

The Prosodic Contours of Jaminjung,
a Language of Northern Australia

A thesis submitted to the University of Manchester
for the degree of doctor of Philosophy
in the Faculty of Humanities
2010

by
Candide Simard
School of Languages, Linguistics and Cultures

ABSTRACT	20
DECLARATION	21
COPYRIGHT STATEMENT	22
ACKNOWLEDGMENTS	23
ABBREVIATIONS IN INTERLINEAR GLOSSES	27
CHAPTER 1 INTRODUCTION	29
1.1 Jaminjung: speakers	31
1.2 Genetic affiliation	32
1.3 Previous studies	33
1.4 Fieldwork setting and data collection	34
CHAPTER 2 PROSODIC RESEARCH	36
2.1 Basic notions in prosodic research	36
2.1.1 Prosodic phenomena	37
2.1.2 Intonational features	42
2.2 Theoretical Models	49
2.2.1 Phonetics or phonology	49
2.2.2 Characterizing phonological models	50
2.2.3 Overview of the Models	53
2.2.3.1 The American Structuralists	53
2.2.3.2 The British school	53
2.2.3.3 Brazil and discourse intonation	55
2.2.3.4 The Dutch school	56
2.2.3.5 Intonational phonology	56
2.2.3.6 Prosodic phonology and the Prosodic Hierarchy	61
2.2.3.7 INTSINT	64
2.2.3.8 Tilt	65

2.2.3.9 KIM	66
2.2.3.10 Fujisaki	66
2.3 The PENTA model	68
2.3.1 Parallel encoding	69
2.3.2 Pitch target approximation	70
2.3.3 Evaluating pitch targets	71
2.4 Methodological considerations	77
2.5 Conclusion	80
 CHAPTER 3 GROUPING	 81
3.1 Prosodic units	82
3.1.1 Intonation Units	83
3.1.2 Units smaller than the IU	85
3.1.3 Units larger than the IU	87
3.1.4 Prosodic cues associated with the boundaries of prosodic units	88
3.1.5 Prosodic units in other Australian languages	91
 3.2 Outline of the grammar of Jaminjung, its (potential) prosodic units and their morpho-syntactic correspondence	 94
3.2.1 Introduction to the grammar of Jaminjung	94
3.2.2 Phonology and syllable structure in Jaminjung	95
3.2.2.1 The phonemic inventory of Jaminjung	95
3.2.2.1.1 Vowels	95
3.2.2.1.2 Consonants	97
3.2.2.2 Phonotactics	99
3.2.2.2.1 Initial position	99
3.2.2.2.2 Final position	102
3.2.2.2.3 Consonant clusters	104
3.2.2.2.4 Reduplication	106
3.2.2.3 Word-level phonological processes	109
3.2.2.3.1 Deletion	109
3.2.2.3.2 Lenition	109
3.2.2.3.3 Fortition	111
3.2.2.4 Stress	111
3.2.2.5 Summary	113
3.2.3 Prosodic units in Jaminjung	114
3.2.3.1 Prosodic words	114
3.2.3.1.1 Nominals	114

3.2.3.1.2 Case markers	116
3.2.3.1.3 Verb forms and bound pronominals	117
3.2.3.1.3.1 Bound pronominals	119
3.2.3.1.4 Coverbs	119
3.2.3.1.5 Particles	120
3.2.3.1.6 Clitics	121
3.2.3.1.7 Summary prosodic words	123
3.2.3.2 Phrasal constituents	124
3.2.3.2.1 Noun phrases	125
3.2.3.2.2 Complex Predicates	126
3.2.3.2.3 Non-finite Predicates	127
3.2.3.2.4 Particle group	127
3.2.3.3 Intonation units and prosodic sentences	128
3.2.3.3.1 Intonation units	129
3.2.3.3.2 Prosodic sentences	129
3.2.3.4 An inventory of clause types in Jaminjung	130
3.2.3.4.1 Independent clauses: Verbal clauses	130
3.2.3.4.2 Independent clauses: Verbless clauses	131
3.2.3.4.3 Finite subordinate clauses	134
3.2.3.4.4 Non-finite subordinate clauses	134
3.2.3.4.5 Non-clausal IUs	136
3.2.3.5 Summary	139
3.3 Analysis: prosodic constituents in Jaminjung	140
3.3.1 Methodology	141
3.3.2 Dataset	142
3.3.3 Right-edge boundaries	145
3.3.3.1 Duration	146
3.3.3.1.1 Absolute duration	146
3.3.3.1.2 Relative duration	147
3.3.3.2 Pauses	148
3.3.3.3 Affinity index	149
3.3.3.3.1 Coverbs as prosodic words	150
3.3.3.4 Pitch and phonation events	152
3.3.3.4.1 Lowering of F0	152
3.3.3.4.2 Final Velocity	155
3.3.3.4.3 Phonation events	156
3.3.3.5 Summary: cues at right-edge boundaries	158
3.3.4 Left-edge boundaries	159
3.3.4.1 Pitch reset	159
3.3.4.2 Initial high tone	161

