Stop Oppositions in Warumungu: A Distributional and Acoustic Analysis

Mitchell Browne[1,2], Michael Proctor[1], Jane Simpson[3], Mark Harvey[4], Robert Mailhammer[5], and Harriet Carpenter[1]

[1]Macquarie University, [2]The University of Western Australia, [3]The Australian National University, [4]The University of Newcastle, [5]Western Sydney University

mitchell.browne@mq.edu.au

Abstract

Warumungu is a language of Central Australia with a reported contrast for stops. There has been no formal distributional or acoustic analysis of this contrast; this study begins to address this gap. Analysis of 439 stops produced by three speakers reveals that the consistent phonetic correlates of the manner contrast in morpheme-medial positions are duration and voicing. The primary distinction in stop manner is duration; voicing distinctions are predictable from duration.

Index Terms: stops, plosives, Warumungu, acoustic phonetics, Australian languages, duration, voicing

1. Introduction

The Warumungu (Pama-Nyungan) consonant inventory departs from the prototypical inventory of Australian languages (see [1], [2, pp. 605–615]) in having two intervocalically contrastive stop series at each place of articulation, illustrated in (1) and (2). The contrast has previously been described [3] as a distinction between short voiced ('Mode 1'), and long voiceless stops ('Mode 2'); however, the phonetic basis of the contrast has not been systematically examined.

	MODE 1	MODE 2
1)	/kantu/ 'inside'	/nant:u/ 'windbreak'
2)	/wacala/ 'subsection name'	/cac:ula/ 'dew, fog'

Warumungu stops have a long history of analysis. Previous linguistic analysis has recognised a phonetic manner opposition in stops. The first analysis was presented in 1953 by Arthur Capell, who observed stronger stop devoicing than in most Australian languages, as well as "a tendency to gemination of plosives", but claimed that this was neither consistent nor phonemic [4]. Subsequent analyses have treated the distinction as phonemic: Ken Hale described a contrast between 'tense' and 'lax' stops [5], and Jeffrey Heath characterized the contrast as one of fortis/lenis [6].

One reason that analyses of this contrast differ is because MODE 2 stops occur in a restricted environment. Hale noted a general lengthening process of post-tonic consonants, which is only distinctive for stops (primary stress in Warumungu is generally word-initial; with secondary stress on non-word-final alternating syllables, with some lexical exceptions). This analysis was continued by Prithindra Chakravarti, who further observed a complementary distribution between long vowels and long consonants in stressed syllables, which he labelled "bimorism or bisyllabism" [7].

In addition to stress-related effects, there are also morphophonological interactions involving stops in Warumungu. Firstly, medial MODE 2 stops in disyllables become short when certain suffixes are added [3], [6]. For example, the MODE 2 stop in /ŋap:a/ is shortened when the dative suffix attaches: /ŋapa-ka/. Secondly, initial stops of these alternation-triggering suffixes themselves may be voiced or voiceless, depending on whether they were suffixed to disyllabic words or polysyllabic words, and on whether the suffix was monosyllabic or polysyllabic [5], [6]. In these environments, the 'short' stops are described as alternating between voiced (for disyllables) and voiceless (elsewhere) [3].

In sum, over the past seventy years, the existence of at least two series of oral stops has been recognised in Warumungu. What distinguishes the two series has been variously described as lax/lenis/short versus tense/fortis/long; however, no systematic quantitative phonetic analysis has been conducted to test these analyses and inform the characterization of this contrast. It is not clear what the primary phonological distinction between stops is; whether they are specified for more than one parameter; and how categorically the two stop series are distinguished.

1.1. Stop oppositions in Australian languages

Stop contrasts in Australian languages are rare [8]. Warlmanpa, a language neighbouring Warumungu, is reported to have a stop contrast based on length [9], however this requires further phonetic analysis. While it is not reported to be contrastive, medial stops in Warlpiri are reported to have high constriction duration in post-tonic position, with durations of approximately 100-140ms [10], [11]. In the Top End (the northern part of the Northern Territory), stop oppositions similar to the Warumungu system have been described in Bininj Gunwok [12], Burarra [13], Jawoyn [14], Murrinh-Patha [13], and Ngalakgan [15].

