PROSODY AND WORD ORDER IN MARKING FOCUS WITHIN DJAMBARRPUYŊU NOUN PHRASES

Kathleen Jepson

Institute of Phonetics and Speech Processing (IPS), LMU Munich
k.jepson@phonetik.uni-muenchen.de

ABSTRACT

This paper examines word order and the prosodic realisation of focus within noun phrases in Djambarrpuyŋu, an Australian Indigenous language with free word order.

Within noun phrases, the prosodic marking of focus differs across languages and may include phonological and phonetic means. Languages with greater syntactic flexibility, however, may not prosodically encode information status within noun phrases in the same ways, or at all, though this is under-explored.

Djambarrpuyŋu noun phrases with focus on an adjective, a noun, or the noun phrase were elicited. Analysis of the data showed that focus condition was a predictor of word order within the noun phrase, specifically, focused adjectives occurred before given nouns. But neither phonological (accent distribution, accent type) nor phonetic (f0 peak, RMS amplitude peak) correlates of prominence consistently contributed to distinguishing the focus conditions.

Keywords: prosody, focus, information status, word order, Djambarrpuyŋu

1. INTRODUCTION

This paper examines word order and the prosodic realisation of focus within noun phrases in Djambarrpuyŋu (ISO 639-3: djr), an Australian Indigenous language with free word order, spoken in northern Australia by ~4,000 people (see Fig. 1).

The relationship between information structure and information status and prosody is well established cross-linguistically [1], [2]. Focus, which indicates the presence of alternatives [3], and givenness, which considers how activated or accessible information is in the minds of interlocutors [4], are commonly examined. Focused and new elements are often accented, sometimes with a special pitch accent, and are realised with concomitant phonetic cues to the phonological designation of "accented"; longer duration, increased intensity, and higher fundamental frequency (henceforth f0), resulting in greater phonetic prominence. Given information is acoustically less prominent and may be deaccented [2], [5]. In languages that resist deaccenting (see e.g., [6], [7]), phonetic cues may still encode focus and information status. Languages that allow greater syntactic variability may also rely less on prosodic cues to information structure or status [7], [8].

Figure 1: A map of Australia showing where Djambarrpuyŋu is spoken.

The prosodic realisation of focus within noun phrases (henceforth NPs)—a noun and its semantically-related modifiers such as adjectives—shows considerable cross-linguistic variation [9]. However, the aforementioned trends of focus marking are observed [7], [9], [10]. In some languages such as Italian and Japanese, distribution of intonational accents is not a significant predictor of focus nor information status within NPs [7], [11]. However, phonetic cues may still be employed in these languages. In Japanese NPs, adjectives and nouns have higher f0 range when new information focused than given [11]. Similar NP-internal pitch variability has also been reported for Hungarian, which has a syntactic focus position [12].

Djambarrpuyŋu presents a point of difference from previously investigated languages because it allows free word order, including within noun phrases [13]–[15].[1] Speakers may considerably alter the syntactic structure of an utterance as a means of conveying information structure, and clause initial position has pragmatic importance or is a position of prominence in Australian languages [16]. However, this is not a formal structural expression of focus. Prosodic investigations of Australian languages have found that focused constituents that are fronted may additionally occur in their own intonational phrase, prosodically dislocated from following material, and have the highest f0 (associated with a pitch accent) in the utterance [17]. Narrow focus may also be encoded by a rising pitch accent [18], [19].

Djambarrpuyŋu has fixed stress on the first syllable of the word, and shows similarities to head/edge-prominence languages [20]. Previous prosodic analysis of Djambarrpuyŋu [21] suggests that deaccenting is an uncommon strategy for encoding given information, as reported for other Australian languages [18], [19]. But there is evidence that if syntactic position is maintained (see [22]), deaccenting can be used to encode accessible information in Djambarrpuyŋu [21].

In this paper, word order and prosodic cues to focus are examined within contiguous NPs (i.e., in which target words are adjacent), considering patterns of accentuation and accent type, as well as the acoustic measures of peak f0 and peak root mean square (henceforth RMS) amplitude. This study aims to describe if and how focus is encoded in Djambarrpuyŋu noun phrases by these measures, whether the order of the adjective and noun in NPs varies consistently by focus condition, and how syntactic and prosodic encoding may interact. To investigate these topics, data from a production task eliciting noun phrases were examined.

2. METHODS

2.1. Participants

Thirteen native Djambarrpuyŋu speakers (7 women and 6 men, mean age 45, age range 20-69) were recorded in Milingimbi, N.T., Australia. All participants were familiar with related language varieties, other Aboriginal languages, and Australian English. Participants were paid for their time.

