Scaling processes of clause chains in Pitjantjatjara

Rebecca Defina, Catalina Torres, Hywel Stoakes

School of Languages and Linguistics, The University of Melbourne, Australia

rebecca.defina@unimelb.edu.au, catalinat@student.unimelb.edu.au,
hstoakes@unimelb.edu.au

Abstract

Clause chains are a syntactic strategy for combining multiple clauses into a single unit. They are reported in many languages, including Korean and Turkish. However, they have seen relatively little focused research. In particular, prosodic features are often mentioned in descriptions of clause chaining, however there have been vanishingly few investigations. Corpus-based studies of the prosody of clause chains in two unrelated languages of Papua New Guinea report that they are typically produced as a sequence of Intonation phrases united by pitch-scaling of the L% boundary tones in each clause with only the final, finite, clause descending to a full L%. The present study is the first experimental investigation of the prosody of clause chains in Pitjantjatjara.

This paper focuses on one type of clause chain found in the Australian Indigenous language Pitjantjatjara. We examine a set of 120 clause chains read out by three native Pitjantjatjara speakers. Prosodic analysis reveals that these Pitjantjatjara clause chains are produced within a single Intonational Phrase. Speakers do not pause between the clauses in the chain, there is consistent linear downstep throughout the phrase and additionally phrase final lowering occurs at the end of the utterance. This differs from previous impressionistic studies of the prosody of clause chains.

Index Terms: Australian languages, speech prosody, syntax-prosody interface, F0 scaling

1. Introduction

This study presents an experimental investigation of scaling processes and associated prosodic structure of clause chains in Pitjantjatjara, an Indigenous language of Australia. Clause chains are a syntactic construction which combine multiple clauses into a single coherent unit. They consist of one or more non-finite clauses and one finite clause. The non-finite verbs are marked with a distinctive affix which marks them as dependent and are commonly referred to as medial verbs. The tense, aspect, and mood of the non-finite clauses are inferred from the values of the finite verb. A classic example from Pitjantjatjara can be seen in (1), with verbs marked in bold.

1. *Munu kuka* **nyuti-_ra_ tjali-_ra_**
   and.SS meat **tie-MV lift.onto.head-MV**

ma-kati-ra   *ngura-ngka* **ma-tjarpa-_ra_**
away-bring-MV camp-LOC   **away-enter-MV**

wa_ni_-ra   **tja_ra_rpu-ngkula** **pau-ra**   **wanti-ngu**
throw-MV   **dig.pit-MV**   **cook-MV**   **leave-PST**[1]

'And/then he tied up the meat (with sinew), put it on his head and carried it (to camp), he entered the camp and threw it down and dug a cooking pit, left it to cook.'   [1, p. 220]

Clause chains are common across a wide range of languages including Turkish, Korean, and many languages of Papua New Guinea. They are sometimes referred to as the spoken equivalent of a paragraph, providing a formal syntactic unit for grouping together a series of clauses. While impressionistic prosodic features are often discussed in the descriptions of these constructions, there have been no dedicated experimental investigations of prosodic structure within clause chains.

This study aims at providing an analysis of scaling processes related to clause chains within the autosegmental-metrical (AM) approach [2], [3]. In AM, the term tonal space refers to an interval where high (H) tones are at the top and low (L) tones towards the bottom level of a band of Fundamental frequency (F0) values [3]. Variations in the tonal space can serve as phonological representation of prosodic constituency [4], [5] or interact with other linguistic domains, such as syntactic structure. As an example, it has been shown that a major syntactic break between a noun phrase and a verb phrase in French can trigger an intermediate phrase (ip) boundary [6], [7]. The scaling of H tonal targets is influenced by *downstep*, a language specific phonological process that causes a narrowing of the tonal space [2]. Final lowering, by contrast, is a tonal event that cannot be predicted by the downward trend induced through downstep and is realized as a much lower F0 [8].

Based on impressionistic, corpus-based investigations of clause chains in Matukar Panau, an Oceanic language of Papua New Guinea, Mansfield and Barth [9] propose that clause chains are typically mapped to a prosodic unit above the Intonation Phrase (IP), namely the utterance phrase. They report that Matukar Panau clause chains are typically realized as a sequence of IPs with pitch scaling of the L% boundary tones uniting the whole chain as a single prosodic unit. Similarly, Sarvasy [10] reports that clause chains in Nungon, a Finisterre-Huon language of Papua New Guinea, are also prosodically linked as a whole via pitch-scaling of the L% boundary tones in the medial clauses with only the final, finite, clause descending to a full L%. However, no experimental

---

[1] Abbreviations used in glosses: ACC 'accusative', 'DAT 'dative', ERG 'ergative', LOC 'locative', MV 'medial verb', PST 'past'.

studies have been devoted to the examination of clause chains and their relation to the prosodic hierarchy.

