Technical University of Berlin
M. Sc. in Audio Communication & Technology

# Exposé

## An investigation of musical expertise on the neurological processing of the spoken language rhythm

**Canberk Turan**
c.turan@campus.tu-berlin.de

Supervisor: Dr. Athanasios Lykartsis

## 1 Introduction

Rhythm is one of the foremost mechanisms for the human brain to translate a piece of auditory information. It is providing the link to the rhythmic oscillations in brain activity which makes it crucial for verbal learning and language development. As is known, phone numbers are preferably separated into shorter groups to make it easier to be remembered with its evenly paced sequence. This is because human brain is more prone to process rhythmic grouping patterns. In this regard, speech has inherently a serial structure that makes it possible for us to employ the knowledge established in the field of musical rhythm analysis. This leads us to the idea of incorporating the recently developed metrics of speech periodicity with a comprehensive study that encompasses both neurobiological and behavioral aspects of language perception. The principal focus of the thesis is to elucidate the relationship between the listener's evaluation of the certain speech stimuli from a rhythmical point of view and rhythmic qualities that are analytically obtained from the respective audio signals. These parameters, which are also called as audio features, are approached in a similar sense with the conventional descriptive concepts of MIR (Music Information Retrieval) domain. In addition, the relevance of musical skills will be investigated in the context of perceptual evaluation of the speech rhythm by running listening experiments on groups with varying music background.

While researchers in many fields are engaged in rhythm, there is little awareness of how auditory expertise shapes these mechanisms. This project builds a bridge across multiple scientific disciplines; the focus lies in the validation for impacts of music aptitude, conceptualizing speech rhythm as a calculable neurobehavioral trait. We anticipate results that could inspire future training methods to develop general auditory skills for children who have difficulty in learning language and efficient therapies for aphasia patients needing rehabilitation.

## 2 Literature review

Speech and music are both evident on time, unlike multiple visual art forms and communication modalities [1]. They both are composed by temporal intervals that possibly differ in duration and acoustic marking by different spectral properties leading to metrical anticipations as 'beats'. This notional relation has been the topic of much dialogue as it leads to new findings about the natural evolution of human cognitive skills. Therefore works in this multidisciplinary field spans a range of methodological techniques, including neuroimaging [2], psychophysics [3], and traditional behavioral experiments as well as studies of animal synchronization[4], engaging an evolutionary approach on human rhythmic perception and cognition.

Rhythm can be defined as a systematic temporal repetition of certain events, whereby these repetitions themselves become temporal units. In the context of a linguistics, the rhythm of a particular language is always associated with the pattern of its speech phenomena, which are characteristic of that specific language. These repetitions can occur as a result of the interaction of serially ordered segments(e.g. consonants and vowels) or suprasegmental levels that are not limited to single sounds but often extend over syllables, words, or phrases[5]. That generates a pattern of time intervals within a sequence that constitutes the quasi-rhythmic acoustic cues which could latter be examined in a metric fashion. An early and influential theory of speech rhythm argued that the languages fell into rhythm classes [6]. According to this paradigm, in stress-timed languages such as English and German, stressed syllables were claimed to occur at regular temporal intervals, whereas in syllable-timed languages such as Italian and French, syllable onsets claimed to be evenly timed. The rhythm class hypothesis was criticized as in [7] with being an overly discrete categorization and limited applicability. Nonetheless, it encouraged other studies for the development of certain measurement methods based the consonant/vowel segmentation and variation [8, 9]. This quantification attempts contributed in conjunction with many neuroimaging and behavioural experiments, yet it has failed occasionally to give globally consistent and robust results as discussed in [10].

Behavioral studies on rhythm perception mostly favored to measure the abilities with motor response tasks such as finger tapping in which subjects were observed while they perform certain physical actions synchronized to given stimuli [11]. Furthermore, various studies involved subjects' judgments on the deviation from isochrony based on single events as well as the tempo changes occurring on an entire auditory sequence [12].

On one hand, most of these approaches had certain drawbacks considering the difficulty level of the tasks demanded from the test persons. In case of lacking cognitive and memory skills required for following instructions and beat production, beat perception induced by the stimulus only is altered by the quality of the task performance [13]. On the other hand, musical forms of stimuli have primarily been more under the spotlight of most studies as opposed to long linguistic sequences, since it consists of noticeably periodic pulses - beats [14], which is preferred for motor skill activities and speech has a relatively irregular temporal structure. This also brings up the question about how speech can possess abstract periodic patterns and still induce perceptual regularities in human brain[15]. Moreover, it is also reported that listeners who grow up in a musical environment with temporally irregular beats are capable of synchronizing with this asymmetric sound structure. This indicates that synchronizing with or detecting periodicities to a large extent depends on familiarity and enculturation rather than on physically measurable periodicity [16].

