Music and Choreography Metaphors in Spoken Language Rhythm Modelling and Their Application to Computer-Assisted Pronunciation Training for Mora-Timed Japanese

Evgeny Pyshkin, University of Aizu, Japan
Akemi Kusakari, National Institute of Technology, Hachinohe College, Japan
Natalia Bogach, Peter the Great St. Petersburg Polytechnic University, Russia
John Blake, University of Aizu, Japan

The Kyoto Conference on Arts, Media & Culture 2023
Official Conference Proceedings

Abstract

The article describes the development and assessment of a computer-assisted pronunciation training (CAPT) environment to improve prosody practice focusing its particular setup for Japanese. Users interact with the system by replicating model utterances, while the system feedback is provided in the form of pitch curves contrasting the learner's pronunciation against the model. Complementing the pitch graphs with the audible output and quantitative metrics provide objective evaluation of pitch quality. However, we observed that learners of Japanese, a mora-timed language, faced significant challenges in understanding and applying the feedback effectively, since Japanese rhythmic patterns are not encoded well in the pitch curves. To address this, we explore the use of additional instructive feedback models such as musical and even choreographic notations. Along with rhythmically enhanced phonetic transcription, these interfaces have sufficient flexibility to include syncopations and silence. Music notation has standard features, which can represent the important components of language rhythm including the time signatures capturing the assigned rhythm for a phrase, the notes symbolizing variations in pitch, and the timed rests that can be helpful in visualizing the mapping of the phrase to rhythmic units. We share practical CAPT scenarios supporting pronunciation exercises in Japanese. In particular, we use examples of Yodel music to model the phrases characterized by complex language rhythmic patterns. Vocal music can be helpful and inject fun into the learning process. We also noticed that rhythmic structure of utterances can be described with dance movement patterns, for example using known graphic dance notations.

Keywords: Speech Visualization, Language Rhythm, Pronunciation Training, Music Notation, Choreographic Notation, Cross-Disciplinary Study

Introduction

Present-day intelligent language learning environment are complex multiaspect and cross-disciplinary systems involving the contributions from various areas of knowledge and expertise, naturally including language instruction, applied linguistics, psychology, and phonology, but also the recent technology achievements enabled by speech processing algorithms, human-machine interfaces, multimedia technology, mobile and communication solutions, to mention a few (Figure 1). Instruments delivered by the technology advances created new learning possibilities, which do not simply engage the digitization of learning materials but create new scenarios, which are impossible to support without using intelligent technologies of knowledge representation management.

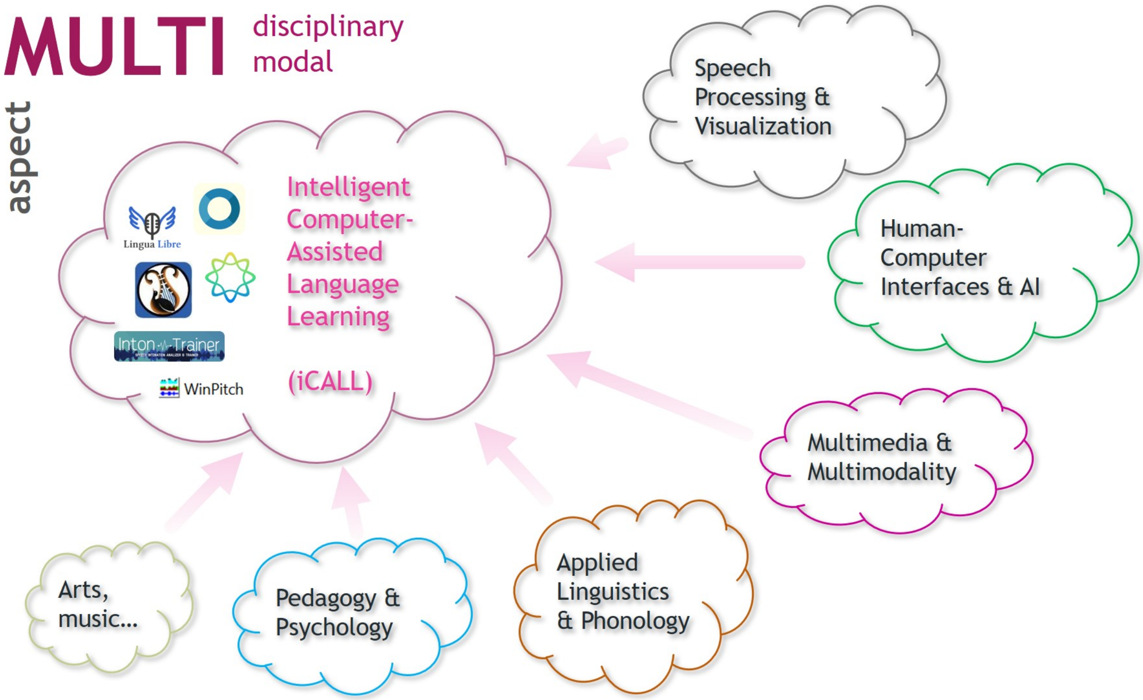


Figure 1: A variety of disciplines contributing to iCALL.

In addition to the symbiotic connection between non-technical education and computer and software engineering, achieving the higher levels of multimodality of learning system, improving descriptive and instructive feedback production, and their better personalization can be significantly enhanced through leveraging the user's personal experience beyond the pure scope of language learning. In this perspective, models and metaphors borrowed from arts, music, and literature can provide a novel more engaging learning experience. By integrating these disciplines into computer-assisted pronunciation training (CAPT), we cater to different learning styles and cognitive preferences, allowing learners to interact with the material in more personalized and intuitive ways. The incorporation of familiar aesthetic and rhythmic structures from the arts may serve as mnemonic devices, aiding in the retention and reproduction of linguistic patterns, such as prosody. This interdisciplinary approach not only enriches the language learning landscape but also helps to foster a deeper appreciation for the intricacies and interconnectedness to the fundamental principles of art and music. This innovative fusion, therefore, holds promise for more effective and enjoyable language acquisition experiences.