Cross-linguistically, stop contrasts characterized as voicing oppositions are typically realized through systematic differences in voice onset time (VOT) [16], [17]. In Murrinh-Patha and Ngan'gityemerri – languages of the Daly River region – VOT is the primary correlate of the stop manner opposition [18], [19]. In other languages of the Top End, VOT is not a consistent correlate of stop mode [12], [13], [14], [15], [20], so it does not appear that the stop manner opposition in these languages can be categorized as a voicing opposition. Analyses of these languages have proposed that the basis of the opposition is duration. In Bininj Gunwok for example, the mean length of lenis stops is 76 ms, compared to 161 ms for fortis stops [12, p. 173]. In other Top End languages, duration is not an independent predictor of stop type: Evans & Merlan [14, p. 216] propose that the phonological opposition in Jawoyn stops is determined by comparative quantitative weighting of both duration and voicing. No phonetic analysis of stops has yet been conducted to examine the basis of similar manner contrasts in Warumungu.

1.2. Research questions

The aim of this study is to examine the distributional and acoustic phonetic properties of Warumungu stops, to determine:

1. the relative lexical distribution of Mode 1 and Mode 2 stops inside monomorphemic words
2. whether duration differs between Mode 1 and Mode 2 stops
3. whether voicing differs between Mode 1 and Mode 2 stops
4. how duration and voicing interact

Based on the answers to these questions, we will address the following hypotheses about the Warumungu stop contrast: (H1) duration alone is phonologically specified; (H2) voicing alone is phonologically specified; and (H3) both duration and voicing are independently specified in the phonology.

2. Method

2.1. Distributional analysis

To establish the distribution of stops of different type, we examined an existing word list of Warumungu. Because inflection influences stop mode, we excluded any morphologically complex entries (e.g. verbs, which require an inflection to be well-formed) and words which are orthographically recognised as complex (e.g. complex verb constructions), as well as bound morphemes. After these exclusions, the total number of words considered in our distributional analysis was 804.

2.2. Acoustic analysis

Data used for acoustic analysis is drawn from two sources, both comprising careful speech. The first source is an audio archive containing acoustic recordings of a single adult female speaker (age 40+) producing citation form examples for the Warumungu Picture Dictionary [22], either in isolation or short sentences. The second data source for the phonetic study consists of acoustic recordings made during experimental fieldwork conducted around Tennant Creek in 2023 and 2024. These data were elicited from two adult female (ages 50+ and 60+), both native speakers of Warumungu. Participants were shown a series of 32 images corresponding to Warumungu words containing target segments, checked by speakers as to whether they represented the intended target word.

Participants were instructed to say the corresponding Warumungu word twice for each image. Each participant was shown the series of images in a different (random) order. Each image was displayed until the participant provided a response. Responses were recorded using a Zoom H5 with an LMF-1 Lavalier microphone. Speech audio was recorded at a sampling rate of 44.1 kHz and saved as 16-bit uncompressed WAV files. Participants were remunerated for their time.

970 Warumungu words containing stops were identified in the data sources, of which 461 met the criteria for acoustic analysis: only medial stops in tautomorphemic words of at least two syllables were analyzed.

Matlab [23] tools were developed for data inspection, audition, and analysis, to facilitate systematic processing and acoustic segmentation of the dataset. Medial stop intervals were located in each word using a semi-automatic algorithm. The interval of occlusion was selected on a waveform/spectrogram plot by a phonetically trained analyst (Author 1) and a supervised student (Author 6). Occlusion boundaries were estimated automatically from amplitude and energy thresholds calculated over overlapping 20 ms windows, and manually corrected when necessary. Stop duration was calculated over the total interval of occlusion including any carryover voicing and negative voice onset intervals (Figure 1). Voicing was estimated from harmonic ratios (HR), which calculate the distribution of energy across frequency bands [24], [25]. Voiced speech is characterized by greater concentration of energy at multiples of the fundamental frequency (HR → 1), while voiceless/devoiced speech is produced with more even spectral distribution of energy, resulting in lower harmonic ratios (HR → 0). Harmonic ratios were calculated over 25 ms overlapping windows throughout each word, and the degree of voicing for each stop was calculated as the mean HR over the occlusion interval. Using mean HR allows for consistent analysis between stops, particularly where clear landmarks are not identifiable (e.g. approximant realisations).

