Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Stylised sustained prosody in three Australian languages

Kathleen Jepson[1], *Rasmus Puggaard-Rode*[1], *John Mansfield*[2]

[1]Institute of Phonetics and Speech Processing, Ludwig Maximilian University of Munich
[2] Department of Comparative Language Science, University of Zurich
{k.jepson,r.puggaard}@phonetik.uni-muenchen.de, john.mansfield@uzh.ch

Abstract

A similar, striking prosodic pattern is reported to occur in languages around Australia. It is characterised by a stretch of level, high pitch, and lengthening of the IP-final vowel. This pattern appears to have a similar meaning in each language, expressing the extension of an event in space or time. However, there are some differences in the form described. In this paper, we present a study of the acoustics of this pattern in three unrelated Australian Indigenous languages, and propose a method for automatically identifying examples within an audio file. Hence, the purpose of this paper is twofold: 1) to provide a cross-linguistic description of this prosodic pattern with the aim of acoustically describing cross-linguistic variation, and 2) to provide a proof-of-concept for a method to automatically identify this pattern which could allow other language data to be incorporated into the typological description in the future.

Index Terms: intonation, typology, Australian Indigenous languages, iconicity, automated speech processing

1. Introduction

A striking prosodic pattern is described for many Australian languages: a stretch of level, high pitch over an intonational phrase (IP) accompanied by extreme lengthening of the IP-final vowel [1-10]. It has been given various names including "narrative high monotone", "sustained level pitch", "Linear Lengthening Intonation", and variations on "stylised sustained high contour/plateau/tone/intonation". Examples of this pattern are presented in Fig. 1a–c. We use "stylised sustained prosody" (SSP) to refer to this pattern, a term we discuss further in §4. Across languages, this pattern is proposed to have the same function: to extend an event in space or time. For example, to express that people travelled over a long distance, or continuity in doing a task for a long time [9, 11]. It is most often found in narratives, and is said to have an iconic use, with the extent of duration of the vowel related to the magnitude of the distance or time of the event, dramatising the ongoing nature [2, 12].

This pattern is utterly distinct from other prosodically conditioned lengthening phenomena such as phrase final lengthening (e.g., [13]). In [13]'s cross-linguistic survey, finally lengthened vowels in languages without a length contrast were 1.15 to 1.79 times longer than non-final vowels, corresponding to finally lengthened vowels of ~100-150 ms and non-final vowels of ~70-100 ms. Lengthened vowels in the SSP pattern, however, are reported to be up to several seconds in duration (e.g., [3, 14]). [15] found for Jaminjung that lengthened vowels in her dataset were 641.97 ms on average, with the longest vowel being ~3700 ms. However, the scope of durational modification more generally across languages is not known.

SSP affects IP-final vowels, which are most often not stressed, and the languages for which the pattern is described have free word order, meaning that the word classes and morphemes that occur in IP-final position vary within and between languages. This is distinct from similar iconic lengthening patterns of vowels in languages such as English or Italian, where the stressed vowel of adjectives show this type of lengthening, providing an intensification of the adjective's meaning [16]. For example, "the tree was enooooormous", "they live faaaaar out of town", or "the kitten was teeeeeny". In some Australian languages, words other than the one with the relevant semantics can show this lengthening, with the meaning applying over the entire event, irrespective of whether or not the verb occurs within the IP with SSP. For example, a literal translation of Fig. 1a would be "collected the twooooo, many". Further, this feature while optional, is used somewhat systematically, and is a marker of proficient story-telling [11].