Intelligent computer-assisted pronunciation training (iCAPT) environments represent a particular class of iCALL systems with their specific challenges. CAPT feedback personalization, as well as the problems of training and evaluating the pronunciation remain an important topic in the scope of language instruction research.

In process of developing CAPT environment based on pitch digital processing (Pyshkin et al., 2019), and its multilingual setup for using for prosody practice in languages from different phonology groups, we discovered the numerous challenges in adequate modeling of language rhythm required to tailor the multimodal feedback to language learners, particularly for syllable-timed and mora-timed languages. The following text examines possible ways to enhance the CAPT system feedback at both descriptive and instructive levels through the incorporation of music- and choreography-based explanations to the CAPT exercise interfaces, to address the above-mentioned challenges.

Nowadays, implementing a successful learning environment often assumes exploring the indirect connections between the disciplines, incorporating gamification scenarios into the learning workflow, using technology solutions enhancing the feedback from the learning system to its users as ironically depicted in the famous scene of George Cukor's *"My Fair Lady"* movie with Audrey Hepburn and Rex Harrison, where Prof. Higgins used a candle for evaluating the correct pronunciation of a plosive sound (Figure 2). This candle is a humorous and apt metaphor of the cross-disciplinary application of engineering technology to pronunciation training; in a sense, that device is a prototype of a technology-assisted pronunciation training system.

Figure 2: Pronunciation training feedback can involve engineering devices.

In our research, we are trying to bring music and even choreography models to language learning process. Indeed, music and even choreography are arts, where rhythm stands for an essential systematic feature, therefore, applying music and choreography metaphors for modeling and visualization of mora-timed rhythmic patterns can complement other known ways of pitch visualization, so that to address such language features, which are usually out of possibilities in traditional pronunciation training systems, including syncopations, necessary periods of silence, superimposition of mora-timed phrase to a characteristic rhythmic template, all necessary for meaningful spoken language interpretation.

StudyIntonation CAPT Environment at a Glance

The testbed for the current study is provided by *StudyIntonation*, a pronunciation training environment developed in the frame of ongoing project on using signal and speech processing algorithms applied to the construction of prosody-based CAPT system (Bogach et al., 2021). Figure 3 depicts a simplified system architecture that illustrates the flow of interaction and data through various components of the CAPT system. Starting from the learner's interface, the diagram likely shows the process by which the learner's spoken input is captured, processed, and analyzed.

The main interface of the system accessed by language learners is provided by the mobile application, enabling the user's access to the pronunciation exercises, each of them is supported by the plotted pitch graphs representing both the model utterance recorded by native speakers and the user's attempts visually contrasted against the model. Pitch visualization and its further evaluation is possible because of using digital signal processing algorithms and dynamic time warping for pitch curve analysis, the latter enabling constructing pitch similarity metrics.

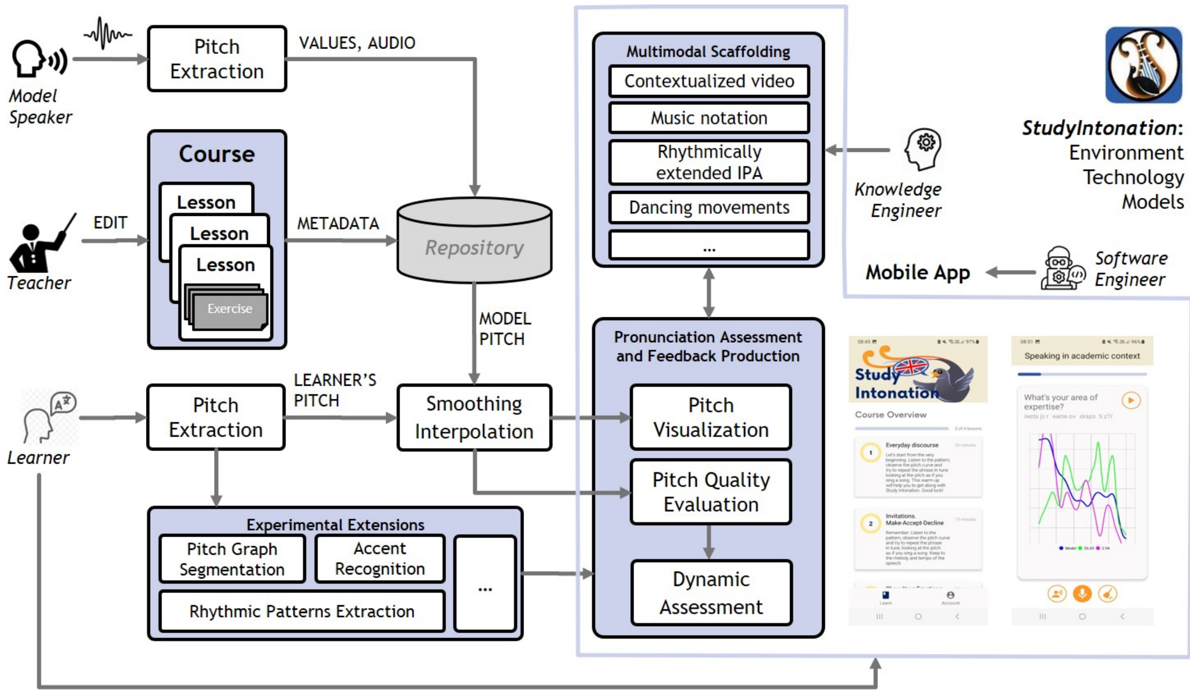


Figure 3: StudyIntonation CAPT system architecture.