Pitjantjatjara (Glottocode: pitj1243, ISO 639-3: pjt) is spoken by roughly 3000 people [11] in the desert region of central Australia. The language is classified as part of the Wati-Nyungic branch of Pama-Nyungan [12] and forms part of the Western Desert dialect chain which extends across much of central and western Australia. Pitjantjatjara and its closest sister dialect, Yankunytjatjara, are largely mutually intelligible and are both described in one grammar [13] and dictionary [14]. Pitjantjatjara has a left-headed binary foot structure, with final syllables treated as extra-metrical [15]. Primary stress is word-initial and initial/stressed syllables are strongly associated with a pitch peak which shows significantly longer duration than other syllables in the word. We are specifically interested in the F0 height of H tonal targets associated with accented syllables in Pitjantjatjara clause-chains.

Clause chains are the standard method for describing connected sequences of actions in Pitjantjatjara. They can be fairly long, such as in example (1) with eight verbs. More commonly, however, they consist of two or three verbs [16]. The finite verb is typically, but not necessarily, chain final.

Goddard [13], [17] describes two main types of clause chain (serial verb constructions in his account) in the very closely related dialect, Yankunytjatjara. He refers to these as Loose and Tight. Loose clause chains may contain many verbs, each with their own nominal arguments. In contrast, Tight clause chains contain only two verbs and these verbs must be adjacent with no intervening material. Importantly for the present paper, Goddard also notes prosody as a discriminating feature. Tight clause chains are produced as a single intonational unit, while the clauses in Loose chains typically contain pauses or intonation boundaries. Defina [16] reports that the distinction between Loose and Tight clause chains is often ambiguous and difficult to draw in practice, since the majority of Loose clause chains consist of two adjacent verbs.

2.   Research questions

The present study is the first part of a wider project on the prosodic structure of Pitjantjatjara clause chains. In this experiment, we examine Pitjantjatjara utterances consisting of Loose clause chains to determine their prosodic correlates. We test the predictions made by earlier impressionistic research, specifically: Goddard's [13], [17] report of a phrase boundary between verbs and the reports of clause chains in PNG realized as a sequence of IPs with a constant downstep rule [9], [10]. We additionally manipulate the length of clause chains to examine whether the size of the syntactic constituent has an effect on prosodic realization. Within our experimental paradigm, we evaluate four hypotheses:

H1. Prosodic constituency and clause chains
This hypothesis predicts that verbs in clause chains are preferably produced as single prosodic units, namely IPs, marked through intervening pauses.

H2. Downstep of H peaks
We hypothesize that prosodic constituency is marked through F0 scaling. We predict that the H peaks of consecutive verbs in the clause chain obey a constant downstep rule. This would mean that F0 peaks in clause chains in the two length conditions below share a positive linear relationship.

H3. Restructuring of IP internal scaling
We predict that there will be clause chain internal intermediate phrase boundaries, with the F0 of peaks following this break reset to a higher level.

H4. Final lowering is utterance final
Further, we predict that final lowering is utterance final. This means that when the final verb is not utterance final, neither it, nor any of the medial verbs, trigger final lowering.

3.   Methods

3.1   Participants

Three native speakers of Pitjantjatjara, all female, aged 48-58 (M = 54) participated in this study. Speakers are also fluent in English and literate in both languages.

3.2   Materials

Examples (2) and (3) show the two conditions included in the experiment: Two-verb Loose clause chains, with 12 token sentences containing two consecutive verbs and Three-verb Loose clause chains with 12 token sentences containing three consecutive verbs. Stimuli were constructed in consultation with native speakers. Word order is highly flexible in Pitjantjatjara, but we ensured that all sentences were Subject $Verb_{Medial}$ ($Verb_{Medial}$) $Verb_{Finite}$ Object. The Two-verb clause chains consisted of the medial verb *mira_ra* 'call for' and the finite verb *ngurinu* 'looked for' as in (2). Twelve distinct items were created by varying the subject and object to other proper names or pronouns. To ensure a continuous F0 trace, stops were avoided as much as possible.