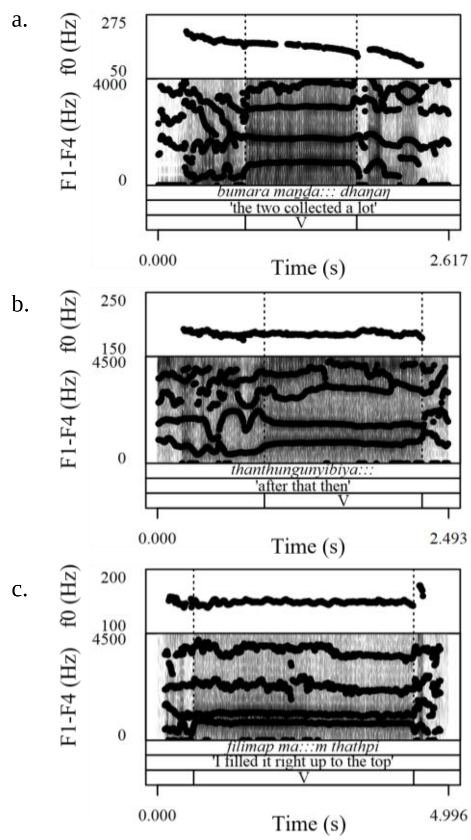

Figure 1: *Examples of the stylised sustained prosody pattern; f0, F1-F4, & spectrogram in a. Djambarrpuyŋu, b. Jaminjung-Ngaliwurru, c. Murrinhpatha. Tier 1: orthographic transcription (":::" indicates lengthened vowel), Tier 2: free translation, Tier 3: annotation of the lengthened vowel. Created using* [17].

10.21437/SpeechProsody.2024-198

While the pitch of SSP is often described to be high and level over the IP and lengthened vowel, there do appear to be differences between languages. In Iwaidja, for example, the contour can be either a high plateau or a rise [7]. While in Djambarrpuyŋu, following the level high pitch of the IP, the pitch over the lengthened vowel can remain high or fall steadily, reflecting continuation or finality, respectively [6]. In Kuninjku, on the other hand, the high, level pitch may spread up to the IP boundary tone; if this is L%, signalling finality, then there is a sharp fall over a short portion of the lengthened vowel, equivalent to the duration of a regular vowel, and any following sonorant in the rime [12]. Reduplicated verb stems in utterances with SSP are reported to occur in Nunggubuyu [4], Wik-Mungkan [8], and Anindilyakwa [1]; in the first language the IP-final vowel is lengthened, in the second the lengthened vowel is proposed to be an additional clitic =*a::*, while in the last a "post-lengthening" clitic =*wa* occurs following the lengthened vowel, often accompanied by falling/low pitch.

This study aims to provide a cross-linguistic, acoustically informed description of the SSP pattern including identifying some aspects of the variation in three unrelated Australian Indigenous languages: Djambarrpuyŋu (Dj.), Jaminjung-Ngaliwurru (JN), and Murrinhpatha (MP). Specifically, in light of the variability reported across languages, we aim to describe this pattern considering: i) the difference in f0 between the IP with SSP and the mean f0 for the speaker, to assess if the contour is higher than most regular speech; ii) the f0 trajectory over the lengthened vowel which is often the focus in impressionistic descriptions; and iii) the duration of these vowels. A further aim is to provide a proof-of-concept for a method to automatically identify this pattern within long audio files to be able to incorporate data from other languages and expand this preliminary typological study.

1.1. Automatically finding instances of SSP

Instances of the SSP pattern are relatively rare, and the relevant data are often found in sizable collections of long sound files (i.e., corpora of narratives collected over the course of years of fieldwork). This poses a problem for a potential larger typological study of the pattern: manually finding instances of SSP is time-intensive for those reasons, and because SSP may not be indicated in accompanying transcriptions. A larger study of SSP could be feasible with a method of annotation that automatically identifies candidate stretches of speech.

There are several reasons why general-purpose tools like forced alignment [18] cannot help in this task. For one, these tools generally perform best for major languages where a lot of training data is available (although note that MAUS [19] provides a general purpose forced alignment option for Australian languages). Furthermore, the extreme vowel duration in instances of SSP would likely result in errors within forced alignment since such tools generally perform best when the input reflects the training data [20].

Given 1) manually finding instances of SSP in large corpora would be prohibitively time-demanding, and 2) existing tools for the automatic labelling of speech are ill-equipped to deal with such unusual speech segments, we propose a specific-purpose algorithm for finding instances of SSP in long audio files. One or more consistent cues would be required to automatically identify candidate instances of SSP. Inspection of the data (see Fig. 1) suggested highly stable F1 of the lengthened vowel could be a suitable cue. The composition of the algorithm is discussed in §2.2.2.