As described in our earlier work (Mikhailava et al., 2022), there are many practical ways to harness a higher level of multimodality through tailoring both the CAPT interface and CAPT feedback production, including support for contextualized attitudinal exercises, using model video, learner's recorded pitch reproduction; and, the most important in the scope of this study, finding better ways for utterance rhythm visualization, and rhythm-based pronunciation improvement recommendations. Though pitch curves, phonetic transcription, and pitch similarity metrics all together provide an objective intonation visualization, they lack an instructive value and do not completely address the aspects of CAPT personalization and support of variety of learning styles.

Music Notation as an Instrument to Model Language Prosody

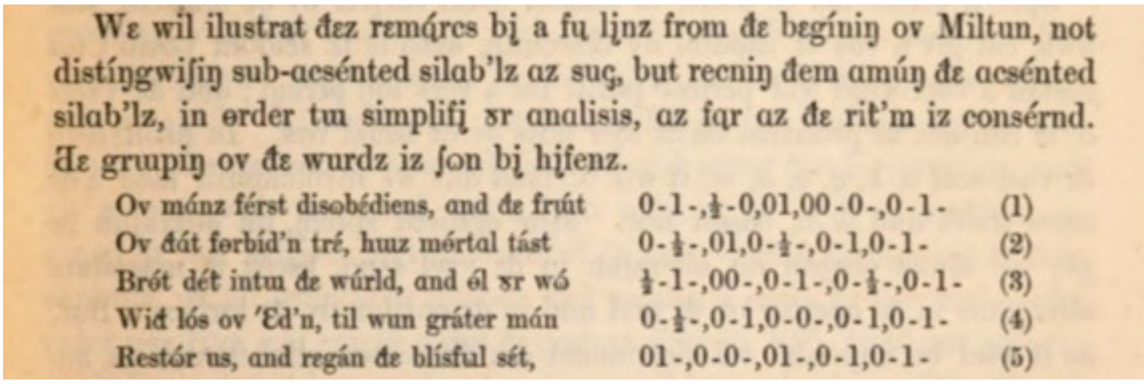
Using graphic representations similar to music notation in the application domains other than music itself is not a novel concept. The mathematics of music provides a formalized language that meticulously structures timing and rhythm. As brilliantly pointed by John Thelwall as far as in 1812, *"Nothing is in present better ascertained or more clearly demonstrated, than the* *mathematical proportion of the bars of music"* (Thelwall, 1812). Etienne-Jules Marey, the French scholar and physiologist (who was one the firsts to accommodate a music-like notation to describe movements), admired music notation (Marey, 1868) as a universal language which is a graphic expression of *"very fugitive, delicate, and complex movements, no language could express"* (translated from French – E.P.). "Old times" considerations of Thelwall and Marey are in harmony with more recent research on applying music models to mora-timed language constructions: *"Radically divergent contexts can share similar musical structures. As musicians know, feeling the sense of rhythm, and sharing rhythmic structures from beyond one's own shores creates a bridge across languages and cultures"* (Gilbert & Yoneoka, 2000).

Back to History

In Thelwall's work (1812), the concept of cadence (a central component of music harmony) was extended for its application to language prosody:

> *I lay it down, therefore, as the first principle and basis of all rhythmical theory and analysis, and of all instruction for the improvement of human utterance and composition,- that a cadence; is a portion of tuneable sound, beginning heavy and ending light;-Secondly, that a foot is a syllable, or number of syllables, occupying the space or duration of such a cadence.* (Thelwall, 1812)

In turn, in his *"The Essentials of Phonetics"* (Ellis, 1848), John Ellis defined a formalized model to describe a spoken language rhythm with (using contemporary computing definition) a temporal ternary model for English rhythmic pattern definition (Figure 4).


Wɛ wil ilustrat đɛz rɛmárcs bi̧ a fu̧ li̧nz from đɛ bɛgíniŋ ov Miltun, not distíŋgwiʃiŋ sub-acsénted silab'lz az suç, but recniŋ đem amúŋ đɛ acsénted silab'lz, in ɵrder tu̧ simplifi̧ ɜr analisis, az fa̧r az đɛ rit'm iz consérnd. Đɛ grupiŋ ov đɛ wurdz iz ʃon bi̧ hi̧fenz.

Ov mánz férst disobédiens, and đɛ frút	0-1-,½-0,01,00-0-,0-1-	(1)
Ov đát fɵrbíd'n tré, huz mórtal tást	0-½-,01,0-½-,0-1,0-1-	(2)
Brót dét intu̧ đɛ wúrld, and ól ɜr wó	½-1-,00-,0-1-,0-½-,0-1-	(3)
Wiđ lós ov 'Ɛd'n, til wun gráter mán	0-½-,0-1,0-0-,0-1,0-1-	(4)
Restór us, and regán đɛ blísful sét,	01-,0-0-,01-,0-1,0-1-	(5)

Figure 4: Rhythmic patterns from *"The Essentials of Phonetics"* (Ellis, 1848).

Figure 5 is composed from two examples of simplified music notation for modeling quite complex declamation rhythmic patterns that we can find in the amazing works of William Thomson (Thomson, 1904).