Studying beat perception in the context of neural entrainment and resonance has received an enthusiastic response from the scientific community [17], as it has notably has the potential to deliver a neural marker of beat perception. Previous neuroscientific research has also remarked a shared mechanism for both music and speech processing that takes place in anatomically overlapping brain structures and highlighted that supplementary motor area [18] as a crucial brain region involved in the temporal processing of speech. Latter studies on the processing of linguistic rhythm [2] have revealed that the human brain responds to the interchange of strong and weak beats under certain conditions, displaying a similar processing pattern to musical stimuli. However, the validity of the idea that assumes brain signals are representing the stimulus envelopes faithfully, is also a topic of discussion. In [19], it has been made evident that frequency-domain representations of rhythms are dissociable from beat perception. Therefore, it is recommended to combine EEG measurements of neural signals with creative behavioral paradigms is of more benefit to our understanding of beat perception.

In this regard, with the rapid expansion of its commercial application areas, rhythm analysis in the realm of Music Information Retrieval stands out as a promising tool for temporal analysis of higher-level speech elements mentioned above. Motivated by the urge to classify genres and detect languages [20], statistical and characteristic representations such as the Beat Histogram and the Beat Spectrum raised the potential of extracting audio descriptors from more complex patterns [21, 22]. The use of the novelty functions has also served as an important source of information for temporal behaviour of onsets in a signal. [23] introduced multiple calculation methods based on the frequency shifts, phase irregularity and amplitude variations. Lykartsis extended these methods by specifying the relevant signal quantities and illustrating their temporal trajectories as most explanatory novelty functions. Relying on satisfactory results achieved in the framework of language identification tasks, we believe it could provide as a foundation for further behavioral experiments and shed light on the ambiguities of previously proposed rhythm typology [24].

# 3 Method

Although this topic extends from neural fundamentals of timing to novel derivations of linguistic and musical metrics, it is intended to give an overview of the current state of music-psychological research on natural speech comprehension as well as it is correlation with the rhythm's analysis.

It is still not clear, if there is an influence of musical expertise on the processing of speech rhythm. In part of the EU-funded project "The NEurobiology of RHYthm: effects of MUSical expertise on natural speech comprehension" (NERHYMUS), it is investigated how rhythm expertise of musicians may affect the way they process the rhythm of spoken language. In order to achieve this, both musicians and non-musicians will go through a variety of tests: online questionnaire for quantification of musical experience, behavioral measures of beat perception in language and EEG. Whilst these tasks are conducted at the Maastricht University - Cognitive Neuroscience department, the infrastructure of the behavioural experiment and data analysis will be carried out within the framework of this thesis.

The guidelines of the experiment concept as well as the stimuli material are provided by team at the Maastricht University. The listening experiment is designed accordingly using the well-known Matlab package called Psychtoolbox. This free toolbox provides an exhaustive set of Matlab functions that are employable for neuroscience research offering a high level of control on auditory stimuli and the design of experiment interface. Later on, the technical maintenance of the experiment setup will be handled remotely.

The collected data will be processed together with the features extracted from the stimuli signals. For the signal analysis, techniques similar to the [24] will be implemented using conventional toolboxes such as MIRtoolbox and AKtoolbox. As opposed to calculating regularities between intervals of language prominence units, this novel technique suggests that the investigation of temporal quantities that are captured directly from the signal deliver more robust results. Since it combines extraction of signal-inherent periodicities and their properties, the process does not require any manual annotations and does not exhibit variability with respect to other non-rhythm-related abstract parameters.

Finally, the harmonisation between the experiment data and the acoustic features will be validated using mutual information toolbox, since it is capable of addressing non-linear dependencies between two variables .Thereby, it will be possible to detect causal links rather than simple associations[25]. Additionally, upon project team's agreement, the analysis may be expanded with the inclusion of the previous neuroimaging (EEG) data.