1.2. Languages

The three languages described in this paper are spoken in relatively remote areas of northern Australia (see Fig. 2), and belong to three different language families: Dj. is Pama-Nyungan, JN is Mirndi, and MP is a Southern Daly language. They share similar, small vowel inventories — Dj. /i iː a aː u uː/, JN /i e a u/, MP /i e a u/ — but have distinct intonational systems and patterns of word-level prosody [6, 21, 22]. Dj. and MP are acquired by children, JN is spoken by a very small number of elderly people.

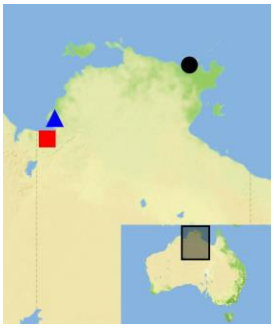

Figure 2: *Map showing where Djambarrpuyŋu (black circle), Jaminjung-Ngaliwurru (red square), and Murrinhpatha (blue triangle) are spoken in northern Australia. Created using* [23].

2. Methods

2.1. Materials

The data used in this study are narratives, and include both contemporary and archived recordings. The narratives were told by women and men, and covered mythological stories and recounts of personal events. In total, the data analysed is 63:09 minutes of audio, comprised of three Dj. narratives totalling 22:47 minutes, four JN narratives totalling 29:09 minutes, and one MP narrative totalling 11:13 minutes. Dj. audio data were recorded in Milingimbi in 2016 and 2017 using a Zoom H6 digital recorder and Countryman H6 headset microphone with a hypercardioid pattern directional capsule covered with a windshield, made at 24 bit bit-depth and a 48 kHz sample rate [24]. JN audio data were collected in Timber Creek and Kununurra in the 2000s and early 2010s, and were made using a range of recording devices and sample rates (44.1 – 96 kHz) [25]. MP audio data were collected in Wadeye in the early 1980s on tape, and were digitised by and archived in PARADISEC [26]. The narratives were transcribed and translated with assistance of Dj.-, JN-, and MP-speaking language consultants. Files were downsampled to 44.1 kHz and stereo files were converted to mono using Praat [27].

2.2. Data preparation and measures

To begin to develop a tool for automatically identifying SSP, instances were manually annotated. In this way it was possible to assess the accuracy of the tool and make adjustments to increase performance.

2.2.1. Manual annotation

Instances of the SSP pattern were identified in Praat, based on impressionistic features of pitch and vowel duration, in consultation with the existing transcriptions. SSP is distinct when it occurs, and could usually be straightforwardly determined based on those features. The vowel was annotated as well as a unit corresponding to an intonational phrase. Other annotation tiers included the orthographic transcription and free translation.

2.2.2. Automatic annotation

A function was written in R [28] to automatically identify the lengthened vowel in instances of the SSP pattern in the audio data introduced in §2.1 using the following procedure.

Formant and root-mean-squared (RMS) amplitude tracks for the entire narrative sequences are generated using the *wrassp* package [29] in R with a sampling rate of 200 Hz. By default, the algorithm loops over the F1 and F2 tracks in 500 ms windows taken at each 50 ms intervals. Since F1 is often stable in instances of SSP, for each window, it is calculated how many F1 measures fall below an automatically calculated dynamic range threshold corresponding to the 60% sample quantile of the RMS amplitude track for each sound file, and what the average difference is in Hz between adjacent F1 samples within the window. Windows with at most three samples below the dynamic range threshold and where the average difference in adjacent F1 values is below 10 Hz are considered candidates; consecutive windows that match these criteria are considered part of the same SSP instance. In order to filter out spurious cases, for each window, the average difference in F2 values that are 100 ms apart is also calculated, and windows are rejected if that difference is above 100 Hz. The F2 filter is optional. An alternative filter based on adjacent f0 measurements (also calculated using *wrassp*) is also implemented; this filter improves performance when formant tracking is poor. The defaults' performance matches human annotators fairly well, but the function is flexible; users can define their own dynamic range thresholds, window lengths, window shifts, and lags and thresholds for f0, F1, and F2. The outcomes of varying these parameters for each language and recording are discussed in §3.1.