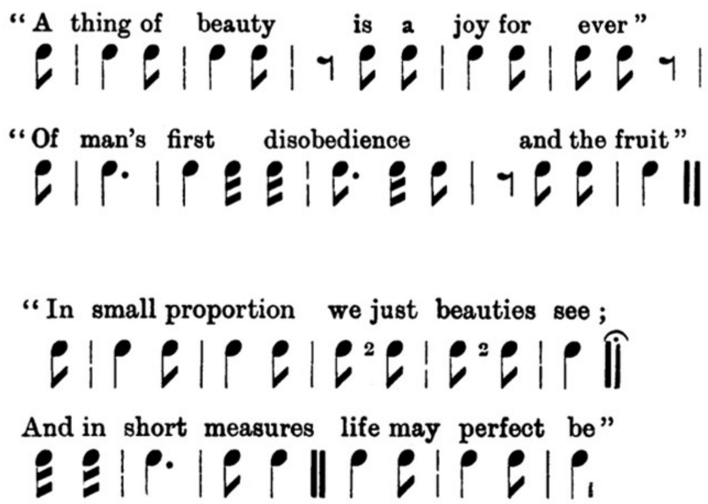


Figure 5: Simplified music notation for language rhythm modeling (Thomson, 1904).

The historical insights into the musicality of language resonate with the aims of modern CAPT systems. Applying music notation to language prosody in CAPT thus taps into time-tested methods, providing learners with a familiar visual and auditory schema that can capture the subtleties of pronunciation from phonetics to prosody. This blending of tradition with technology showcases a multidisciplinary strategy, marrying historical language theories with contemporary learning solutions in a concise, effective teaching methodology.

Mapping of Rhythmic and Syllable Patterns

For mora-timed languages such as Japanese, understanding spoken language rhythmic divisions of the utterance into portions (known as language isochrony) is one of the critical elements of spoken language proficiency. Japanese poetry could be considered as an excellent natural source for understanding the rhythmic formatting of spoken Japanese. Many know that traditional Japanese *haiku* is composed based on regular syllable-based model, where 17 syllables (mora) are split across three small portions of poetic text with the first one of five, the second one of seven, and the third one of five syllables, thus making the 5+7+5=17 mora structure. Fewer know that for the purposes of authentic spoken language interpretation and haiku translation, one needs to understand that the mora patterns are not the same as rhythmic patterns (Gilbert & Yoneoka, 2000). In fact, to assure a correct haiku reading, 17-mora verse must be mapped to a 24-unit three bars four beat rhythmic pattern as Figure 6 illustrates using the example of famous Basho's "Frog" haiku: *"An old pond – / A frog dives in / Water sound"* as translated in (Basho et. al., 2006). Only because of such mapping, the rests and syncopations can be respected for appropriate meaningful reading.

Interestingly, though the effect of haiku reading for learning Japanese rhythmic patterns is quite obvious, the implications of Japanese haiku reading can be helpful not only for learners of Japanese language, but also for English language education, especially by Japanese learners. Thus, there is a definitive cross-cultural context (Kono, 2011).

Figure 6: Haiku: mapping a 17-mora verse to 24-unit rhythmic pattern.

Modeling Japanese Rhythmic Patterns: Combining Music Notation With Rhythmically Formatted IPA

The above mentioned rhythmically enhanced phonetic transcription introduced by Ellis (Ellis, 1848) is one of first examples of using additional rhythmic formatting of language prosody. While using common phonetic models (such as IPA), many Japanese language textbooks still lack a good instrument to introduce the rhythmic patterns. Computerized language learning systems could serve as research environments to experiment with different approaches to enhance traditional phonetic transcription with rhythmic information, for example, as shown in Figure 7.

Kana (Phonology feature)	Extended IPA Pitch within the pendulum rhythm	Time signature
おはようございます (mora)	(o-ha)-(jo-ː-go)-(dza-i)-(ma-sɯ̥)-(●-●)-(●-●) (LH)-(HHH)-(HH)-(HL)-(●●)-(●●)	3/4
こんにちは (mora)	(ko-n)-(ɲi-tʃi)-(βa-●)-(●-●) (LH)-(HH)-(H●)-(●●)	4/4
こうえいです (long vowel)	(ko-ː)-(e-ː)-(de-sɯ̥)-(●-●) (LH)-(HH)-(HL)-(●●)	4/4
しゅっぱついたします (staccato)	(ʃɯ-p-pa-tsɯ)-(i-ta-ʃi-ma)-(sɯ̥-●-●-●) (LHHH)-(LHHH)-(L●●●)	4/4

Figure 7: A sketch of an approach to extend IPA with rhythmic information.

Such enhancements can naturally be presented along pitch graphs, model speech video recording, and music notation, the latter can be more beneficial for those users having at least elementary music background (as it can be seen from the *StudyIntonation* mobile app screenshots for Japanese pronunciation training presented in Figure 8).