4.3.4 Comparison: pitch target, final velocity and mean F0	241
4.3.5 Comparison: pitch excursion	243
4.3.6 Comparison: duration	244
4.4 Conclusion and Discussion: focus domains	246
4.5 Topics	249
4.5.1 Dataset	250
4.5.2 Hypotheses	258
4.5.3 Methodology	258
4.5.4 Given topics	259
4.5.4.1 Given topics at the left periphery	259
4.5.4.2 Comparison: subtypes given topics	265
4.5.4.3 Summary: given topics	269
4.5.5 New topics	269
4.5.5.1 Thetic and/or new topic	273
4.5.6 Comparison: topics	276
4.6 Contrast	281
4.6.1 Contrastive argument focus	282
4.6.1.1 Dataset	282
4.6.1.2 Comparison between contrastive and non-contrastive focused arguments	283
4.6.2 Contrastive topics	286
4.6.3 Contrastive clitic = <i>gun</i>	290
4.6.4 Comparing contrastive constructions	293
4.7 Conclusion: marking information structure	295
CHAPTER 5 DISTINGUISHING SENTENCE TYPES	297
5.1 Methodology	300
5.2 Declaratives	301
5.2.1 Hypotheses	301
5.2.2 Simple sentences (1 IU)	302
5.2.2.1 Dataset	303
5.2.2.2 Average pitch, final velocity and pitch targets	304
5.2.3 Complex sentences (2 IUs)	309
5.2.3.1 Dataset	310
5.2.3.2 First IUs	319
5.2.3.2.1 Summary: First IUs	325
5.2.3.3 Second IUs: Verbal clauses	326

5.2.3.3.1 Summary: 2 verbal clauses	330
5.2.3.4 Comparison of second IUs, verbal, non-verbal and NPs	331
5.2.3.4.1 Mean F0	332
5.2.3.4.2 Pitch excursion	334
5.2.3.4.3 Final velocity	336
5.2.3.4.4 Duration	339
5.2.3.4.5 Comparing second IUs that follow lengthened first IUs to other second IUs in complex sentences	341
5.2.4 Discussion: complex declarative sentences and iconic lengthening	344
5.3 Interrogatives	353
5.3.1 Hypotheses	353
5.3.2 Dataset	354
5.3.3 Information and polar questions	358
5.3.3.1 The falling contour	361
5.3.3.2 The rising contour	363
5.3.3.3 The fall-rise contours	365
5.3.4 Tag questions	367
5.3.5 Summary: interrogatives	368
5.4 Imperatives	369
5.4.1 Hypothesis and dataset	369
5.4.2 The correlates of imperative sentences	370
5.4.3 Summary: imperatives	373
5.5 Comparison between declaratives, interrogatives and imperatives	373
5.5.1 Conclusion and discussion: comparison of sentence types	375
5.6 Vocatives	378
5.6.1 Dataset	379
5.6.2 The correlates of vocatives	380
5.6.2.1 Comparing the vocatives to the surrounding IUs	381
5.7 Reported or ‘quotation’ speech	384
5.7.1 Dataset	385
5.7.2 The correlates of the quotative expressions	386
5.8 Conclusion: sentence types	392
 CHAPTER 6 CONCLUSION: THE PROSODIC CONTOURS OF JAMINJUNG	 399

