Interspeech 2024
1-5 September 2024, Kos, Greece

Collecting Mandible Movement in Brazilian Portuguese

Donna Erickson [1], *Albert Rilliard* [2], *Malin Svensson Lundmark* [3], *Adelaide Silva* [4],
Leticia Rebollo Couto [5], *Oliver Niebuhr* [6], *João Antônio de Moraes* [5,7]

[1] Haskins labs; [2] Université Paris Saclay, CNRS, LISN; [3] Lund University; [4] Universidade Federal do Paraná; [5] Universidade Federal do Rio de Janeiro; [6] University of Southern Denmark; [7] CNPq

`EricksonDonna2000@gmail.com, albert.rilliard@lisn.fr, malin.svensson_lundmark@ling.lu.se, adelaidehpsilva@gmail.com, leticiarebollocouto@letras.ufrj.br, olni@sdu.dk, jamoraes3@gmail.com`

Abstract

This paper reports on a corpus of Brazilian Portuguese (BP) mandible movements. The data was collected using a recently available technique, the MARRYS helmet, which allows for quick and reliable collection of mandible data of a large number of speakers. Audio and mandible were recorded from more than 90 L1 and L2 BP speakers. The recording process and signal synchronization are presented. A partial set of the corpus, based on 37 L1 speakers producing three sentences mostly composed of /a/ vowels, was segmented at the phone level. The jaw movements are compared to the sentence's prosodic structure. Results indicate that, similar to other languages, Brazilian Portuguese speakers show increased mandible lowering for stressed syllables but also quick mandible closing. Some interesting mandible opening and closing patterns are also reported on the prestress or post-stress positions in relation to the prosodic structure of Brazilian Portuguese.

Index Terms: speech articulation, mandible movement, Brazilian Portuguese, phrasing, stress

1. Introduction

Recent work, e.g., [1], [2], [3], addresses ways the mandible plays an important and necessary role in speech production. Specifically, in order to produce a syllable, the mandible opens and closes, and as such, the mandible can be referred to as the syllable articulator. In addition, the amount of mandible opening changes not only due to segmental effects, i.e., consonants and vowels but also due to changes in prosody, i.e., syllable prominence patterns in the language. Each language has its own utterance prominence patterns and, consequently, language-specific mandible patterns, e.g., [1], [5], [6], [7], [8]. The mandible lowering patterns of an English utterance are such that nuclear stress syllables have the largest amount of mandible opening, followed by phrasal stress and then foot stress, e.g., [1], [6]. However, languages like Mandarin, Japanese, and French, which may be described as edge-strengthening languages [9], show increased mandible lowering at the end of each phrase, with the largest amount of mandible lowering at the end of the utterance [1], [10], [11], [12]. In learning a second language, there is a tendency to transfer first language mandible movement patterns to the second language [1]; focusing on mandible movement patterns helps second language learners improve the prosody of the new language [13]. In addition to prominence articulation, the mandible can be seen to articulate prosodic boundaries, i.e., word, foot, phrase, utterance [14], [15], [16]. Thus, the mandible is not only a syllable articulator; it is also a prosodic articulator.

Given that the mandible plays this important role in speech production, extensive research into mandible movement patterns of many speakers for a large number of languages is called for. However, the collection of mandible data needs to be more easily available. Currently, mandible movement, as well as tongue and lip movement, have been measured using high-tech approaches, such as x-ray microbeam [17] or electromagnetic articulography (EMA) [18]. Both of these techniques require glueing a sensor onto the lower incisor just above the gum line to measure the mandible movement (or to the lips or tongue to measure the other articulators). Generally, the experiments take several hours, including setting up the equipment, gluing the sensors, running the experiment, and then processing the data; as such, it is difficult to get a large number of speakers recorded. A new, less-invasive, and simpler method of recording mandible movement—the MARRYS helmet, a helmet with two bending sensors for bidirectional mandible movement measurement [19], [20], [21], has recently become available [22]. With this device, it is possible to have 15-minute recording sessions, thus making data collection much faster and easier.