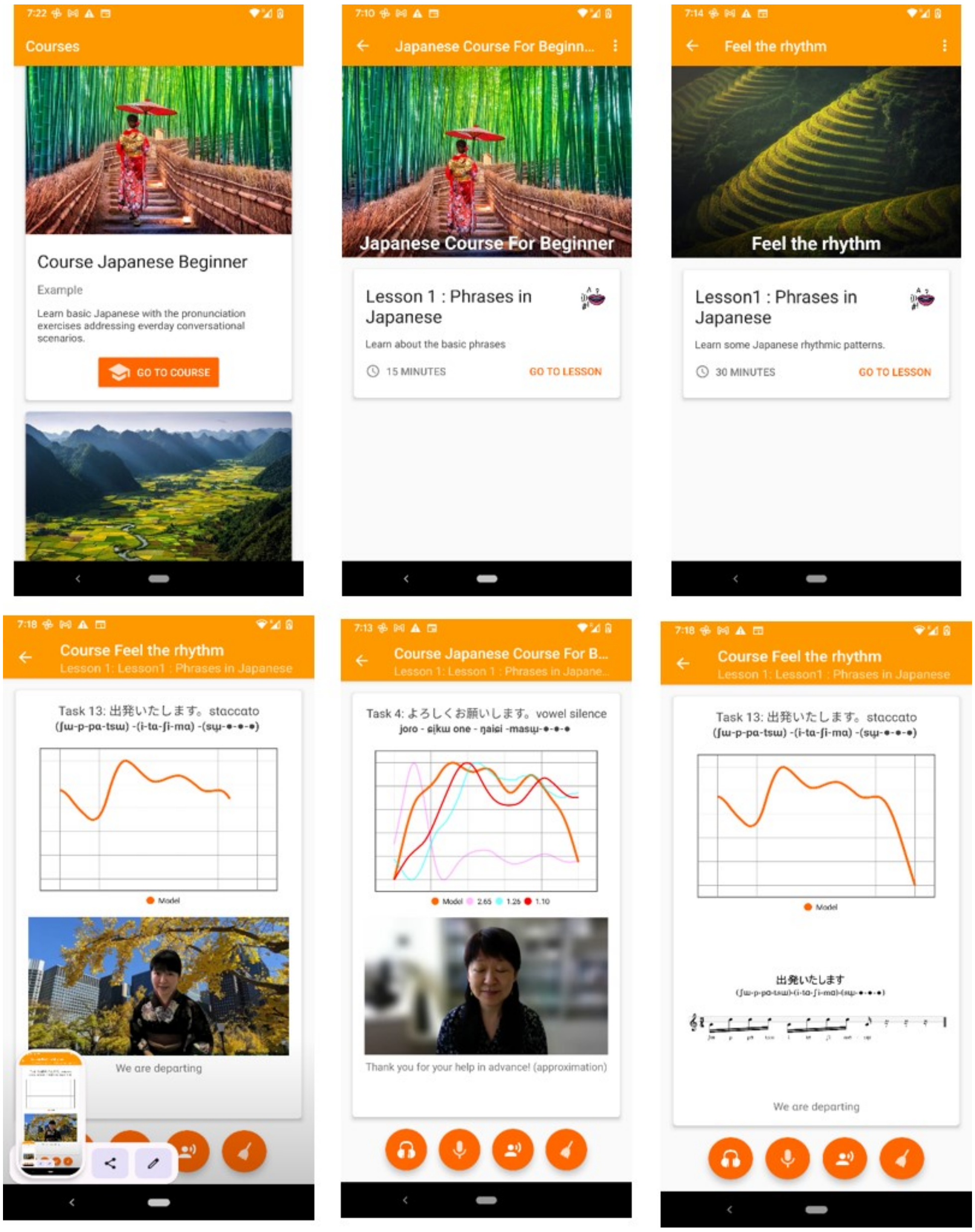


Figure 8: *StudyIntonation* mobile app interface for learning Japanese.

Music notation has standard features, which can be helpful in representing the important components of language rhythm including the time signatures that can naturally represent the constant rhythm assigned to a phrase, the notes that can represent pitch higher and lower tones, and the timed rests that can be helpful in visualizing the mapping of the phrase to rhythmic units (Pyshkin et al., 2023). As we demonstrate with the examples of Yodel music (Figure 9) used to model the phrases characterized by complex language rhythmic patterns, vocal music can be helpful and can also inject fun into pronunciation training.

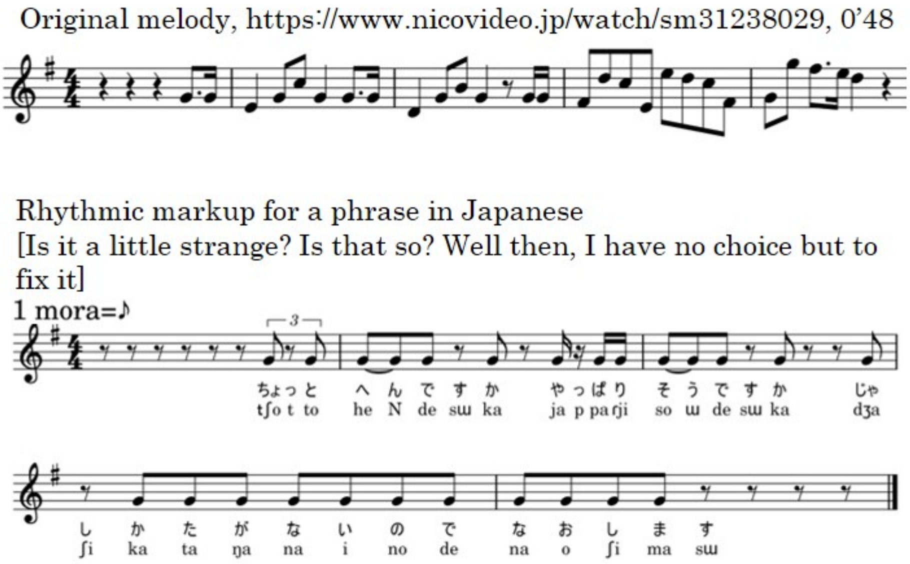


Figure 9: Constructing complex rhythmic patterns using Yodel music.

Breath control, an often-overlooked aspect in language learning, is crucial for mastering the ebb and flow of a language's rhythm. By engaging with the breathing patterns intrinsic to singing, learners can develop a more intuitive feel for the pauses and stresses within a language's melodic structure. This kinesthetic connection helps internalize the prosody of the target language, enabling learners to feel the melody in their speech.