# 4 Timeline

| **Gantt chart** | Jan 20 | | | | Feb 20 | | | | Mar 20 | | | | Apr 20 | | | | Mai 20 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Literature survey | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | |
| Listening experiment | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | | | |
| Data handling | | | | | | | | ■ | ■ | ■ | | | | | | | | | | |
| Signal processing | | | | | | | | ■ | ■ | ■ | ■ | ■ | | | | | | | | |
| Statistical analysis | | | | | | | | | | ■ | ■ | ■ | ■ | | | | | | | |
| Writing | | | | | ■ | ■ | ■ | | | | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |

The expected outline for the master's thesis will include the following points:

- Introduction
  - Background of temporal perception and rhythm
  - Motivation and social aspects
  - Project concept (Nerhymus)
- Theory
  - Speech rhythm description
  - Neurological processes
  - Behavioural experiment
- Method
  - Experimental design
  - Data handling
  - Feature extraction
- Results
  - Mutual information analysis
  - Discussion
  - Future work

# References

[1] P. Fraisse, "Perception and estimation of time.," *Annual review of psychology*, vol. 35, pp. 1–36, 1984.

[2] K. D. Kandylaki, K. Henrich, A. Nagels, T. Kircher, U. Domahs, M. Schlesewsky, I. Bornkessel-Schlesewsky, and R. Wiese, "Where is the beat? the neural correlates of lexical stress and rhythmical well-formedness in auditory story comprehension," *Journal of Cognitive Neuroscience*, vol. 29, no. 7, pp. 1119–1131, 2017.

[3] R. Vanrullen and J. Dubois, "The psychophysics of brain rhythms," *Frontiers in Psychology*, vol. 2, p. 203, 2011.

[4] C. ten Cate, M. Spierings, J. Hubert, and H. Honing, "Can birds perceive rhythmic patterns? a review and experiments on a songbird and a parrot species," *Frontiers in Psychology*, vol. 7, p. 730, 2016.

[5] J. London, *Hearing in Time: Psychological Aspects of Musical Meter.* Oxford University Press, USA, 1 ed., 2004.

[6] D. Abercrombie, *Elements Of General Phonetics*, vol. 363. Edinburgh University Press, 1966.

[7] R. Dauer, "Stress-timing and syllable-timing reanalyzed," *Journal of Phonetics*, vol. 11, no. 1, pp. 51 – 62, 1983.

[8] F. Ramus, M. Nespor, and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, vol. 73, pp. 265–292, 12 1999.

[9] L. Low, E. Grabe, and F. Nolan, "Quantitative characterizations of speech rhythm: Syllable-timing in singapore english," *Language and Speech*, vol. 43, no. 4, pp. 377–401, 2000. PMID: 11419223.

[10] A. Turk and S. Shattuck-Hufnagel, "What is speech rhythm? a commentary on arvaniti and rodriquez, krivokapić, and goswami and leong," vol. 4, no. 1.

[11] J. Iversen and A. Patel, "The beat alignment test (bat): Surveying beat processing abilities in the general population," *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10)*, 09 2008.

[12] S. Fujii and G. Schlaug, "The harvard beat assessment test (h-bat): A battery for assessing beat perception and production and their dissociation," *Frontiers in human neuroscience*, vol. 7, p. 771, 11 2013.

[13] J. Grahn and M. Brett, "Impairment of beat-based rhythm discrimination in parkinsonŁs disease," *Cortex; a journal devoted to the study of the nervous system and behavior*, vol. 45, pp. 54–61, 11 2008.

[14] W. E. Benjamin, "A theory of musical meter," vol. 1, no. 4, pp. 355–413.

[15] M. Schmidt-Kassow and S. Kotz, "Attention and perceptual regularity in speech," *Neuroreport*, vol. 20, pp. 1643–7, 11 2009.

[16] E. E. Hannon and S. E. Trehub, "Metrical categories in infancy and adulthood," *Psychological Science*, vol. 16, no. 1, pp. 48–55, 2005. PMID: 15660851.

[17] E. W. Large, "Resonating to musical rhythm : Theory and experiment," 2008.

[18] J. A. Grahn and M. Brett, "Rhythm and beat perception in motor areas of the brain," *Journal of Cognitive Neuroscience*, vol. 19, no. 5, pp. 893–906, 2007.

[19] M. J. Henry, B. Herrmann, and J. A. Grahn, "What can we learn about beat perception by comparing brain signals and stimulus envelopes?," vol. 12, no. 2, pp. 1–17.

[20] J. J. Burred and A. Lerch, "Hierarchical automatic audio signal classification," *Journal of the Audio Engineering Society (JAES)*, vol. 52, pp. 724–739, 07 2004.

[21] G. Tzanetakis, "Tempo extraction using beat histograms," 01 2005.

[22] J. Foote and S. Uchihashi, "The beat spectrum: A new approach to rhythm analysis," vol. 22, 08 2001.

[23] J. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, pp. 1035 – 1047, 10 2005.

[24] A. Lykartsis and S. Weinzierl, "Using the beat histogram for rhythm description and language identification," 09 2015.

[25] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," 2007.