The purpose of this paper is (1) a preliminary report on a new method of collecting mandible data for a large number of speakers and (2) an initial examination of how mandible movements of BP speakers reflect the prosodic organization of spoken Brazilian Portuguese. BP is a stress language in which the lexical stress can fall on one of the last three syllables of the word, like in "TÁbata," "saMÁra," and "maraJÁ," respectively referred to as proparoxytone, paroxytone, and oxytone in BP linguistic terminology [23], [24]. According to the prosodic phonology concepts [25], the utterance is hierarchically divided into phonological phrases and phonological words. Within each phonological phrase, one word has more stress than the others. In BP, when there is a clash of lexical stresses in the same phonological phrase, frequently there occurs a deaccenting of the first stressed word, e.g., catTÁR LÁtas > catar LÁtas. Nuclear stress usually falls on the last word in the utterance. In addition, there tends to be an F0 rise at the end of each phrase, except for the final word in declarative utterance, which is characterized by a fall on the last stressed syllable of the last word of the utterance, which is in contrast to a rise over the nuclear stress characterizing the interrogative utterance.

2. Methods

2.1. Dataset

This study reports on 37 L1 speakers of BP. Three sentences were chosen to contain all the same low /a/ vowel, as mandible lowering is affected by vowel height. In addition, the stressed

10.21437/Interspeech.2024-1216

syllable's position within the word varied. Each speaker produced three utterances (A, B, C; see below), with three repetitions each. According to prosodic phonology [25], each sentence's expected phrasing would be with only three phonological phrases, even though they have four words with lexical stresses (or four phonological words). The phonetic transcription places a ' before the stress, also written in bold letters. The boundaries of the phonological phrases are indicated by |, while the || marks the sentence boundary. The positions of stress were the only differences between sentences in terms of prominence. Productions with pauses or deviant pronunciations (repetitions, hesitations, etc.) were discarded. A total of 277 utterances are considered here (74 were discarded).

Sentence A: “Tábata foi chamada pra falar parada.”
['**ta**batɐ | foj ʃɐ'**ma**dɐ | pɾa fa'**lax** pa'**ra**dɐ ||]
Tabata was asked to talk without moving

Sentence B: “Samara foi paga pra catar latas.”
[sa'**ma**ɾɐ | foj '**pa**ga | pɾa ka'**tax** '**la**tɐs ||]

Samara was paid to collect cans.
Sentence C: “O Marajá parava pra matar barata.”
[o maɾa'**ʒa** | pa'**ra**vɐ | pɾa ma'**tax** ba'**ra**tɐ ||]
The maharaja stopped to kill cockroach.

Table 1: *information on the 37 speakers: median age and number by gender and Brazilian macro-regions – and the main states composing them, when relevant.*

Region *(state)*	**Female**	**Male**	**Age**
North	**3**	**0**	**38**
North-east	**-**	**-**	**-**
Center-west	**0**	**1**	**21**
South-east	**6**	**9**	**33**
(São Paulo)	*4*	*5*	*23*
(Rio de Janeiro)	*1*	*2*	*55*
(Minas Gerais)	*1*	*2*	*40*
South (+1)	**9**	**8**	**26**
(Paraná)	*5*	*5*	*24*
(Sta. Catarina)	*2*	*1*	*25*
(Rio Grande do Sul)	*2*	*2*	*35*
Total (+1)	**18**	**18**	**33**

The mandible data was collected in a sound-proof booth at the Federal University of Paraná, Curitiba, Brazil, during the ABRALIN conference (October 2023): this was intended to ease reaching individuals from various places of Brazil (for regional variation of BP, see [26]). The subjects were volunteers, most of them with a linguistic background, who signed consent forms. The project was approved by the UFPR Ethics Committee (CAE registration no. 69008123.6.0000. 0214). Some information about the speaker's age, gender, and geographic origin is given in Table 1 (note one speaker declined to inform their gender and appears only for age in the southern region). The geographic distribution reflects the location of the conference, with a majority of speakers from the southern Brazil Paraná state and an important demographic (as well as economic and proximal) weight from the south-eastern states, in particular, São Paulo; we lament having been unable to record any speakers from the north-eastern states.