Connecting Music Rhythmic Patterns to Choreography Elements

As singing can contribute to breathing control improvements, the similar effect can be reached by body movement exercises. Are there better body movements connected to musical patterns and rhythmic models rather than movements of dancing?! Indeed, why not to try to model the rhythmic structure of utterance with dancing patterns, represented in some way, for example using known graphic notations.

Dancing Patterns for Speech Rhythm Modeling: A Primer

Unlike integration music representations to pronunciation training exercises shown in Figure 8, incorporation of dancing patterns is not currently supported in our CAPT system and presented here as a potential future research direction. Exploring the kinaesthetic synergy between dance and speech, we posit that learners may benefit from the integration of dance movements to reinforce language rhythm and pronunciation. Based on Steiner's *"Lemons Lemons Lemons Lemons Lemons"* (Steiner, 2016), the Royal Ballet's 2023 production *"The Limit"* (Winship, 2023) nicely demonstrates how the spoken text can be explicitly included to the choreographic construction in a way that choreography shapes the text and creates a visual and meaningful reflection of a language rhythm. The possibility to apply this concept to language learning (thus, in opposite direction) pivots on the premise that learners could map the dynamic features of pronunciation—such as stress, intonation, and timing—to corresponding dance steps and sequences. By embodying the rhythm of speech through dance, learners may discover a more visceral understanding of the language's prosody. Dance, with its structured movement patterns, offers a tangible counterpart to the abstract rhythms of speech, providing a physicality to pronunciation that could enhance memory retention and intuitive learning.

For this discussion, we took the pronunciation exercise for utterance repetitions of "Good morning!" in Japanese. A possible way to help learners to feel the correct rhythmic model of a phrase is to encourage them to repeat the phase several times, while keeping the necessary rhythmically adjusted moments of silence between the repeated utterances. The rhythmics pattern of a single phrase becomes better understood after repeating the same phrase within one pronunciation exercise attempt, just as one repeats movements or movement sequences in dancing phrase. Figure 10 illustrates a possibility to map the spoken phrase rhythmic pattern to the repeated sequence of relatively simple ballet steps for Japanese "Good morning!". Indeed, why not to begin the day with a sequence of dance steps?

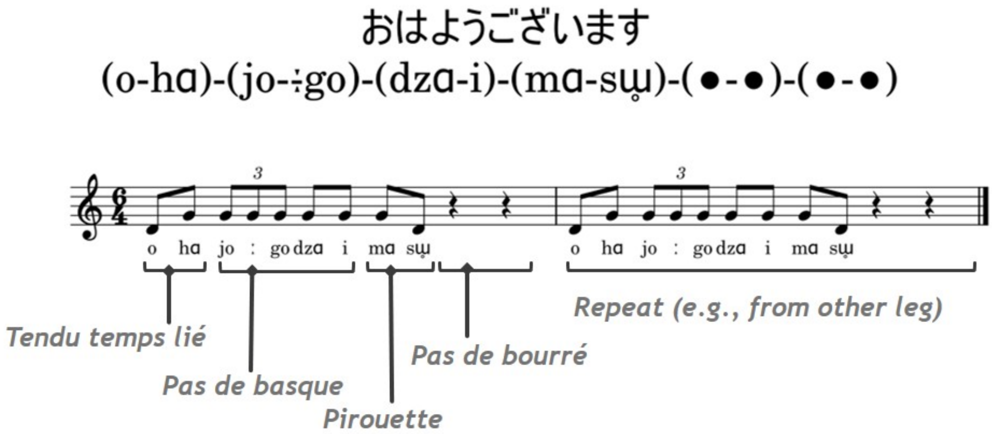


Figure 10: Following phrasal rhythm with ballet steps.

Dance Graphic Notations: Case Study

As soon as we want to introduce choreography elements to complement a music and phonetic rhythmic pattern, we need to find a suitable approach to enable such elements inside a pronunciation training application. Model video would be a perfect way to incorporate choreography patterns to the system, however the exercises could already be linked with multimedia frames including contextualized video and audio recordings, and the system itself is not assumed to be a dancing virtual instructor. Therefore, searching for kinds of schematically presented diagrams or simplified images looks like more affordable, from the perspective of its practical feasibility, as well as organizational, implementational, and (by and large) computational expenses. Such diagrams could be sketched using one of known graphic notations. Surprisingly, in the long history of dance notation, there are not so many practical approaches applicable to our case. Figure 11 draws on a timeline-mapped visualization of the brief history of such notations.

The ornamentally-beautiful model known as Beauchamp–Feuillet notation was apparently the first documented formalized system to describe Baroque dance movements (mostly footwork) and staging, where Pierre Beauchamp, the French choreographer, dancer, musician and composer states for a probable inventor (Powell, 1995), while Raoul Auger Feuillet, choreographer and dance theoretician, put the approach to practically usable description and implementation (Feuillet, 1701). Presenting the movements along with (and across) music lines was one of innovations to clarify the relation of the steps to the music.