By default, the function outputs a Praat TextGrid with the predicted locations of lengthened vowels (i.e., instances of SSP), but the locations can also be stored as a data object in R, added as a new tier to an existing TextGrid, or added as a separate annotation layer to an EMU-SDMS database. The package *tuneR* [30] is used to process audio data, and *rPraat* [31] is used to write TextGrids automatically using R.

2.2.3. Database construction and extraction

An EMU-SDMS database [32, 33] was created that included both the manual and automatically annotated tiers (default and recording-specific). Using the manual annotations, vowel duration was extracted, f0 trajectories for the lengthened vowel was calculated using *wrassp*, the mean pitch over the SSP IP (including the lengthened vowel) was calculated, as well as the average pitch for each speaker over the entire file.

3. Results

In total, 83 instances of SSP were manually annotated in the data, including 37 from Dj., 21 from JN, and 25 from MP. The frequency of SSP therefore varies considerably between the language samples, occurring 1.6 times per minute of audio in Dj., 0.7 times per minute in JN, and 2.3 times per minute in MP.

3.1. Automatic identification

The algorithm for identifying instances of SSP annotated a total of 125 candidate stretches of speech using the default settings. Because instances of SSP are relatively rare, we prioritise finding *all* clear instances of the pattern over finding *only* clear instances of the pattern; this means that the annotation method unavoidably sometimes falsely identifies cases of irrelevant speech or silence as an instance of SSP (see Table 1). We aim to minimise this in future iterations of this method; however, output must still be manually checked. Since there are relatively few instances overall, this should still save time relative to finding instances manually, especially in unfamiliar audio files.

We found that setting the parameters specifically for each audio file greatly improved performance. Narrative data is collected by many linguists, and may not always be recorded under optimal conditions. Setting a dynamic range specifically for each file worked better than the current default, which automatically calculates a dynamic range. Furthermore, it was helpful to increase the difference threshold in F1 for recordings where the formant tracking did not perform well due to adverse recording conditions, and in these cases, it was also helpful to filter instances based on f0 rather than F2. Table 1 shows the proportion of correctly identified cases, falsely identified cases, and missed cases for each recording using the default parameters and recording-specific parameters.

Table 1: *The proportion of correctly identified cases, falsely identified cases, and missed cases for each recording when using default parameters of the automatic method / recording-specific parameters, as well as the number of manually annotated instances.*

	correct	**incorrect**	**missed**	**manual**
Dj. (1)	6 / 7	5 / 7	10 / 9	16
Dj. (2)	5 / 6	2 / 4	2 / 1	7
Dj. (3)	8 / 9	36 / 0	6 / 5	14
JN (1)	0 / 2	0 / 2	2 / 0	2
JN (2)	0 / 1	21 / 25	9 / 8	9
JN (3)	3 / 3	0 / 0	0 / 0	3
JN (4)	4 / 5	2 / 2	3 / 2	7
MP	19 / 22	14 / 11	7 / 3	25
Total	46 / 55	80 / 51	39 / 28	83

The algorithm works best when searching only for instances where the lengthened vowel is above 500 ms. Some of the instances we have identified manually are as short as 203 ms (see §3.2.3), and these short instances make up the majority of missed cases in Table 1. We have not found an adjustment to the algorithm that can consistently distinguish short instances of SSP from phonemically long vowels. Finding the shortest instances of SSP likely requires either manual annotation or considerable manual checking of the automatic output.

3.2. Acoustic analysis

3.2.1. SSP f0 relative to the speaker's average

While the contour is described to be high in most discussions of SSP, in these data, the average f0 of the IPs of interest appears to be around the speaker's average (calculated over the entire narrative), or slightly lower. Thus, overall, these contours are usually slightly below the middle of the speaker's range, not in the higher portion. Fig. 3 illustrates the distribution of values calculated by subtracting the average f0 for the speaker from the mean f0 calculated over each target IP; negative values indicate the SSP IP had lower mean f0 than the speaker's average, and vice versa.