2.2. Jaw and audio signals processing

Two signals were recorded: the jaw position through the MARRYS helmet [19] and the speech through an SR-LMX1+ Saramonic omnidirectional Lavalier microphone. The two signals were recorded on separate computers; the sound was recorded with the computer used to present the sentence to be read to the speakers, allowing an easy by-sentence segmentation in a second step. To synchronize the two signals (jaw and audio), a series of three taps on the strap holder of the jaw sensor were made regularly during the recordings. The taps produced a visible (negative) movement on the left jaw sensor and an audible sound. Later, the first and last series of taps were localized on both signals. The average time of the initial and final series of three taps was used to stretch the jaw signal (a time shift between signals is observable after about 10 minutes of recording) and then to align the two signals. Before alignment, each channel of the jaw signal was high-pass filtered with a cutting frequency of 0.05Hz to remove the DC component and eventual slow shift in the signal linked to sensor movements; then, a low-pass filtering was applied (cutoff frequency of 10Hz) to remove noise. The left and right signals were added to get a single measure of jaw aperture. The resulting jaw signal was then standardized for each speaker (centered scaled), considering the values observed on the complete recording (but removing the zones with taps) to estimate its mean and standard deviation.

The audio and jaw signals were then paired in a single WAV file; this signal was then automatically segmented into individual WAV files for each sentence, thanks to synchronization beeps produced by the recording interface. This allowed the application of a forced alignment procedure on the audio channel with the Montreal Forced Aligner [27] and its pre-trained models for Brazilian Portuguese [28], [29]. The alignments were individually checked to detect sentences with pauses, hesitations, repetition, errors, or other incoherence; eventual misalignments were corrected, while sentences with errors were discarded. A phone-level aligned data is thus obtained, with acoustic and jaw movements. Note that the MARRYS helmet was in a development stage; new versions now allow for the smoother obtention of audio and jaw signals. An acoustic analysis of the audio recordings was done to estimate the fundamental frequency (F0, in Hz) and the first five formants, using the standard Praat algorithms [30] with the default parameters for F0, and for formants a 5500Hz frequency ceiling for females and a 5000Hz ceiling for males. F0 values were manually checked and only considered on vowels.

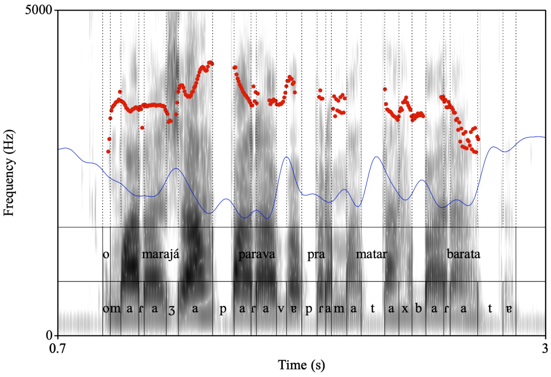

Figure 1: *plot of sentence C: spectrograms, with their segmentation (words and phones), the F0 trace overlaid (red speckles, between 0 and 300Hz), and the jaw movements (blue line, arbitrary unit).*

To compare the productions over speakers, the timing of the sentence had to be normalized across each repetition. This was done by considering the mean duration of each phone of each sentence and for all stimuli. An equivalent of 1 point every 5 ms on this average duration (i.e., if a phone has a mean duration of 87ms, 17 points were considered, starting from the initial boundary of the segment) was used to interpolate the jaw movement observed on each speaker. This led to about 445 points for each sentence and the exact same number of measurements for all repetitions of all speakers of the same sentence.

3. Results

3.1. Example of jaw movement

A Praat display for a typical BP speaker's production of sentence C is presented in Figure 1: the mandible position is overplotted on a spectrographic representation of the sentence, together with a text grid showing the phones' boundaries and the F0 extraction. The typical speaker was chosen as the one whose mandible movement pattern was closest to the average curve, as depicted in Figure 2. The following observations are made: (1) the amount of mandible lowering varies across the utterances, even though the vowels are the same phonologically low vowel; (2) stressed vowels have increased mandible lowering; (3) reduced vowels, including the reduced vowels in the final syllable of the utterance final word, show mandible lowering, but substantially less than the other vowels, and (4) F0 and mandible movement are independent, as one may change independently of the other, that may not change or change in another direction (this does not rule out covariation).