2.2. Stimuli

Five nouns and five colour-term adjectives were selected for creating the materials, and were matched in all combinations resulting in 25 unique picture tiles (see [9]). The target words were selected to be both segmentally and morphologically diverse, and broadly representative of the types of structures observed in Djambarrpuyŋu words; targets words were one to five syllables in length, and included reduplicated forms and compound nouns.

Twelve 2×2 grids were created using the 25 tiles (see e.g., Fig. 2). There were three combination types corresponding to three focus conditions: 1) each tile had a different picture of a different colour resulting in whole noun phrase focus (henceforth NPF); 2) all the pictures were different but of the same colour resulting in focus on the noun (henceforth NF); and 3) all the pictures were the same but of different colours resulting in focus on the adjective (henceforth AF). The design is summarised in Table 1.

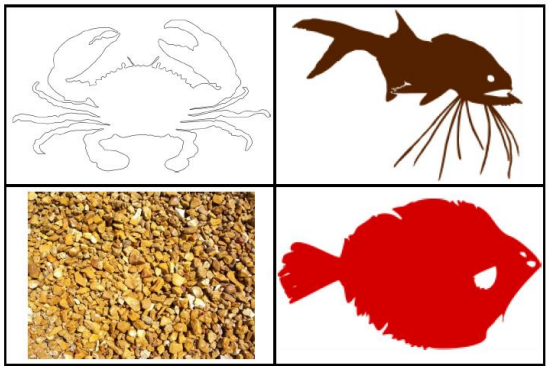

Figure 2: A grid from the NPF condition, using nouns *nyoka* ‘crab’, *malmuŋu* ‘threadfin salmon’, *minyirrminyirr* ‘gravel’, *raŋanŋaniŋ* ‘sole’, and adjectives *watharr* ‘white’, *gurrŋan* ‘brown’, *buthalak* ‘yellow’, *miku* ‘red’. Other target words, not shown, were *buthuruwuŋgan* ‘hammer oyster’, and *mol* ‘black’.

	Focused element	
Condition	Adjective	Noun
NPF	✓	✓
NF		✓
AF	✓	

Table 1: Summary of the focused element in the three focus conditions.

2.3. Procedure

Before commencing the task, participants were shown the materials, and the nouns and colour terms were discussed. The participant and their interlocutor (the author, or fellow participant) were seated across from each other at a table. Each participant completed three blocks of four trials, corresponding to the three focus conditions. In each trial, the participant was shown one of the 2×2 grids on a computer screen not visible to the interlocutor. The interlocutor had a blank, hardcopy 2×2 grid in front of them. Between the participant and interlocutor were five laminated coloured picture tiles that included those on the participant’s screen plus an additional tile appropriate for the focus condition. For each tile in the grid (i.e., turn), the participant was asked to described the tile, and instruct the interlocutor where it was located on the grid using directionals “top”, “bottom”, “left”, and “right”. Example frame sentences, e.g., (1) were discussed with participants in the instructions; however, participants varied word order at the utterance level, and also used other verbs such as intransitive “to go”, though the effect of verb is not considered here (see [23], [24]). Participants were encouraged to use both the noun and adjective words but were free to give instructions however they preferred, resulting in naturalistic speech. There was some variability in the words used (e.g., *buthuruwuŋgan~buthuruwaṯu*), which is not explored

in the current study. The interlocutor had to place the correct tile in the correct position on the blank grid. The trials were presented in three blocks of the NPF-, NF-, and AF-conditions.

(1) *Nyoka watharr nhe dhu rulwaŋdhun*
crab white 2SG FUT put.down
garrwar-lil wiŋ'kuŋu-lil
top-ALL left-ALL
'You will put the white crab to the top left.'

Audio data were collected using a Zoom H6 digital recorder and Countryman H6 headset microphone with a hypercardioid pattern directional capsule covered with a windshield. Recordings were made at 24 bit bit-depth and a 48 kHz sample rate. Recording sessions primarily took place sitting outside, either in the open or on a veranda.

2.4. Data processing and acoustic measures

Audio files were transcribed and prepared in Praat [25], [26], and utterances were forced aligned using WebMAUS [27]. An EMU-SDMS database was created [28], and segmentation was manually corrected. Additional tiers included utterance, target word, orthographic word, and a tone tier for prosodic annotation. Data were labelled prosodically based on visual inspection of the f0 contour and auditory impressions following [21]. F0 (Hz) and RMS amplitude (dB) contours were extracted from the EMU database for the target words, and the f0 and RMS amplitude peaks associated with the accented syllable were calculated.

