. Language discrimination by human newborns and cotton-top tamarin monkeys. – *Science* 288, 2000, 349–351.

Ramus, Mehler 1999: Ramus, F., J. Mehler. Language identification based on suprasegmental cues: A study based on resynthesis. – *The Journal of the Acoustical Society of America* 105(1), 1999, 512–521.

Rathcke et al. 2024: Rathcke, T., E. Smit, Y. Zheng, M. Canzi. Perception of temporal structure in speech is influenced by body movement and individual beat perception ability. – *Attention, Perception and Psychophysics*, Springer US, 2024, 1-17.

Russo, Barry 2008: Russo, M., W. J. Barry. Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. – In: Proceedings of Speech Prosody, Campinas, 2008, 419–422.