3.2. Mean mandible position over time

Time-normalized data for the 37 L1 speakers of BP are shown in Figure 2: the mandible position of each speaker (individual grey lines) and the mean over all speakers (blue curve) for each of the three sentences (vertical lines indicate the position of each phone boundaries). Rectangular color boxes emphasize the stressed syllables at the end of each phonological phrase. These stressed syllables are lengthened and marked by an opening mandibular movement over the vowel that ends at a low level. In most cases, it is the mandible's lowest position in the phonological phrase (i.e., the maximal aperture). Still, articulatory constraints linked to the phonetic content may affect the relative jaw position, as in the case of [maɾa'**ʒa**] (vs. ['**ta**batɐ] and [sa'**ma**ɾɐ]) where the red rectangular box marks the location of stress in the first phonological phrase of each utterance. These initial phonological phrases show examples of the three types of lexical stress in BP, with increased mandible lowering, respectively on the antepenultimate and penultimate syllables of ['**ta**batɐ] and [sa'**ma**ɾɐ], while for the oxytone [maɾa'**ʒa**], the closing gesture linked to the fricative /ʒ/ preceding the stressed vowel may explain the higher minimal aperture for the stress comparatively to the preceding vowels. A strong jaw opening along the vowel still marks the stress. The mandible thus shows an important opening on the phonological phrase's stressed vowel, followed by a closure on the post-stressed syllable(s) of the word if they exist.

That the stress in [maɾa'**ʒa**] does not show the largest mandible aperture in the word is interesting and aligns with the account that the last two syllables, [ɾa.ʒa] may have equivalent weight in terms of acoustic characteristics: the stress is perceived on the last syllable before a reduction or at the end of the word. According to [31], [32], in BP, stress is marked by the duration of the stressed vowel but also by the fact that pre-stressed vowels may be equivalent to the stressed vowel, which is the last of equals before the reduced post-stressed vowels: post-stress reduction is essential for the identification of stress. In the case of oxytonic words lacking post-stress material, this absence of reduction leads to identifying the final syllable as the stressed one.

The green rectangular boxes in Fig. 2 show the stress at the end of the second phonological phrase, occurring on the words [ʃɐ'**ma**dɐ], ['**pa**ɡa], and [pa'**ɾa**vɐ]. Again, the mandible opens maximally during the stress and closes on the post-stressed syllable of the words. Notice that the amount of mandible lowering on the stressed vowel in this second phonological phrase is larger than on the stressed vowel in the previous phonological phrase.