2.5. Analysis

Data were statistically analysed with generalised linear mixed models (henceforth GLMM) and linear mixed effects models (henceforth LMM) in R using the *lme4* [29] and *emmeans* [30] packages. Fixed effects were WORD ORDER (adjective-noun vs. noun-adjective) or POSITION ORDER (first position vs. second position), FOCUS CONDITION (AF- vs. NF- vs. NPF-condition), and WORD TYPE (adjective vs. noun). Random intercepts were SPEAKER and WORD IDENTITY. Details of models are provided in the results subsections where relevant.

3. RESULTS

A total of 654 utterances were collected: the intended 624 (4 tiles × 4 trials × 3 focus conditions × 13 speakers) plus 30 repetitions, predominantly prompted by production errors. In this study, contiguous NPs are considered. Therefore, utterances in which the target words were not adjacent (126), or where only one target word was produced (88) were excluded from the analysis. A further 13 utterances were excluded as they were repetitions, and nine due to production errors, resulting in 418 utterances for analysis; 836 target words. The distribution of utterances by focus condition was AF = 163, NF = 133, NPF = 122.

3.1. Word order

The relationship between word order and focus condition was tested with a GLMM. Word order was predicted, with a fixed effect of focus condition, and a random intercept for speaker.

Focus condition was a significant predictor of the order of the noun and adjective within NPs ($\chi^2(2) = 73.47$, $p < 0.0001$). Word order patterns were similar for NPF- and NF-conditions, with both orders being almost equally used—adjective-noun order was used in 56% of NF-condition and 54% of NPF-condition utterances. In the AF-condition, the adjective-noun order was far preferred, being used in 85% of utterances (AF ~ NF $p < 0.0001$, AF ~ NPF $p < 0.0001$, NF ~ NPF n.s.).

3.2. Accent distribution and pitch accent type

Almost all target words were accented (n = 828). Accent type was predicted in an LMM with fixed effects of focus condition, position order, and word type, with a three-way interaction, and a random intercept for speaker. No effect was significant in predicting accent type.

While accent distribution could not be analysed statistically as deaccenting was used infrequently (n = 8), deaccenting was employed as a strategy to mark given information by one speaker. This speaker deaccented nouns in the AF-condition in five out of 16 utterances, and strongly preferred the adjective-noun order in all focus conditions.

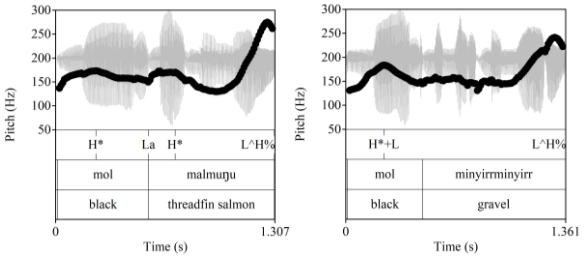

Figure 3: Right: an utterance where the adjective and noun are accented (NPF-condition); Left: a rare example of deaccenting (AF-condition)—the noun *minyirrminyirr* does not have an associated pitch accent.

3.3. F0

F0 peak was predicted in an LMM with fixed effects of focus condition, position order, and word type, with a three-way interaction, and random intercepts for speaker and word identity. The interaction

between focus condition, position order, and word type was significant in predicting f0 peak height ($\chi^2(3) = 11.2$, $p = 0.011$). However, there was only one significant comparison: adjectives in the NPF-condition had a significantly lower f0 peak when they occurred in the second position compared with the first ($\beta = -22.47$, $p < 0.0001$).

While other differences where not significant, there were trends (see raw values in Fig. 4). For example, in the AF-condition, adjectives in either first or second position had higher f0 peak values than nouns.

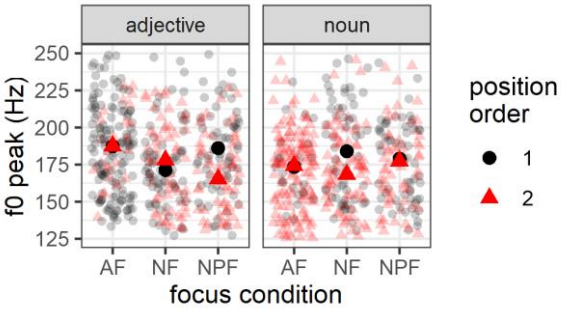

Figure 4: F0 peak values (Hz) and means for adjectives and nouns across focus conditions, grouped by whether they were the first (1) or second (2) word in the NP.

3.4. RMS amplitude

RMS amplitude peak was predicted in an LMM with fixed effects of focus condition, position order, and word type, with a three-way interaction, and random intercepts for speaker and word identity. The interaction between focus condition, position order, and word type was significant ($\chi^2(3) = 31.6$, $p < 0.0001$).