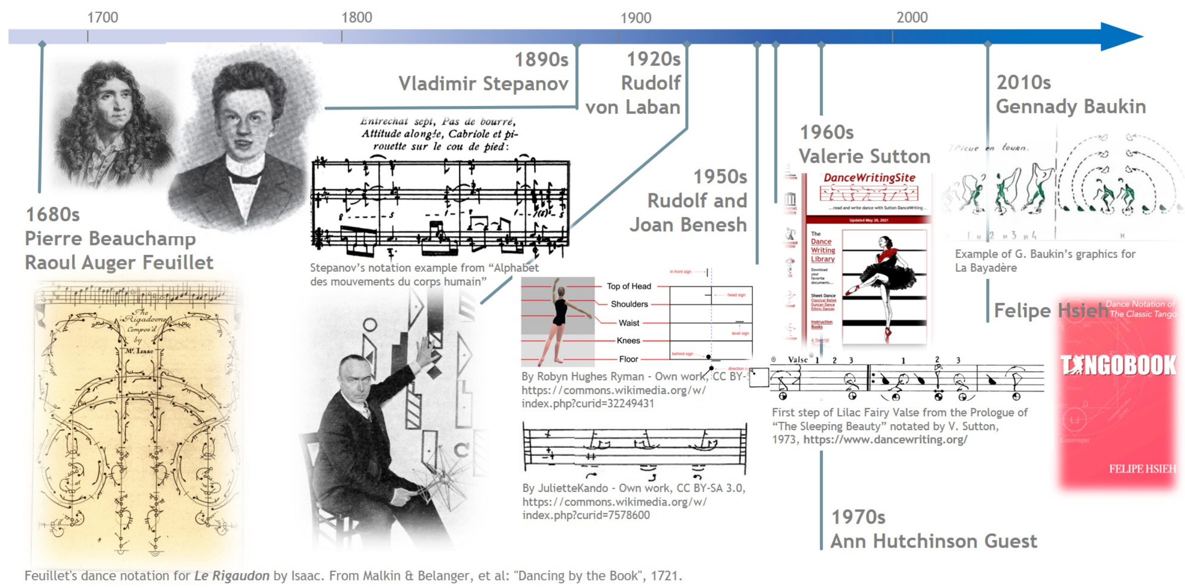


Figure 11: Ballet notations in time.

The next significant contributor to the development of systematic dance notations that must be mentioned in the context of our interests is Russian dancer Vladimir Stepanov (1866 – 1896), who died in the young age but succeeded in creation of the complex system for ballet notation (Stepanov, 1892), the method used and further improved by Alexander Gorsky who notated many classical ballets (mainly by Petipa). The collection of scores created by Gorsky after Stepanov's death and subsequently under the supervision of Nicholas Sergeyev, the Imperial Ballet director with the Mariinsky Theater in 1901 – 1917, is now known as Sergeyev collection (Sergeyev, 1944). The scores based on Stepanov notation are still in active use in modern times, both for completely new productions and reconstructions of Romantic ballet performances. Conceptually, Stepanov followed a novel idea to implement a dancing notation using music score-like graphic model, the concept formulated by Etienne-Jules Marey already mentioned above in the context of discourse on music notations, as far as in 1868 (Marey, 1868), thus creating the first system that, on the one hand, is based on the anatomical structure of the human body but at the same time, uses the graphic model similar to musical writing plays to describe specific dancing movements (Rousseau, 1995) mainly, those used in a classical ballet.

Probably, except the revolutionary notation suggested by Rudolf von Laban and defined in terms of spatial models and concepts rather than body movements modeling, most subsequence efforts to ballet or dancing graphics are based on body modeling, music score mapping, or tracking, including the contributions by Rudolf and Joan Benesh (Benesh & Benesh, 1977) and Valerie Sutton (Sutton, 1981).

Towards Integration With a CAPT System: Expectation and Challenges

In a computer-assisted environment it would be natural to investigate the possibilities to apply movement visualization techniques (Kyan et al., 2015) of virtual reality models, for example such as reported in (Anjos, Ribeiro & Fernandes, 2018). While these techniques excel in guiding precise physical movements, their application to pronunciation training requires overcoming the disjuncture between the kinesthetics of dance and the cadence of spoken language. CAPT systems must capture sound signals and represent perceived prosody rather than the gross motor skills in dance. This requires a specialized approach that goes beyond direct replication of existing computationally expensive models.

The notations briefly introduced here provide important insights to formalized systematic graphic languages for modeling dancing movement modeling. However, we must admit that in most cases a high level of expertise could be required to understand and decode those models, though the meaning of some graphic symbols for relatively simple elements can be understood by many. From the viewpoint of notation "decodability" by non-experts, the less-formal choreography drawing patterns sketched by Baukin in his "Ballet in graphics" (Baukin & Sorochinskaya, 2013) published in connection to La Bayadère full clavier score by Minkus (Minkus, 1877) could be more practical for representing dancing movements in the form that can be easier interpreted by many. However, from the perspective of practicality, there are obvious challenges, since to design such dancing movement sketches, the drawing skills are required in combination with some expertise in linguistics, as well as in music and choreography. Therefore, probably, there is good space for researching a possibility to use AI generated graphics (Guinness, 2023) and corresponding required studies in the emerging domain of AI engine prompt engineering (Liu & Chilton, 2022).

Conclusion

The integration of musical notation into CAPT systems has opened a novel avenue in pronunciation tuition, marrying the strictness of musical structure with the intricacies of spoken language. Drawing on the rich history of prosody research and pedagogy, exemplified by pioneers like Thelwall and Ellis, our approach utilizes the established, formalized language of music to convey the complex concepts, such as rhythm and pitch. The methodological use of musical notation extends beyond mere visual aids, serving as a bridge between the learner's cognitive understanding of language structure and their auditory perception of its musicality. This allows learners to perceive the tonal and rhythmic patterns of language through a musical score. Looking ahead, the exploration of choreography as a metaphor for language prosody presents a hitherto unexplored avenue of language research. By considering the potential of dance as a kinaesthetic framework for pronunciation training, we are on the cusp of broadening the scope of language learning tools to include not just the ears and eyes, but the entire body.