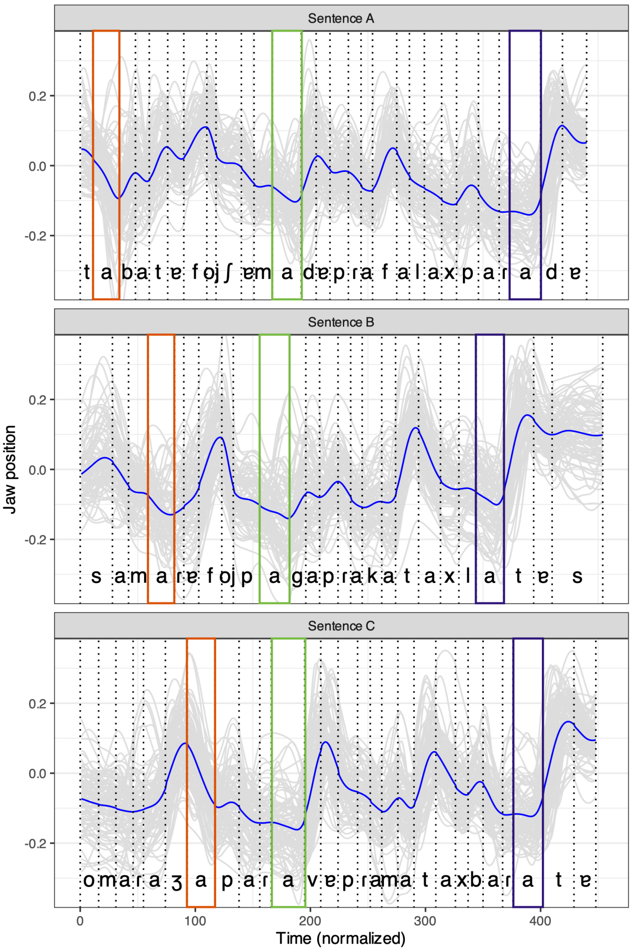

Figure 2: *normalized mandible movements for each of the three utterances: grey lines indicate each stimulus, while the blue curve shows their overall mean.*

The blue rectangular box shows the stress on the third phonological phrase, occurring on the words [pa'**ɾa**dɐ], ['**la**tɐs], and [ba'**ɾa**tɐ]. Again, we see the mandible opens maximally during the stressed vowel and closes on the post-stressed syllable of the word. Notice also that the amount of mandible lowering on the stressed vowel in this utterance-final phrase is the largest in the phrase and sometimes the largest in the utterance (utterance A). This is the nuclear stress, marked by vowel lengthening, F0 fall, and a strong reduction of the post-stressed syllable(s), clearly marked by the strongest and fastest mandibular closing movement over the utterance.

3.3. Intonation patterns

The averaged normalized F0 contours for each utterance are presented in Figure 3. Again, colored rectangles mark the stressed vowels in each phonological phrase, with a fourth box, cyan, marking the stress of the third phonological word. The mandibular movements described above are coherent with intonational patterns observed over the sentences and described in various analyses of BP prosody [31], [33]. For the first three phonological words, it is possible to observe a lower F0 on the stress vowels, with a rise on their final parts and along the post-stressed syllables when they exist. On the nuclear stress (fourth phonological word), the stress is marked for declarative sentences by a high prestress F0 followed by a fall over the stressed vowel (which is lengthened) and a final fall.

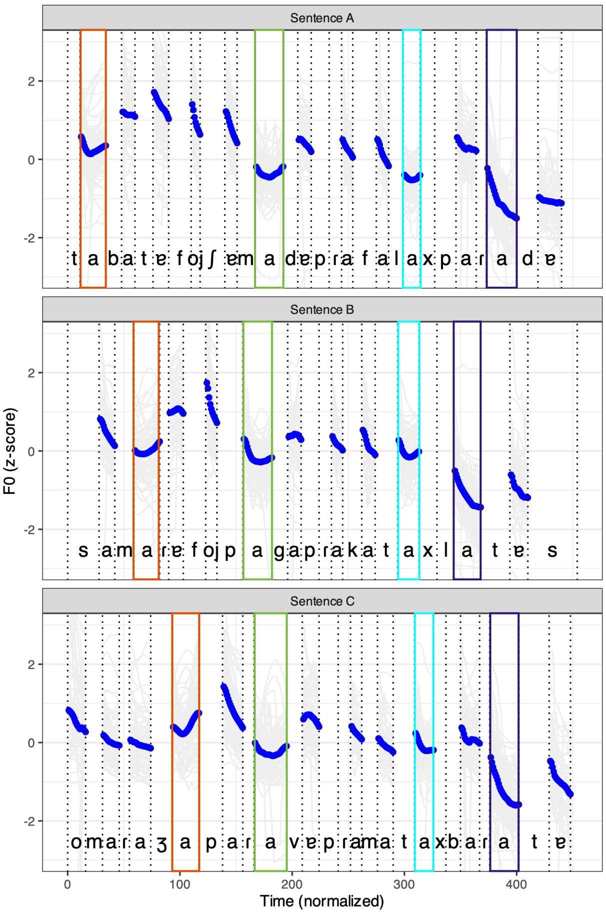

Figure 3: *F0 contours (converted in semitones and centered scaled by speaker) for each of the three sentences; grey curves show individual stimuli, and blue dots indicate the mean curve.*

4. Discussion and conclusions

The results suggest the following pattern of mandible movement for BP speakers: generally, for each stressed syllable, the mandible is lowering/lowered and then closes quickly. There is at least one word with stress within each phonological phrase, and this word tends to occur toward the end of the phonological phrase. In this way, BP is similar to edge-strengthening languages like French, Japanese, and Mandarin; however, since BP has lexical stress marked by F0 movement, it is classified as a "head-prominence" language, similar to some extent to English [9]. In general, compared to American English (AE), for BP, we see less mandible-lowering for each syllable but more of a general pattern of gradual mandible lowering or closing, depending on the position of the stress in the word, for each phonological phrase.