A number of comparisons were significant (see z-score values in Fig. 5; RMS amplitude values were z-scored by speaker for plotting due to strong systematic speaker differences that interacted with position order). Adjectives in second position had significantly lower RMS amplitude peak values when in NF- than AF-conditions ($\beta = -3.33$, $p < 0.001$). Nouns in second position had higher RMS amplitude peak values when in NF- than AF-condition ($\beta = 1.6$, $p < 0.05$). In second position of AF-condition utterances, nouns had significantly lower RMS amplitude peak values than adjectives in that position ($\beta = -3.54$, $p < 0.05$). Adjectives in AF-condition had significantly higher RMS amplitude peaks in second than first position ($\beta = 3.22$, $p < 0.001$).

4. DISCUSSION AND CONCLUSIONS

The aim of this paper was to describe how focus is realised prosodically in Djambarrpuyŋu noun phrases by both phonological and phonetic means, and what

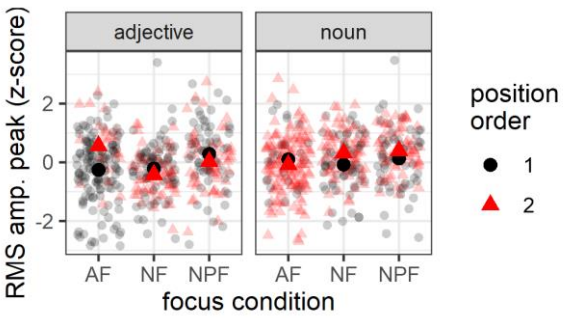

Figure 5: RMS amplitude peak values (z-score) and means for adjectives and nouns across focus conditions, grouped by whether they were the first (1) or second (2) word in the NP.

effect focus condition may have on word order. Specifically, accentuation and accent type, peak f0, and peak RMS amplitude were examined.

Word order significantly varied due to focus condition, with adjective-noun order being more common in the AF-condition than the NF- or NPF-conditions. Neither accentuation nor accent type varied consistently with focus condition. Phonetic prosodic cues did not vary consistently to distinguish between the different focus conditions in these data, and it is not clear whether all the patterns that were observed served to increase the prominence of the focused element c.f. [9]. This aligns Djambarrpuyŋu with languages like K'iche', of Guatemala, for which traditionally investigated phonological and phonetic means of encoding prosodic prominence are not found to reflect focus condition in NPs [9] (see also [31]). However, the current figures and statistical results do suggest that adjectives and nouns are treated differently from one another across focus conditions, paralleling results in [9].

Word order in this study was narrowly considered, only as the order of the target words in contiguous NPs, but discontinuous NPs may shed light on the diversity of ways focus within NPs is encoded in Djambarrpuyŋu. Further, it may be that other methods for encoding givenness assist in distinguishing the focus conditions, such as deletion, and, as mentioned above, deaccenting which while not regularly used, may be employed by some speakers in cases of continuity of word position in combination with givenness, as has been described for English [22].

Investigations are currently underway into discontinuous noun phrases, as well as additional acoustic measures such as duration, dynamic measures of pitch, and other aspects of prosody such as phrasing. However, the current results do not suggest phonological nor phonetic correlates of prominence consistently contribute to distinguishing words' information status within Djambarrpuyŋu NPs in these three focus conditions.

5. ACKNOWLEDGEMENTS

Many thanks to the Yolŋu who participated in this study and in earlier work. Thank you to Elizabeth Räkay Milmilany, Johanna Schirmer, and Rasmus Puggaard-Rode for assistance and advice, and colleagues at the IPS and a reviewer for their comments. This work was supported by the Alexander von Humboldt Foundation, the University of Melbourne and the ARC Centre of Excellence for the Dynamics of Language (CE140100041).

6. REFERENCES

[1] Arvaniti, A. 2022. The autosegmental-metrical model of intonational phonology. In: Barnes, J., Shattuck-Hufnagel, S. (eds), *Prosodic Theory and Practice*. MIT Press, 25–83.

[2] Kügler, F., Calhoun, S. 2020. Prosodic encoding of information structure: A typological perspective. In: Gussenhoven C., Chen, A. (eds), *The Oxford Handbook of Language Prosody*. OUP, 453–467.

[3] Krifka, M. 2008. Basic notions of information structure, *Acta Linguist. Hung.*, 55(3-4), 243–276.

[4] Baumann, S., Grice, M. 2006. The intonation of accessibility, *J. Pragmat.* 38(10), 1636–1657.