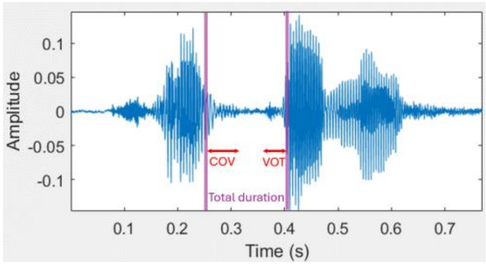

Figure 1: *Total duration example of* /cap:ina/ *'bearded dragon'.*

Each token in the acoustic analysis was coded for place of articulation; mode (based on the current practical orthography used in Warumungu dictionaries in which stops are written with one letter (MODE 1) or two letters (MODE 2) , e.g. *kantu* 'inside', *nanttu* 'windbreak'); speaker; word; total syllable count of word; preceding morae (N); mean HR; and duration. Tokens with excessive background noise were excluded. Final token counts are shown in Table 1.

Table 1: *Number of stops included in acoustic analysis.*

PLACE OF ARTICULATION	MODE 1	MODE 2
Bilabial	77	33
Alveolar	32	25
Retroflex	44	26
Palatal	63	36
Velar	79	46
TOTAL	295	166

2.3. Statistical analysis

Outliers in which mean HR deviated more than two standard deviations from the mean for each stop mode were excluded from statistical analysis. 22 of the original 461 tokens were excluded by this criterion, leaving 439 stops for analysis.

We examined stop duration, voicing, and the interaction between duration and voicing using linear mixed-effects models in R [26], [27], [28]. Stop duration (3) was modelled as a function of stop mode with interactions for place of articulation, and random effects of word (intercepts & slopes) and speaker (intercepts only). Stop voicing (4) was modelled as a function of stop mode (baseline MODE 1) with interactions for place of articulation (baseline ALVEOLAR), and random effects of word (intercepts only) and speaker (intercepts & slopes). The most descriptive convergent model was chosen for each variable.

3) duration ~ Mode*PoA + (Mode|Word) + (1|Speaker)

4) voicing ~ Mode*PoA + (1|Word) + (Mode|Speaker)

To examine the relationship between duration and voicing, we modelled voicing (mean HR) as a dependent variable and duration as the independent variable. We also examined the overlap in the distribution of tokens (with respect to duration and voicing), calculated with the R package OVERLAPPING [29]. Plots were generated with the PLOTLY package [30] in R.

3. Results

3.1. Distributional analysis

MODE 1 stops are relatively unrestricted, able to occur in the onset of any syllable. MODE 2 stops are primarily found in disyllables following primary stress.

Considering only words which have a stop in post-tonic position (n=407), 246 (60%) are documented as MODE 1, and 161 (40%) as MODE 2 (Figure 2). When further restricting this to disyllabic words (n=145) the ratio is reversed: 50 (34%) disyllabic words have a MODE 1 stop following primary stress; and 95 (66%) words have a MODE 2 stop in this environment. At each place of articulation, both stop types occur intervocalically, and as the second member of a cluster, i.e. V.**C**V and VC.**C**V respectively. MODE 1 stops can rarely occur as the first member of a cluster, whereas MODE 2 stops are not evident in this environment.