Some interesting differences are observed between BP and AE mandible movement. (1) For AE, the larger the syllable prominence, the greater the mandible lowering; for BP, mandible lowering increases with stressed syllables. However, mandible closing seems to occur more quickly and strongly than in AE. (2) For BP, but not so much for AE (e.g., the descriptions made by [1]), the position of the stress in the word affects the pattern of mandible movement: for the proparoxytone (e.g., TÁbata), there is marked opening for the stressed syllable, followed by a steep closing, and a general upward trend of mandible vowel opening for subsequent syllables in the word until the end of the phonological phrase; for the paroxytone (e.g., saMÁra), the mandible opens progressively for each vowel until it reaches the stress, whereafter there is a steep closing movement toward the word-final/phrase-final reduced syllable; for the oxytone (e.g., maraJÁ), there is a different pattern such that the final stressed syllable does not receive much more mandible lowering, than the preceding syllable, as discussed in Section 3.2., and (3) for AE, the nuclear stress has the largest amount of mandible lowering in the utterance, while for BP, the nuclear stress is marked by increased lowering on the stressed syllable and by a steep, strong jaw closing. Along these lines, we note a difference between the mandible movement for an utterance non-terminal phrase, e.g., the end of the second phonological phrase, vs. the utterance's final nuclear stress: the former is produced with a distinct mandible opening, while the latter is produced with a distinct mandible closing. More work is needed to corroborate these observations.

Given that speakers of different languages employ different mandible movement patterns to implement stress and phrasing, it follows that language teachers could help learners produce better first language prosody by (1) showing learners their L1 mandible movements and (2) training learners then to produce the new L2 movements, similar to "jaw dancing" techniques used to teach Japanese how to improve their AE prosody [13]. Brazil has received a growing and significant number of immigrants and refugees in recent decades; these results can be applied to the teaching of BP as a foreign language as well as BP as a host language, thus helping to develop an awareness of the interconnection between mandible articulation and prosodic or intonation structure in order to improve oral fluency of these speakers.

Future work will continue analyzing the large corpus of mandible data collected for 90 L1 and L2 speakers. This will include a more in-depth exploration of acoustic cues of stress and phrasing, along with the mandible movements for both L1 and L2 BP speakers. Also, the alignment of the acoustic segment with the mandible position, velocity, and acceleration points will be examined. Mandible movement in spoken languages has heretofore been relatively unexplored, mostly due to challenges in collecting and analyzing big data corpora. With an easier and less expensive method of collecting mandible data, perhaps we can have a prosodic/mandible database of languages around the world.

5. Acknowledgments

The authors wish to thank the speakers for their enthusiastic participation. The authors are also deeply indebted to Prof. Pollianna Milan, responsible for the scientific project at UFPR, to Isabela Pratte for making recordings possible and for her constant help, and to the Literature and Linguistics Department of the Federal University of Paraná for receiving us.

6. References

[1] D. Erickson and O. Niebuhr, "Articulation of prosody and rhythm: Some possible applications to language teaching," in *Proceedings of the 13th International Conference of Nordic Prosody*, Sciendo, 2023, pp. 1–45. doi: 10.2478/9788366675728-001.

[2] M. Svensson Lundmark, "Rapid movements at segment boundaries," *J. Acoust. Soc. Am.*, vol. 153, no. 3, pp. 1452–1467, Mar. 2023, doi: 10.1121/10.0017362.