2. *Wanyima-lu mira-ra nguri-nu   Umatji-ku*
   name-ERG   call-MV   look.for-PST name-DAT
   'Wanyima called for and looked for Umatji.'
   [Subject] [Verb_Medial] [Verb_Finite]   [Object]

The Three-verb clause chains were based on the Two-verb clause chains by adding a third verb as the new finite verb. So *nguri* 'look for' now appears as a medial verb and the third and final finite verb is *nyangu* 'saw', as in example (3).

3. *Wanyima-lu mira-ra nguri-ra   nya-ngu Umatji-nya*
   NAME-ERG   call-MV   look.for-MV see-PST   NAME-ACC
   'Wanyima called for, looked for, and saw Umatji.'
   [Subject] [Verb_Medial^1] [Verb_Medial^2] [Verb_Finite] [Object]

We examine only two and three verb clause chains as longer chains are rare in natural conversation [16]. Comparing two versus three verb chains will provide an initial indication of any difference according to the length of the chain and provide a greater time window with more material for any prosodic phrasing to become apparent.

3.3   Procedure

Participants read each of the 24 stimuli sentences four times, plus 48 filler sentences containing other variants of two- and three-verb clause chains for further studies. The sentence order was randomized for each speaker. If a participant faltered or was unhappy with their reading of a given sentence, they repeated it. Participants were recorded in a quiet room with soft furnishings in the community of Pukatja. All recordings were made with a headset microphone and a H4 ZOOM recording as

mono at 44.1 kHz/16 bit. Participants were paid for their time according to rates recommended by Ninti One. Testing was done in Pitjantjatjara and the experimenter conversed with the participants in Pitjantjatjara for the duration of the experiment.

3.4 Analysis

To construct the database, first, utterances were transcribed noting deviations in recordings from the script, for example minor variations in the subject or object such as the use of the abbreviated Umanya for the name Umatjinya. For this analysis, a subset of 120 of the total 288 target tokens were selected.

All audio files were normalised to -3db and split into single utterance files. Each file was automatically segmented, forced aligned and phonetically labelled using the Munich Automatic Segmentation tool (MAuS) [18]. A User Defined Grapheme to Phoneme mapping was used and input into the G2P tool [19] using The Australian Aboriginal Language Parameter model [20] accessed at the Bavarian Archive for Speech Signals website (http://hdl.handle.net/11858/00-1779-0000-000C-DAAF-B).

Each labelling file was output from MAuS as a Praat textgrid. Each textgrid was compiled and edited using "rPraat" [21] adding five principle tiers (Utterance, Word, Grammatical Category, Syllable number and Phonetic Segment). All utterances were manually checked and hand corrected at a word and phonetic level using Praat [22].

Subsequent to checking, prosodic labelling was added on a separate Tone tier, labelling High (H) and Low (L) tones as shown in Figure 1. The tones were labelled with reference to the pitch (F0) trace and corrected phonetic tier with tones of interest marked with $H/L^{+number}$ from left to right and LF marking a final low tone. Pauses were marked on the tone tier when observed using the automatically labelled phonetic tier as reference.
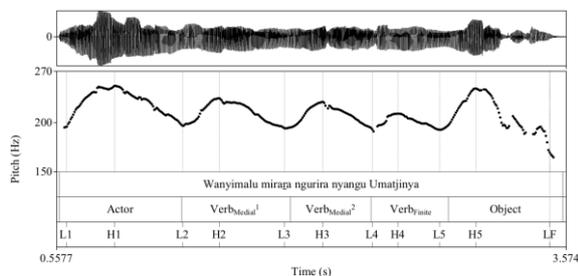

Figure 1.: Pitch (F0) trace of utterance from the Three-verb condition (*Wanyimalu miraṟa nguriṟa nyangu Umatjinya* Eng. 'Wanyima called for, looked for, and saw Umatji.').

All Praat textgrids were then imported into the Emu Speech Database Management System [23]. Prosodic hierarchies were constructed by creating links between each tier (level) and then automatically linking them using time stamps (For more information see [24]).

3.5 Statistical analyses

Measurements taken in Hz for tones were converted into semitones (benchmark 100) [25]. Data were analysed using linear mixed effects models examining differences between H and L tones. Post-pausal tones were excluded from these analyses. Estimated marginal means were obtained for factor comparisons and *p*-values were adjusted with the Tukey

method. Statistical analyses were carried out in R [26] utilising linear mixed models within the statistics packages lme4 and emmeans [27], [28].