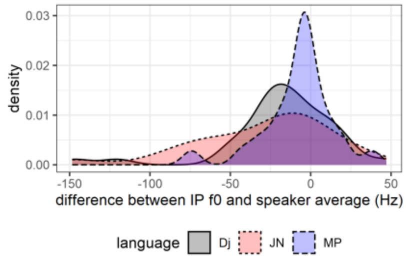

Figure 3. *Density plot showing IP mean f0 – speaker mean f0.*

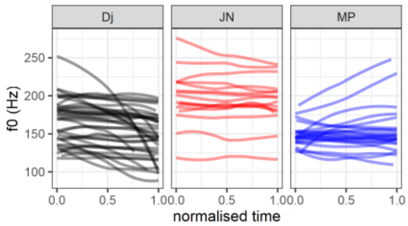

Figure 4: *f0 trajectories of lengthened vowels. Seven trajectories were removed due to tracking errors.*

3.2.2. *Vowel f0*

The f0 trajectories of the lengthened vowels were plotted over normalised time (Fig. 4) to visualise the patterns within and across language samples. It is clear that level f0 over the vowel is most common, in support of cross-linguistic impressionistic descriptions of a plateau in pitch. In Dj., there are additionally a number of falling contours, while in MP variously shaped rises occur.

3.2.3. *Vowel duration*

The durations of the lengthened vowels varied considerably within and between languages (see Table 2 and Fig. 5). At the upper end, vowel durations were within the range described in previous work, with the longest vowel in the dataset being almost 3800 ms, from the MP data. MP lengthened vowels were also longer overall (see Fig. 5). JN had both shortest mean and maximum duration values; the mean duration was similar to the value reported in [15] but the current dataset does not contain vowels of the longer durations previously reported. However, it is anticipated that extremely long vowels can occur in any language, and the upper limit is not language specific.

Table 2: *Duration measurements (ms) of lengthened vowels, across languages.*

	mean dur. (ms)	**SD (ms)**	**max. dur. (ms)**
Dj.	748	422	2123
JN	691	298	1351
MP	1403	749	3774

Fig. 6 shows the duration of phonemically short and long vowels from a controlled Dj. dataset [34] plotted alongside the lengthened vowels from the Dj. narrative dataset to understand how these duration values pattern generally. There is some overlap between the lower range of the lengthened vowels and phonemically long vowels, but overall, the lengthened vowels are considerably longer than all other vowels.

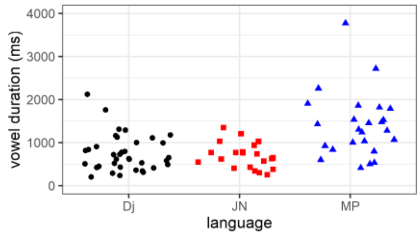

Figure 5: *Duration of lengthened vowels.*

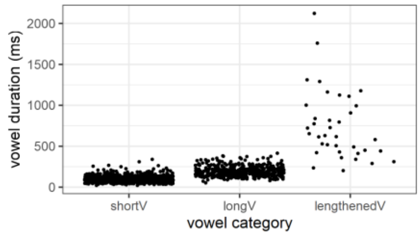

Figure 6: *Duration of phonemically short (n = 991) and long (n = 437) vowels (in word-initial syllables), and IP-final lengthened vowels (n = 37) in Dj.*

4. Discussion and conclusions

We aimed to describe acoustic aspects of a frequently reported, variously named prosodic pattern of extreme vowel lengthening and plateaued pitch in three Australian languages. Overall, there is a strong preference for plateaued f0 over the lengthened vowel, and while the prototypical pattern of a high plateau is observed, there is also clear, and somewhat language-specific diversity of rises, falls, and plateaus at other levels within a speaker's range. Further, the f0 of the target IPs is not always relatively high. Therefore, including "high plateau" in the name obscures the full range of patterns that occur. The term "stylised sustained prosody" proposed in this paper captures both the lengthening of the vowel that is necessary for this prosodic pattern's interpretation as an extended event, as well as the relatively stable patterns we observe in the pitch. Importantly, we did not find complex movements in pitch, which could be amply achieved in the duration of the lengthened vowels.