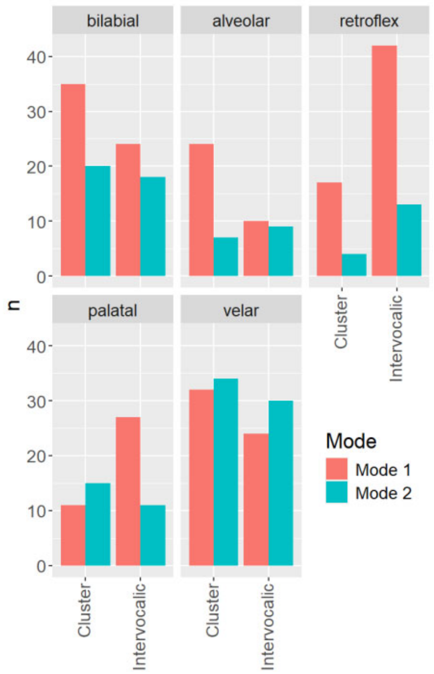

Figure 2: *Distribution of stops found in second syllable onsets, grouped by stop mode and environment.*

The distribution and location of MODE 2 stops in words of different length are summarized in Table 2. 89% of MODE 2 stops are found following a primary stress (i.e. 1 preceding syllable), the majority of which are found in disyllable words; and 7% following secondary stress (i.e. 3 preceding syllables). The remaining 4% of MODE 2 stops are found in the onset of the third syllable, e.g. /kuɭaɲ**c**:ari/ 'diamond dove'. In addition to stress pattern restrictions, MODE 2 stops cannot follow long vowels; and cannot be the first member of a consonant cluster (though they may occur in a consonant cluster following a nasal or a continuant). There are no (morphologically simple) words which contain multiple MODE 2 stops.

It is also notable that Warumungu is distinct from most Australian languages with a stop opposition in that the stop contrast in Warumungu is active following nasals [8], in addition to following continuants.

Table 2: *Distribution of Mode 2 stops by word length and location within word.*

PRECEDING SYLLABLES	TOTAL SYLLABLES IN WORD: 2σ	3σ	4σ	5σ	6σ	7σ
1σ	95	48	14	1	1	2
2σ			4			
3σ			12	1		

3.2. Duration

Duration of MODE 2 stops is significantly longer than that of MODE 1 stops across all places of articulation ($\beta = 136$, $t(424) = 11.6$, $p < .001$). Bilabial and palatal PoAs were significant. Overall mean duration of MODE 1 stops is 43 ms (s.d. 23), compared to 170 ms (s.d. 41) for MODE 2 stops – a duration ratio of approximately 1:4. Figure 3 shows mean durations for each stop mode across each place of articulation. Similar duration differences were found between MODE 1 and MODE 2 stops occurring in post-tonic position in di- or tri-syllabic roots ($\beta = 137.9$, $t(221) = 10.98$, $p < .001$).

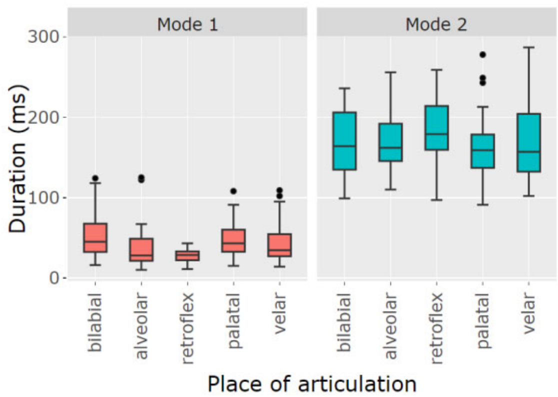

Figure 3: *Stop duration by mode and place of articulation.*

3.3. Voicing

MODE 2 stops have significantly lower mean HR compared to MODE 1 stops across all places of articulation ($\beta = –0.19$, $t(424) = –3.5$, $p < .001$). There was a significant interaction between palatal place of articulation and stop mode ($\beta = –0.04$, $t(424) = –2.2$, $p = .003$). The overall mean HR for MODE 1 stops is 0.66 (s.d. 0.07), and 0.47 (s.d. 0.10) for MODE 2 stops. Higher harmonic ratios are characteristic of voiced speech, so this indicates that MODE 1 stops are more voiced than MODE 2 stops. Mean HR for each stop mode across each place of articulation is shown in Figure 4. Similar differences in voicing were found between MODE 1 and MODE 2 stops occurring in post-tonic position in di- or tri-syllabic roots ($\beta = –0.22$, $t(221) = –4.0$, $p < .001$).