[3] M. Svensson Lundmark and D. Erickson, "Segmental and Syllabic Articulations: A Descriptive Approach," *J. Speech Lang. Hear. Res.*, pp. 1–28, Mar. 2024, doi: 10.1044/2024_JSLHR-23-00092.

[4] M. Svensson Lundmark and J. Frid, "Segmental articulations across prosodic levels," in *Proceedings of the 13th International Conference of Nordic Prosody*, O. Niebuhr, Ed., Sciendo, 2023, pp. 255–261. doi: 10.2478/9788366675728-023.

[5] D. Erickson and S. Kawahara, "Articulatory correlates of metrical structure: Studying jaw displacement patterns," *Linguist. Vanguard*, vol. 2, no. 1, Dec. 2016, doi: 10.1515/lingvan-2015-0025.

[6] D. Erickson, A. Suemitsu, Y. Shibuya, and M. Tiede, "Metrical Structure and Production of English Rhythm," *Phonetica*, vol. 69, no. 3, pp. 180–190, Dec. 2012, doi: 10.1159/000342417.

[7] T. Huang and D. Erickson, "Articulation of English 'prominence'by L1 (English) and L2 (French) speakers," in *Proceedings of the International Congress of Phonetic Sciences*, Melbourne, Australia, 2019, pp. 2480–2484.

[8] D. Erickson, S. Kawahara, Y. Shibuya, A. Suemitsu, and M. Tiede, "Comparison of Jaw Displacement Patterns of Japanese and American Speakers of English: A Preliminary Report." The Phonetic Society of Japan, 2014. doi: 10.24467/onseikenkyu.18.2_88.

[9] S.-A. Jun, Ed., "Prosodic typology: by prominence type, word prosody, and macro-rhythm," in *Prosodic Typology II*, 1st ed., Oxford: Oxford University Press, 2014, pp. 520–539. doi: 10.1093/acprof:oso/9780199567300.003.0017.

[10] D. Erickson, R. Iwata, and A. Suemitsu, "Jaw displacement and phrasal stress in Mandarin Chinese," in *Tonal Aspects of Languages 2016*, Buffalo, New York, 2016, pp. 65–69.

[11] D. Erickson, "Articulatory Prosody: A comparison of Mandarin Chinese and Japanese," in *1st International Conference on Tone and Intonation (TAI)*, ISCA, Dec. 2021, pp. 36–40. doi: 10.21437/TAI.2021-8.

[12] S. Kawahara, D. Erickson, J. Moore, Y. Shibuya, and A. Suemitsu, "Jaw Displacement and Metrical Structure in Japanese: The Effect of Pitch Accent, Foot Structure, and Phrasal Stress," *J. Phon. Soc. Jpn.*, vol. 18, no. 2, pp. 77–87, 2014, doi: 10.24467/onseikenkyu.18.2_77.

[13] I. Wilson, D. Erickson, T. Vance, and J. Moore, "Jaw dancing American style: A way to teach English rhythm," in *Speech Prosody 2020*, ISCA, May 2020, pp. 556–560. doi: 10.21437/SpeechProsody.2020-114.

[14] O. Fujimura, "The C/D Model and Prosodic Control of Articulatory Behavior," *Phonetica*, vol. 57, no. 2–4, pp. 128–138, Dec. 2000, doi: 10.1159/000028467.

[15] D. Erickson *et al.*, "Calculating articulatory syllable duration and prosodic boundaries," in *10th International Seminar on Speech Production, ISSP 2014*, Cologne, Germany, 2014, pp. 102–105.

[16] D. Erickson *et al.*, "Bridging articulation and perception: The C/D model and contrastive emphasis.," in *Proceedings of the International Congresses of Phonetic Sciences*, Glasgow, UK, 2015.

[17] J. R. Westbury, G. Turner, and J. Dembowski, "X-ray microbeam speech production database user's handbook," *Univ. Wis.*, 1994, [Online]. Available: https://ubeam.engr.wisc.edu/pdf/ubdbman.pdf

[18] T. Kaburagi and M. Honda, "Calibration methods of voltage-to-distance function for an electromagnetic articulometer (EMA) system," *J. Acoust. Soc. Am.*, vol. 101, no. 4, pp. 2391–2394, Apr. 1997, doi: 10.1121/1.418255.