4 Results

4.1 Prosodic constituency and clause chains
In our corpus of 120 clause chain utterances, 74% were produced with all verbs within one IP together with the preceding subject and final object. We observed the insertion of a spontaneous short pause in 26% of the data. In 87% of the cases this pause was placed either prior to the onset of the verb series or after it (before the sentence final object). Only in 3% of the utterances was there a pause between the chained clauses. This shows a tendency for clause chains to be mostly realized within the same IP. In the following analysis, we only examine peaks (N=518) and lows (N=568) with no adjacent pause.

4.2 Downstep of H peaks
Figure 2 shows a linear relationship between $Verb_{Medial}$ and $Verb_{Finite}$ peaks in the Two-verb condition. A linear mixed effects model with the fixed factors TONE and CONDITION, together with the random factor participant (N=518) examined whether there is a relationship in height between peaks in the clause chain. A statistically significant relationship was found in all cases. Factor comparisons confirmed this for the Two-verb condition (Est. 1.04, SE 0.1 semitones, *p*<.0001), also for the Three-verb condition, for $Verb_{Medial}^1$ and $Verb_{Finite}$ (Est. 2.02, SE ±0.1 semitones, *p*<.0001) and for $Verb_{Medial}^2$ and $Verb_{Finite}$ (Est. 0.98, SE ±0.1 semitones, *p*<.0001) and for $Verb_{Medial}^2$ and $Verb_{Medial}^1$ (Est. 1.04, SE ±0.1 semitones, *p*<.0001). This suggests that peaks in the chained clauses gradually decrease by approx. 1 semitone.
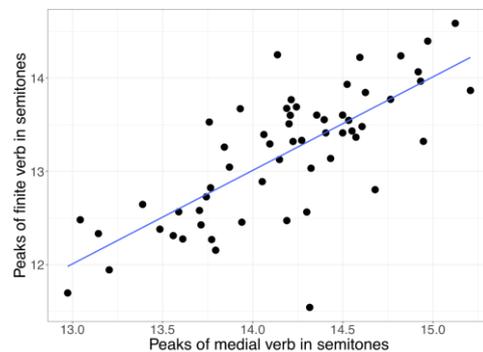

Figure 2.: Scatter plot with fitted regression line of values (in semitones) of $Verb_{Medial}$ peaks as a function of peaks in $Verb_{Finite}$ position in the Two-verb condition.

4.3 Restructuring of IP internal scaling
Figure 3 shows boxplots of peaks in semitones in the Two-verb and Three-verb conditions. Note that in the Three-verb condition, the Object peak is raised to a higher level than that of $Verb_{Finite}$. The linear mixed effects model from 4.2 is used. Factor comparisons reveal that the difference between peaks in Object and $Verb_{Finite}$ is not statistically significant (Est. -0.26, SE ± 0.13 semitones, *p*=0.7 (n.s)). This suggests there is an upwards shift in the peak of Object, blocking the downstep trend found in 4.2. However, no raising of F0 is found in the Two-verb condition.
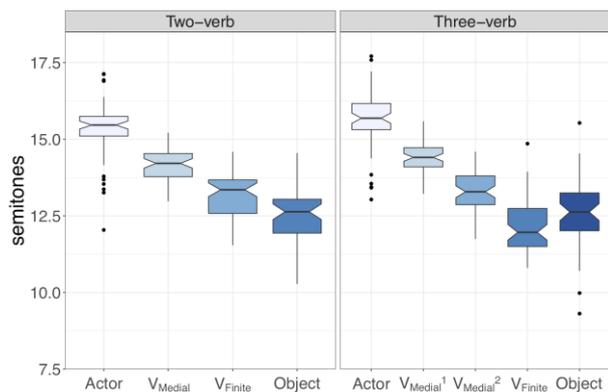
Figure 3.: Notched box plots of successive peaks (in semitones) across utterances in the Two-verb and Three-verb conditions (post-pausal peaks excluded).