The lengthened vowels are considerably longer than vowels in regular speech, but there is a degree of ambiguity in identifying this pattern [12]. The proposed algorithm for automatically identifying SSP performs reasonably well, especially when using recording-specific parameters, but the algorithm cannot solve ambiguous cases and does not perform well for instances of SSP that are within the duration range of 'regular' long vowels. The duration values for Dj. show some overlap between SSP lengthened and long vowels. However, vowel length is contrastive only in word-initial syllables in Dj., so there is rarely positional overlap between these vowels. Further, Fig. 6 compares controlled speech with natural speech, so we might expect the duration of phonemically short and long vowels to be somewhat shorter in comparable narrative data.

SSP is an iconic pattern of narrative speech in many Australian languages, in both senses of the word. It is a marker of skillful narrators, and provides additional information about an event. In this paper we presented a brief overview of SSP in three languages, contributing to the description of these under-documented languages. Future work will incorporate narrative data from other Australian languages, using the automated method presented to speed up identification of candidates, with the aim of undertaking a larger typological study.

5. Acknowledgements

We thank Albena Buyanggirr, Jimmy Ginigiri, Margaret Guwangil, Jessie Murarrgirarrgi, Erich Round, Eva Schultze-Berndt, and Raphael Tunmuck. KJ was supported by the Alexander von Humboldt Foundation and the ARC Centre of Excellence for the Dynamics of Language.

6. References

[1] J. Bednall, “Temporal, aspectual and modal expression in Anindilyakwa, the language of the Groote Eylandt archipelago, Australia,” PhD thesis, the Australian National University; Université Paris Diderot (Paris 7), 2019.

[2] J. Bishop and J. Fletcher, “Intonation in six dialects of Bininj Gun-wok,” in *Prosodic typology: The phonology of intonation and phrasing*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2005, pp. 331-361.

[3] J. Fletcher, “Intonation and prosody in Dalabon,” in *Prosodic typology II: The phonology of intonation and phrasing*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2014, pp. 252-272.

[4] J. Heath, *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies, 1984.

[5] S. Hellmuth, F. Kügler, and R. Singer, “Intonational patterns, tonal alignment and focus in Mawng,” presented at *ICPhS2007* satellite workshop *Intonational Phonology: Understudied or Fieldwork Languages*, Aug. 2007.

[6] K. Jepson and J. Fletcher, “The intonation of Djambarrpuyŋu,” in *Prosodic Typology III: The phonology of intonation and phrasing*, S.-A. Jun and S. D. Khan, Eds. Oxford: Oxford University Press, *provisionally accepted*.

[7] R. Mailhammer and P. Caudal, “Linear Lengthening Intonation in English on Croker Island: Identifying substrate origins” *JournaLIPP*, vol. 6, pp. 40-56, Mar. 2019.

[8] B. Sayers, “Interpenetration of stress and pitch in Wik-Munkan grammar and phonology,” in *Papers in Australian Linguistics*, vol. 9, Series A, no. 42, J. Hudson and B. Sayers, Eds. Canberra: Pacific Linguistics, 1976, pp. 31-79.

[9] M. Sharpe, *Alawa phonology and grammar*. (Australian Aboriginal Studies no. 37, Linguistic series no. 15). Canberra: Australian Institute of Aboriginal Studies, 1972.

[10] C. Yallop, *Alyawarra: An Aboriginal language of central Australia* (Australian Aboriginal Studies, Research and Regional Studies no. 10). Canberra: Australian Institute of Aboriginal Studies, 1977.

[11] M. Walsh, “Ten postulates concerning narrative in Aboriginal Australia,” *Narrative Inquiry*, vol. 26, no. 2, pp. 193-216, 2016.

[12] J. Bishop, “Aspects of intonation and prosody in Bininj Gun-wok: An autosegmental-metrical analysis,” PhD thesis, the University of Melbourne, 2002.

[13] L. Paschen, S. Fuchs, and F. Seifart, “Final lengthening and vowel length in 25 languages,” *Journal of Phonetics*, vol. 94, pp. 1-22, 2022.