[19] M. Svensson Lundmark, D. Erickson, O. Niebuhr, M. Tiede, and W.-R. Chen, "A new articulatory tool: Comparison of EMA and MARRYS," in *Phonetics and Phonology in Europe 2023: Speech variation in the wild*, Radboud University, Nijmegen, Netherlands, Jun. 2023, pp. 33–34. [Online]. Available: https://pape-conference.org/

[20] D. Erickson, O. Niebuhr, W. Gu, T. Huang, and P. Geng, "The MARRYS cap: A new method for analyzing and teaching the importance of jaw movements in speech production," in *Proceedings of the 12th International Seminar on Speech Production (ISSP 2020)*, M. Tiede, D. H. Whalen, and V. Gracco, Eds., Haskins Press, 2021, pp. 48–51.

[21] V. F. Gudmundsson, K. M. Gönczi, M. Svensson Lundmark, D. Erickson, and O. Niebuhr, "The MARRYS helmet: A new device for researching and training 'jaw dancing'.," in *Proceedings of the 25th Interspeech Conference*, Kos, Greece: ISCA, 2024, pp. 1–5.

[22] All Good Speakers, "Flyer Marrys-Cap version 2024," Flipsnack. Accessed: Jun. 05, 2024. [Online]. Available: https://www.flipsnack.com/89B9ADCC5A8/flyer-marrys-cap-version-2024/full-view.html

[23] J. A. de Moraes, "Corrélats acoustiques de l'accent de mot en portugais brésilien," in *Proc. XIth International Congress of Phonetic Sciences*, T. Gamkrelidze, Ed., Tallinn,, Estonia, USSR: Academy of Sciences of the Estoniam S.S.R, 1987, pp. 313–316.

[24] J. A. de Moraes, "Acentuação Lexical e Acentuação Frasal em Português. Um Estudo Acústico-Perceptivo julho de 1995, pp. 39-57.," *Estud. Lingüíst. E Literários*, vol. 17, pp. 39–57, 1995.

[25] M. Nespor and I. Vogel, *Prosodic phonology*. in Studies in generative grammar, no. 28. Dordrecht, Holland ; Riverton, N.J., U.S.A: Foris, 1986.

[26] C. D. S. Cunha, "Os estudos prosódicos no atlas linguístico do Brasil," *Rev. Diadorim*, vol. 20, pp. 291–309, Dec. 2018, doi: 10.35520/diadorim.2018.v20n0a23278.

[27] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi," in *Interspeech 2017*, ISCA, Aug. 2017, pp. 498–502. doi: 10.21437/Interspeech.2017-1386.

[28] M. McAuliffe and M. Sonderegger, "Portuguese (Brazil) MFA G2P model v2.0.0a," https://mfa-models.readthedocs.io/G2P model/Portuguese/Portuguese (Brazil) MFA G2P model v2_0_0a.html, Jun. 2022.

[29] M. McAuliffe and M. Sonderegger, "Portuguese MFA acoustic model v2.0.0a," https://mfa-models.readthedocs.io/acoustic/Portuguese/Portuguese MFA acoustic model v2_0_0a.html, May 2022.

[30] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer program]. Version 6.3.20." Oct. 24, 2023. [Online]. Available: http://www.praat.org/

[31] J. A. de Moraes, "Intonation in Brazilian Portuguese," in *Intonation systems: a survey of twenty languages*, D. Hirst and A. Di Cristo, Eds., Cambridge, U.K. ; New York, NY: Cambridge University Press, 1998, pp. 179–194.

[32] J. A. de Moraes, "The Pitch Accents in Brazilian Portuguese: analysis by synthesis," in *4th International Conference on Speech Prosody*, Campinas, Brazil, 2008, pp. 389–397.

[33] S. Frota and J. A. de Moraes, "Intonation in European and Brazilian Portuguese," in *The Handbook of Portuguese Linguistics*, W. L. Wetzels and S. Menuzzi, Eds., New York: John Wiley & Sons, 2016, pp. 141–166.