4.4 Final lowering is utterance final

Figure 4 shows that LF exhibits the lowest values (across all L tones) in the Two-verb and Three-verb conditions. A linear mixed effects model with the fixed factors TONE and CONDITION, together with the random factor participant (N=568) evaluated whether there is a difference between L tones. Results show that the difference between the sentence final tone (LF) and preceding low tones (L1, L2, L2, L4, L5) is statistically significant (Est. -4.83, SE ±0.27 semitones, $p<.0001$). This confirms LF is the lowest tone across the two conditions.
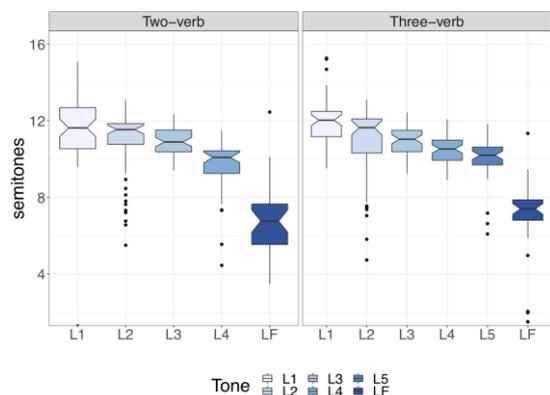

Figure 4.: Notched box plots of successive low tones (in semitones) across utterances in the Two-verb and Three-verb conditions (post-pausal low tones excluded).

5 Discussion

We set out to investigate the prosodic correlates of Pitjantjatjara Loose clause chains. In this experimental context, speakers produced two and three-verb clause chains within a single IP. This was indicated by the lack of internal pauses, the linear downstep in H peaks throughout the utterances, and the presence of a single final lowering L% at the end of the utterance.

We observed one difference between two- and three-verb clause chains. Three-verb clause chains tended to feature an F0 reset between the final verb and the object which could be associated with the demarcation of an intermediate phrase

boundary. Intriguingly, this prosodic break occurred between the verb and the object rather than between the constituent clauses as previously reported for clause chains in other languages [9], [10]. The patterning of and motivation for this F0 reset remain questions for further investigation. It may be related to variations in Pitjantjatjara word order or related to particular verb and object pairings. We intend to investigate these possibilities in future work.

Our finding that Pitjantjatjara Loose clause chains are typically produced within a single IP with no pauses between the verbs is counter to Goddard's [13], [17] description of the very similar dialect Yankunytjatjara. Goddard describes a prosodic phrase break or pause as a typical feature which distinguished Loose clause chains from Tight clause chains. It remains a question for further research whether there are any other prosodic features which distinguish Loose and Tight clause chains in Pitjantjatjara. However, this study suggests that Intonation Phrase boundaries and pauses will not be useful diagnostic features.

Our findings also differ from previous investigations of the prosody of clause chains cross-linguistically. Separate corpus-based investigations of two unrelated languages of PNG found that clause chains tend to be produced as a series of IPs united by a downstepped pitch-scaling of the L boundary tones [9], [10]. It is an intriguing question for further research whether these differences relate to differences between the languages or between the methods used to investigate them. Some researchers have suggested that clause chains and other similar syntactic constructions may have been exoticized in the literature with a focus on less typical examples [29], [30]. It is possible that impressionistic studies may have been influenced by this bias. It is also likely that the experimental conditions and restricted set of verbs used in the present study do not fully capture the range of prosodic realizations of Pitjantjatjara clause chains.

6 Conclusions

This study presents the first experimental investigation of the prosodic correlates of clause chains. We find that Pitjantjatjara clause chains are typically produced within Intonation Phrases. This differs from previous corpus-based and impressionistic reports of the closely related dialect Yankunytjatjara as well as unrelated languages of Papua New Guinea. This divergence raises questions about potential variation among clause chains cross-linguistically and across different contexts.

7 Acknowledgements

8 References

[1] P. Eckert and J. Hudson, *Wangka wiru: a handbook for the Pitjantjatjara language learner*. Adelaide: South Australian College of Advanced Education, 1988.

[2] J. Pierrehumbert, "The phonology and phonetics of English intonation," PhD Thesis, Massachusetts Institute of Technology, 1980.

[3] D. R. Ladd, *Intonational phonology*. Cambridge University Press, 2008.

[4] H. Truckenbrodt, F. Sandalo, and B. Abaurre, "Elements of Brazilian Portuguese intonation," *J. Port. Linguist.*, vol. 8, no. 1, pp. 75–114, 2009.

[5] M. E. Beckman and J. Pierrehumbert, "Intonational Structure in Japanese and English," *Phonol. Yearb. III*, pp. 15–70, 1986.

[6] M. D'Imperio and A. Michelas, "Pitch scaling and the internal structuring of the Intonation Phrase in French," *Phonology*, vol. 31, no. 1, pp. 95–122, 2014.