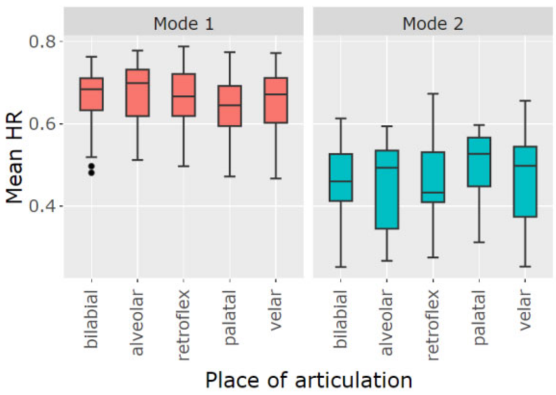

Figure 4: *Mean harmonic ratio for Mode 1 (left) and Mode 2 (right) stops, by place of articulation.*

3.4. Interaction between duration & voicing

Duration was a significant factor predicting voicing (β = –0.001, $t(433) = –15.7$, $p < .001$). The Pearson correlation between duration and voicing is –0.73, independent of stop mode (Figure 5). Density plots (Figure 6) reveal bi-modal distributions for both duration and mean HR (grouped by stop mode). The stop modes have an overlapping area of 0.06 for duration and 0.31 for mean HR.

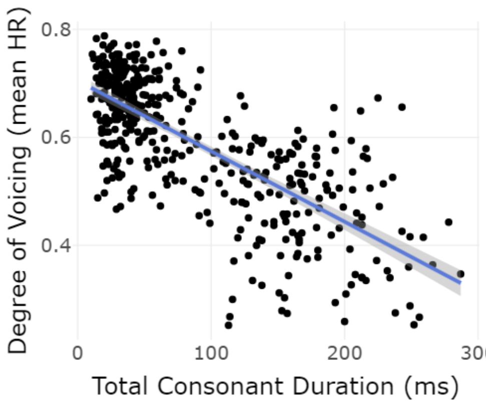

Figure 5: *Total stop duration as a predictor of degree of voicing.*

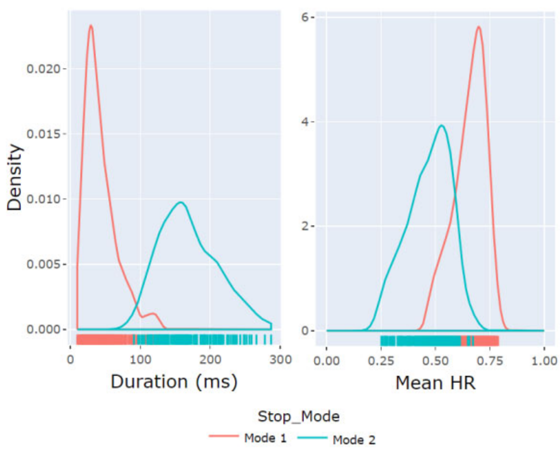

Figure 6: *Distribution of duration (left) and mean HR (right) across Mode 1(red) and Mode 2 (aqua) stops.*

4. Discussion

These data reveal that the contrast between stop modes in Warumungu is phonetically realised with duration and voicing. These phonetic contrasts are active across all places of articulation, for stops occurring intervocalically and in clusters.

Our research questions relate to whether one or both of these parameters is phonologically specified. Given the strong correlation between duration and voicing, we do not find support for H3, as duration and voicing reliably predict one another in root-medial environments. H1 is supported by the distributional and phonetic evidence. The opposition is not active word-initially, and consonant sequences are not found word-initially, and so the gemination analysis aligns with these phonotactic expectations. In contrast, a voicing opposition would be expected to occur in these environments.

Based on the density distributions, duration appears more distinctive, with very little overlap between the stop distributions (6%; in the 91-125ms duration range); compared with mean harmonic ratio (30%). This is quite unlike other analyses of stop duration, such as that of Jawoyn, in which duration was an important factor, yet there was considerable overlap between the duration of short and long tokens, and thus other factors are necessary to account for the stop distinction [14]. This is also evidenced by the higher stop duration ratio (1:4) compared to those reported for other languages, e.g. 1:2 in Bininj Gunwok [12], and (up to) 1:3 cross-linguistically [31]. Similarly, Warumungu Mode 2 stops are longer than (phonetically) long stops in Warlpiri [10], [32].