[5] Chodroff, E., Cole, J. 2019. The phonological and phonetic encoding of information structure in American English nuclear accents. *Proc. 19th ICPhS* Melbourne, 1570-1574.

[6] Chahal, D., Hellmuth, S. 2014.The intonation of Lebanese and Egyptian Arabic. In: Jun, S.-A. (ed), *Prosodic Typology*, vol. II. OUP, 365–404.

[7] Swerts, M., Krahmer, E., Avesani, C. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *J. Phon.* 30(4), 629–654.

[8] Ladd, D. R. 2008. *Intonational Phonology*, 2nd edn. CUP.

[9] Burdin, R. S., Phillips-Bourass, S., Turnbull, R., Yasavul, M., Clopper, C. G., Tonhauser, J. 2015. Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua* 165, 254–276.

[10] Ito K., Speer, S. R. 2006. Using interactive tasks to elicit natural dialogue. In: Sudhoff, S. et al. (eds), *Methods in Empirical Prosody Research*. De Gruyter.

[11] Swerts, M., Taniguchi, M., Katagiri, Y. 2000. Prosodic marking of information status in Tokyo Japanese. *6th ICSLP* Beijing 1, 78–81.

[12] Langer, C., Kügler, F. 2021. Focus and prosodic cues in Hungarian noun phrases. *1st TAI* Sonderborg, 219–223.

[13] Wilkinson, M. 2012. *Djambarrpuyŋu: A Yolŋu Variety of Northern Australia*. LINCOM Europa.

[14] Nordlinger, R. 2014. Constituency and grammatical relations in Australian languages. In: Koch, H., Nordlinger, R (eds), *The Languages and Linguistics of Australia*. Mouton de Gruyter, 215–261.

[15] Louagie, D., Verstraete, J.-C. 2016. Noun phrase constituency in Australian languages: A typological study. *Linguist. Typology* 20(1), 25–80.

[16] Simpson, J., Mushin, I. 2008. Clause-initial position in four Australian languages. In: Mushin, I., Baker, B. (eds), *Discourse and Grammar in Australian Languages*. John Benjamins, 25–57.

[17] Fletcher J., Butcher, A. 2014. Sound patterns of Australian languages. In: Koch, H., Nordlinger, R (eds), *The Languages and Linguistics of Australia.* Mouton de Gruyter, 91–138.

[18] Bishop, J., Fletcher, J. 2005. Intonation in six dialects of Bininj Gun-wok. In: Jun, S.-A. (ed), *Prosodic Typology*. OUP, 331–361.

[19] Fletcher, J. 2014. Intonation and prosody in Dalabon. In: Jun, S.-A. (ed), *Prosodic Typology*, vol. II. OUP, 252–272.

[20] Jun, S-A. 2014. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In: Jun, S.-A. (ed), *Prosodic Typology*, vol. II. OUP, 520–539.

[21] Jepson, K., Fletcher, J. Under review. The intonation of Djambarrpuyŋu.

[22] Terken, J., Hirschberg, J. 1994. Deaccentuation of words representing "given" information: Effects of persistence of grammatical function and surface position. *Lang. Speech* 37(2), 125–145.

[23] Röhr, C. T., Baumann, S., Grice, M. 2015. The effect of verbs on the prosodic marking of information status: Production and perception in German. *Proc. 18th ICPhS* Glasgow.

[24] Fletcher, J., Stoakes, H., Singer, R., Loakes, D. 2016. Intonational correlates of Subject and Object realisation in Mawng (Australian). *Speech Prosody* Boston, 188–192.

[25] Boersma P., Weenink, D. 2022. Praat: Doing phonetics by computer. (v. 6.2.23)

[26] Zhang, C. 2021. Praat (GitHub repository). https://github.com/congzhang365/Praat

[27] Kisler, T., Reichel, U. D., Schiel, F. 2017. Multilingual processing of speech via web services. *Comput. Speech Lang.* 45, 326–347.

[28] Winkelmann, R., Harrington, J., Jänsch, K. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Comput. Speech Lang.* 45, 392–410. (v. 2.3.0)

[29] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, 67(1), 1–48. (v. 1.1-31)

[30] Lenth, R. V. 2022. emmeans: Estimated Marginal Means, aka Least-Squares Means. (v. 1.8.2)

[31] Kaland, C., Swerts, M., Himmelmann, N., Red and blue bananas: Time-series f0 analysis of contrastively focused noun phrases in Papuan Malay and Dutch. *J. Phon.* 96, 101200.

[1] Discontinuous NPs (in which semantically-related nominal elements are distributed throughout the clause) are also permitted in Djambarrpuyŋu though not explored here.