[7] C. Torres, J. Fletcher, and G. Wigglesworth, "Phrasing and constituent boundaries in Lifou French," in *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*, 2019.

[8] J. Pierrehumbert and M. E. Beckman, *Japanese tone structure*. Cambridge, MA: MIT Press, 1988.

[9] J. Mansfield and D. Barth, "Clause chaining and the utterance phrase: Syntax–prosody mapping in Matukar Panau," *Manuscr. Based Talk Present. Univ. Melb. 2018*, in prep.

[10] H. Sarvasy, "Breaking the clause chains: Non-canonical medial clauses in Nungon," *Stud. Lang.*, vol. 39, no. 3, pp. 664–696, 2015.

[11] Australian Bureau of Statistics, "Australia, Language spoken at home by sex (LGA)." 2016, Accessed: Jul. 11, 2017. [Online]. Available: http://stat.data.abs.gov.au/Index.aspx?DataSetCode=A BS_C16_T09_LGA.

[12] C. Bowern and Q. Atkinson, "Computational phylogenetics and the internal structure of Pama-Nyungan," *Language*, vol. 88, no. 4, pp. 817–845, 2012.

[13] C. Goddard, *A grammar of Yankunytjatjara*. Alice Springs: Institute for Aboriginal Development, 1985.

[14] C. Goddard, *Pitjantjatjara/Yankunytjatjara to English dictionary*, Rev. 2nd ed. Alice Springs: IAD Press, 1996.

[15] M. Tabain, J. Fletcher, and A. Butcher, "Lexical stress in Pitjantjatjara," *J. Phon.*, vol. 42, pp. 52–66, 2014.

[16] R. Defina, "Acquisition of Pitjantjatjara Clause Chains," *Front. Psychol.*, vol. 11, p. 541, 2020, doi: 10.3389/fpsyg.2020.00541.

[17] C. Goddard, "Verb serialisation and the circumstantial construction in Yankunytjatjara," in *Complex sentence constructions in Australian languages*, P. Austin, Ed. Amsterdam: John Benjamins, 1988, pp. 177–192.

[18] F. Schiel, "Automatic Phonetic Transcription of Non-Prompted Speech," presented at the Proceedings of the ICPhS, San Francisco, 1999, pp. 607–610.

[19] U. D. Reichel, "PermA and Balloon: Tools for string alignment and text processing," presented at the Proceedings of Interspeech, Portland, Oregon, 2012, p. paper no. 346.

[20] H. Stoakes and F. Schiel, "A Pan-Australian acoustic model: Automatic alignment using the Munich Automatic Segmentation tool," presented at the Paper presented at ALS 2017: Conference of The Australian Linguistics Society, 2017.

[21] T. Bořil and R. Skarnitzl, "Tools rPraat and mPraat," in *Proceedings of Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic*, Cham, Sep. 2016, pp. 367–374.

[22] P. Boersma and D. Weenink, *Praat: doing phonetics by computer*. 2020.

[23] R. Winkelmann, J. Harrington, and K. Jänsch, "EMU-SDMS: Advanced speech database management and analysis in R," *Comput. Speech Lang.*, 2017.

[24] R. Winkelmann, K. Jänsch, S. Cassidy, and J. Harrington, *Manual for the EMU Speech Database Management System*. Institute of Phonetics and Speech Processing, Ludwig Maximillian University, Munich, 2020.

[25] F. Nolan, "Intonational equivalence: an experimental evaluation of pitch scales," in *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, 2003, vol. 39.

[26] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017.

[27] D. Bates, M. Maechler, B. Bolker, and S. Walker, *lme4: Linear mixed-effects models using Eigen and S4*. 2015.

[28] R. V. Lenth, H. Singmann, J. Love, P. Buerkner, and M. Herve, *Estimated marginal means, aka least-squares means*. 2019.

[29] M. Ponsonnet, "Nominal Subclasses in Dalabon (South-western Arnhem Land)," *Aust. J. Linguist.*, vol. 35, no. 1, pp. 1–52, 2015, doi: 10.1080/07268602.2015.976900.

[30] A. Rumsey, L. W. Reed, and F. Merlan, "Ku Waru Clause Chaining and the Acquisition of Complex Syntax," *Front. Commun.*, vol. 5, p. 19, 2020, doi: 10.3389/fcomm.2020.00019.