Taken together, the results provide strong evidence in favour of H1, in which duration is the sole phonological specification distinguishing Mode 1 stops and Mode 2 stops in Warumungu. In this analysis, stops have the same gestural specification, and differ in the number of linked timing units.

This analysis has a number of advantages over H2 (in which voicing is the sole phonological specification). Primarily, the overlap between modes is minimal for duration, compared to voicing. It avoids the proliferation of contrastive features in the Warumungu inventory: [voice] is not phonologically active in Warumungu, and the phonetic realisation of voicing arises through constriction duration. This is supported by the strong correlation between duration and mean harmonic ratio, which suggests that voicing is predictable from the implementation of duration. However, many acoustic properties of Warumungu stops remain unexplored. This includes the complex morphophonological interactions, as well as more fine-grained voicing parameters (e.g. VOT) and variation in the actual manner of articulation of (underlying) stops.

5. Conclusion

This study has found robust support for a stop opposition in Warumungu. Phonetically, this opposition is realised by duration and voicing. Phonologically, this opposition is reliably differentiated by total occlusion duration. The difference in voicing can be captured by the strong negative correlation between duration and voicing, rather than being phonologically specified. The data suggest a geminate analysis for Warumungu stops, rather than a fortis/lenis analysis. However, the geminate analysis requires an expansion of the syllable inventory of Warumungu—specifically either the existence of coda clusters (e.g. '*maŋk.ka* 'hole') or complex onsets (e.g. '*maŋ.kka*)—which warrants further consideration, as neither of these cluster types are otherwise evident in Warumungu.

6. Acknowledgements

We are grateful to the Warumungu speakers who teach us about their language. We also thank the two anonymous reviewers for their detailed comments on this paper. We gratefully acknowledge funding support from the Australian Research Council grant DP220102925 and from the 2024 Macquarie University ECR Enabling Scheme.

7. References

[1] J. Fletcher and A. Butcher, "Sound patterns of Australian Languages," in *The Languages and Linguistics of Australia*, H. Koch and R. Nordlinger, Eds., Berlin, Boston: De Gruyter Mouton, 2014, pp. 91–138. doi: 10.1515/9783110279771.91.

[2] R. M. W. Dixon, *Australian languages : their nature and development*. New York: Cambridge University Press, 2002. [Online]. Available: http://www.loc.gov/catdir/toc/cam024/2001037958.html

[3] J. Simpson, "Warumungu (Australian – Pama-Nyungan)," in *The Handbook of Morphology*, A. M. Zwicky and A. Spencer, Eds., 2017, pp. 707–736. doi: 10.1002/9781405166348.ch32.

[4] A. Capell, "Notes on the Waramunga language, Central Australia," *Oceania*, vol. 23, no. 4, pp. 297–311, Jun. 1953, doi: 10.1002/j.1834-4461.1953.tb00199.x.

[5] K. Hale, *Warramunga notes. Sandy Nandy and George Bruce from Tennant Creek*. MS 863. Australian Institute of Aboriginal and Torres Strait Islander Studies, 1959.

[6] J. Heath, *Warramunga grammatical notes, based on fieldwork in April/June 1977*. PMS 3804. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies, 1977.

[7] P. Chakravarti, *WaRumungu*. PMS 332. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies.

[8] E. R. Round, "Segment inventories," in *The Oxford Guide to Australian Languages*, 1st ed., C. Bowern, Ed., Oxford University PressOxford, 2023, pp. 96–105. doi: 10.1093/oso/9780198824978.003.0010.

[9] M. Browne, *A Grammar of Warlmanpa*. Canberra: ANU Press, 2024. doi: 10.22459/GW.2024.