[14] F. Merlan, *Mangarayi*. Amsterdam: North-Holland Publishing Company, 1982.

[15] C. Simard, “Prosody and function of “iconic lengthening” in Jaminjung,” *SOAS Working Papers in Linguistics*, vol. 16, pp. 65-77, 2013.

[16] J. Guerrini, “Vowel quality and iconic lengthening,” in *Proceedings of Sinn und Bedeutung*, M. Franke, N. Kompa, M. Liu, J. L. Mueller, and J. Schwab, Eds., 2020, vol. 1, no. 24, pp. 242-255.

[17] R. Puggaard-Rode, “praatpicture: 'Praat Picture' Style Plots of Acoustic Data”. R package version 1.0.0, 2024. https://CRAN.R-project.org/package=praatpicture

[18] S. Gonzalez, J. Grama, and C. E. Travis, “Comparing the performance of forced aligners used in sociophonetic research,” *Linguistics Vanguard*, vol. 6, no. 1, art. 20190058, 2020.

[19] T. Kisler, U. D. Reichel, and F. Schiel, “Multilingual processing of speech via web services,” *Computer Speech & Language*, vol. 45, pp. 326-347, 2017.

[20] L. MacKenzie and D. Turton, “Assessing the accuracy of existing forced alignment software on varieties of British English,” *Linguistics Vanguard*, vol. 6, no. s1, art. 20180061, 2020.

[21] C. Simard, “The prosodic contours of Jaminjung, a language of Northern Australia,” PhD thesis, University of Manchester, 2010.

[22] J. Mansfield, *Murrinhpatha words: Morphology and phonology*. Berlin: De Gruyter Mouton, 2019.

[23] Stamen Design, “Terrain map tiles,” under CC BY 4.0. Data by OpenStreetMap, under ODbL, 2023. https://maps.stamen.com/

[24] K. Jepson, “Recordings from Milingimbi Island/Yurrwi, northeast Arnhem Land, Northern Territory, Australia (KJ1)”. Digital collection managed by PARADISEC, 2019.

[25] E. Schultze-Berndt, C. Simard, M. Harvey, D. Hoffmann, and Jaminjung and Ngaliwurru communities. “Annotated audiovisual corpus of Jaminjung and Ngaliwurru, 1993–2017.” Digital collection managed by DOBES Archive, 2017.

[26] C. Street, “Chester Street recordings (CS1)”. Digital collection managed by PARADISEC, 1972.

[27] P. Boersma and D. Weenink, “Praat: Doing phonetics by computer”, version 6.4.01, 2023. https://fon.hum.uva.nl/praat

[28] R Core Team, “R: A language and environment for statistical computing”, version 4.3.2, 2023. https://www.R-project.org/

[29] R. Winkelmann, L. Bombien, M. Scheffers, and M. Jochim, “wrassp: an R wrapper to the ASSP Library”, R package version 1.0.4, 2023. https://CRAN.R-project.org/package=wrassp

[30] U. Ligges, S. Krey, O. Mersmann, and S. Schnackenberg, “tuneR: Analysis of Music and Speech”. R package version 1.4.6, 2023. https://CRAN.R-project.org/package=tuneR

[31] T. Bořil and R. Skarnitzl, “Tools rPraat and mPraat,” in *Proceedings of the International Conference on Text, Speech, and Dialogue*, book series *Lecture notes in computer science*, vol. 9924, pp. 367-374, Sep. 2016.

[32] R. Winkelmann, J. Harrington, and K. Jänsch, “EMU-SDMS: Advanced speech database management and analysis in R,” *Computer Speech & Language*, vol. 45, pp. 392-410, 2017.

[33] M. Jochim, R. Winkelmann, K. Jänsch, S. Cassidy, and J. Harrington, “emuR: Main Package of the EMU Speech Database Management System”. R package version 2.4.2, 2023. https://CRAN.R-project.org/package=emuR

[34] K. Jepson, “Contrastive vowel length and segment duration in production and perception in Djambarrpuyŋu,” *under review*.