References

Anjos, R. K. D., Ribeiro, C., & Fernandes, C. (2018). Three-dimensional visualization of movement qualities in contemporary dance. In *Proceedings of the 5th International Conference on Movement and Computing* (pp. 1–7). https://doi.org/10.1145/3212721.3212812

Basho, M., Buson, Y., Issa, K., Saito, T., Nelson, W.R., & Sakaguchi, M. (2006). *1020 Haiku in translation: The heart of Basho, Buson and Issa.* BookSurge, 2006.

Baukin, G. & Sorochinskaya, L. (2013). *Ballet in graphics.* http://www.forumklassika.ru/showthread.php?t=95054

Benesh, R., & Benesh, J. (1977). *Reading dance: The birth of choreology.* London: Souvenir Press.

Bogach, N., Boitsova, E., Chernonog, S., Lamtev, A., Lesnychaya, M., Lezhenin, I., Novopashenny, A., Svechnikov, R., Tsikach, D., Vasiliev, K., Blake, J., & Pyshkin, E. (2021). Speech processing for language learning: A practical approach to computer-assisted pronunciation teaching. *Electronics, 10* (3), 235. https://doi.org/10.3390/electronics10030235

Ellis, A.J. (1848). *The essentials of phonetics.* London, F. Pitman. Princeton Prosody Archive. https://hdl.handle.net/2027/nyp.33433082185335

Feuillet, R. A. (1701). *Choregraphie: ou l'art de décrire la dance, par caracteres, figures et signes démonstratifs, avec lesquels on apprend facilement de soy-même toutes sortes de dances. Ouvrage tres-utile aux maîtres à dancer & à toutes les personnes qui s' appliquent à la dance.* Chez l'auteur et chez Michel Brunet.

Gilbert, R. & Yoneoka, J. (2000). From 5-7-5 to 8-8-8: An investigation of Japanese haiku metrics and implications for English haiku. *Language Issues: Journal of the Foreign Language Education Center, 3* (1).

Guinness, H. (2023). The best AI image generators in 2024. *Zapier blogs.* https://zapier.com/blog/best-ai-image-generator/

Kono, M. (2011). “Haiku” in English education. *Akita Journal on English Literature,* 53, 11–16.

Kyan, M., Sun, G., Li, H., Zhong, L., Muneesawang, P., Dong, N., Elder, B., & Guan, L. (2015). An approach to ballet dance training through MS Kinect and visualization in a CAVE virtual reality environment. *ACM Transactions on Intelligent Systems and Technology (TIST), 6*(2), 1–37. http://dx.doi.org/10.1145/2735951

Laban, R. (1975). *Laban’s Principles of Dance and Movement Notation.* 2nd ed. edited and annotated by Roderyk Lange. London, MacDonald and Evans.

Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–23).

Marey, E. J. (1868). *Du mouvement dans les fonctions de la vie: Leçons faites au Collège de France.* Germer Baillière.

Mikhailava, V., Pyshkin, E., Blake, J., Chernonog, S., Lezhenin, I., Svechnikov, R., & Bogach, N. (2022). Tailoring computer-assisted pronunciation teaching: Mixing and matching the mode and manner of feedback to learners. *INTED-2022* (pp. 767–773). IATED. https://doi.org/10.21125/inted.2022.0263

Minkus, L. (1877). La Bayadère (The Temple Dancer) after Kalidasa's Sanskrit play. Ballet in 4 Acts. Ed. Anna Marie Holmes.

Powell, J. S. (1995). Pierre Beauchamps, Choreographer to Molière's Troupe du Roy. *Music & Letters, 76* (2), 168–186. http://www.jstor.org/stable/737729

Pyshkin, E., Blake, J., Lamtev, A., Lezhenin, I., Zhuikov, A., &. Bogach, N. (2019). Prosody training mobile application: Early design assessment and lessons learned. *10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems* (pp. 735–740). https://doi.org/10.1109/IDAACS.2019.8924359

Pyshkin, E., Kusakari, A., Blake, J., Pham, N.B., & Bogach, N. (2023). Multimodal modeling of the mora-timed rhythm of Japanese and its application to computer-assisted pronunciation training. *14th International Conference on Learning Technologies and Learning Environments (LTLE 2023) as an event of 14th IIAI International Congress on Advanced Applied Informatics,* Koriyama. In press.

Rousseau, P. (1995). Figures de déplacement. L'écriture du corps en mouvement. *Exposé, REAC.* http://editions-hyx.com/sites/default/files/public/media/expose_2_p.rousseau.pdf

Sergeyev, N. (1944). *Dance Notations and Music Scores for Ballets, 1888-1944* (MS Thr 245). Harvard Theatre Collection, Houghton Library, Harvard University. https://id.lib.harvard.edu/ead/hou01987/catalog

Steiner, S. (2016). *Lemons Lemons Lemons Lemons Lemons.* Nick Hern Books.

Stepanov, V. I. (1892). *Alphabet des mouvements du corps humain: essai d'enregistrement des mouvements du corps humain au moyen des signes musicaux.* Librarie Paul Vigot, Paris.

Sutton, V. (1981). Sutton movement writing & shorthand. *Dance Research Journal, 14* (1-2), 78–85.

Thelwall, J. (1812). *Illustrations of English rhythms.* London, McCreery. Princeton Prosody Archive. https://hdl.handle.net/2027/nnc1.1002383080

Thomson, W. (1904). *The basis of English rhythm.* Glasgow, W.R. Holmes. Princeton Prosody Archive. https://hdl.handle.net/2027/uc1.$b31655

Winship, L. (2023). The Limit review – oppressed pair make words and steps count. *The Guardian,* Oct 24. https://www.theguardian.com/stage/2023/oct/24/the-limit-review-royal-opera-house.

Contact email: pyshe@u-aizu.ac.jp