[10] C. Pentland and M. Laughren, "Distinguishing prosodic word and phonological word in Warlpiri: prosodic constituency in morphologically complex words," *Proceedings of the 2004 Conference of the Australian Linguistics Society*, no. 1995, p. 17, 2004.

[11] R. L. Bundgaard-Nielsen and C. O'Shannessy, "Voice onset time and constriction duration in Warlpiri stops (Australia)," *Phonetica*, vol. 78, no. 2, pp. 113–140, Apr. 2021, doi: 10.1515/phon-2021-2001.

[12] H. Stoakes, "An Acoustic and Aerodynamic Analysis of Consonant Articulation in Bininj Gun-wok," PhD, The University of Melbourne, Melbourne, 2013.

[13] A. Butcher, "'Fortis/lenis' revisited one more time: the aerodynamics of some oral stop contrasts in three continents," *Clinical Linguistics & Phonetics*, vol. 18, no. 6–8, pp. 547–557, Sep. 2004, doi: 10.1080/02699200410001703565.

[14] B. Evans and F. Merlan, "Stop contrasts in languages of Arnhem Land: From the perspective of Jawoyn, Southern Arnhem Land," *Australian Journal of Linguistics*, vol. 24, no. 2, pp. 185–224, Oct. 2004, doi: 10.1080/0726860042000271825.

[15] B. J. Baker, *Word structure in Ngalakgan*. in CSLI publications. Stanford, Calif: Center for the Study of Language and Information, 2008.

[16] C. Henton, P. Ladefoged, and I. Maddieson, "Stops in the World's Languages," *Phonetica*, vol. 49, no. 2, pp. 65–101, Mar. 1992, doi: 10.1159/000261905.

[17] L. Lisker and A. S. Abramson, "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements," *WORD*, vol. 20, no. 3, pp. 384–422, Jan. 1964, doi: 10.1080/00437956.1964.11659830.

[18] N. Reid, *Ngan'gityemerri: a language of the Daly River region, Northern Territory of Australia (Outstanding Grammars from Australia 6)*. Munich: Lincom Europa, 2011.

[19] J. Mansfield, *Murrinhpatha morphology and phonology (Pacific Linguistics 653)*. Berlin & Boston: De Gruyter Mouton, 2019.

[20] R. Mailhammer, S. Sherwood, and H. Stoakes, "The inconspicuous substratum: Indigenous Australian languages and the phonetics of stop contrasts in English on Croker Island," *English World-Wide*, vol. 41, no. 2, pp. 162–192, 2020.

[21] W. R. Leben, "A Metrical Analysis of Length," *Linguistic Inquiry*, vol. 11, no. 3, pp. 497–509, 1980.

[22] S. Disbray and D. N. Stokes, *Warumungu Picture Dictionary*. Alice Springs, N.T.: IAD Press, 2005.

[23] T. M. Inc, *MATLAB version: 24.1.0 (R2024a)*. (2024). The MathWorks Inc., Natick, Massachusetts, United States. [Online]. Available: https://www.mathworks.com

[24] X. Sun, "Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio," *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, p. I-333-I–336, 2002.

[25] T. Drugman and A. Alwan, "Joint Robust Voicing Detection and Pitch Estimation Based on Residual Harmonics," 2020, *arXiv*. doi: 10.48550/ARXIV.2001.00459.

[26] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2024. [Online]. Available: https://www.R-project.org/

[27] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015, doi: 10.18637/jss.v067.i01.

[28] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017, doi: 10.18637/jss.v082.i13.

[29] M. Pastore, P. A. D. Loro, M. Mingione, and A. Calcagni', *overlapping: Estimation of Overlapping in Empirical Distributions*. 2022. [Online]. Available: https://CRAN.R-project.org/package=overlapping

[30] C. Sievert, *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC, 2020. [Online]. Available: https://plotly-r.com

[31] P. Ladefoged and I. Maddieson, *The sounds of the world's languages*. Oxford: Blackwell Publishers, 1996.

[32] R. L. Bundgaard-Nielsen and C. O'Shannessy, "Voice Onset Time and Constriction Duration in Warlpiri Stops (Australia)," in *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, 2019.