REFERENCES

409

Word count: 88926

List of figures

<i>Figure 1-1 Map of part of the Northern Territory, Australia.</i>	32
<i>Figure 2-1 Spectrograms of the word voice produced with different phonations. From top to bottom, it is produced with modal voice, breathy voice, and creaky voice, respectively (adapted from Yoon et al. 2008:3).</i>	40
<i>Figure 2-2 A pitch track showing how the F0 is disturbed by consonants with an extreme constriction in the vocal tract (from Ladd 2008).</i>	42
<i>Figure 2-3 Pitch track.</i>	43
<i>Figure 2-4 An illustration of declination and downstep (adapted from Liu and Xu 2005).</i>	46
<i>Figure 2-5 Functions of intonation and their intonational realization. On the left side is a scale from the paralinguistic to the linguistic, on the right is a scale from the gradient to the categorical. In the middle are the functions of intonation (from Grice and Bauman 2007:14).</i>	52
<i>Figure 2-6 Halliday's tones (1967:16).</i>	54
<i>Figure 2-7 The grammar of English intonation patterns (Pierrehumbert 1980, figure from Ladd 1996:80) in which (A) Prosodic phrase boundaries coincide with boundaries of intonational tunes, and (B) Every intermediate phrase contains at least one accent.</i>	58
<i>Figure 2-8 Tone inventory for English, with the ToBI system of annotation inventory (adapted from Clark, et al. 2007:364-370).</i>	59
<i>Figure 2-9 An example of the alignment of the pitch accents and boundary tones in English, annotated with the ToBI system. The upper part shows the speech signal, the middle part the annotation tiers (tone, syllable and break) and the lower part of the figure shows the associated pitch track (from Beckman and Ayers, 1997).</i>	60
<i>Figure 2-10 The hierarchy of units proposed by Selkirk (1986) in the Prosodic Phonology theory.</i>	62
<i>Figure 2-11 Ladd's examples of coordinated recursive structures (adapted from Ladd 1996:242).</i>	62
<i>Figure 2-12 INTSINT annotation with an English sentence (from Hirst 2005:243).</i>	65
<i>Figure 2-13 Intonation as articulatorily implemented encoding schemes defined and organized by communicative functions (from Xu 2005:243).</i>	68
<i>Figure 2-14 Possible symbolic values of the melodic primitives pitch target, pitch range, strength and duration (adapted from Xu 2005:243).</i>	69
<i>Figure 2-15 Decomposition of the F0 contours of "Lee may mimic my niece" according to the PENTA model (from Xu and Xu 2005).</i>	70
<i>Figure 2-16 Illustration of the TA model (from Xu 2007:2).</i>	71
<i>Figure 2-17 An illustration of a derivative. The function is drawn in black, and a tangent line to that function is drawn in red (Wikipedia, 'derivative', accessed 01/03/2010).</i>	72
<i>Figure 2-18 The profiles of Mandarin tones, obtained in a simulation (adapted from Gauthier et al. 2006:98). The black vertical line shows the point that is measured in the Jaminjung dataset.</i>	73

<i>Figure 2-19</i> Stylised representation of the pitch targets. A [high], B [rise], C [fall], D [low]. The thinner line represents the pitch, the thicker line represents the velocity. The dots on the pitch line indicate the highest and lowest F0 points during the syllable.	75
<i>Figure 2-20</i> Examples of velocity curves for the targets in Jaminjung, over one syllable; and the pitch track of the IU from which they are extracted.	76
<i>Figure 2-21</i> The relation between Hertz and semitones in the range 50 to 500 Hz (from Nolan 2003:771).	79
<i>Figure 3-1</i> The vowels of Jaminjung and the rates of their occurrences (in brackets) in the lexical entries of the database - as of 2009: the database is continuously updated.	95
<i>Figure 3-2</i> The rates of occurrences of segments in initial position in nominals and others word classes (interjections, particles, etc), coverbs and inflected verb forms expressed as percentages.	101
<i>Figure 3-3</i> The rates of occurrences of segments in final position in nominals and others word classes (interjections, particles, etc.), coverbs and inflected verbs expressed as percentages.	103
<i>Figure 3-4</i> An example with 2 NPs in apposition, the first formed of the Kriol demonstrative with a proper noun, and the second of a possessive pronoun, an adjectival nominal, and a common noun. The speech signal and pitch track – speckled line – are also displayed.	116
<i>Figure 3-5</i> The adverbial nominal balarrgu.	116
<i>Figure 3-6</i> The noun jarriny followed by the locative case marker –gi.	117
<i>Figure 3-7</i> An example sentence containing the intransitive verb form bunyjuyu.	118
<i>Figure 3-8</i> The particle bugu.	121
<i>Figure 3-9</i> The clitic =biyang attaching to a finite verb, a nominal, and a coverb.	123
<i>Figure 3-10</i> A sentence segmented into prosodic words.	124
<i>Figure 3-11</i> The first constituent is a complex predicate, formed by a coverb and an inflected verb.	126
<i>Figure 3-12</i> The non-finite predicate mubayib at the end of the sentence is labeled as a phrasal constituent.	127
<i>Figure 3-13</i> The group formed by the particle and a noun phrase is labeled ‘particle group’.	128
<i>Figure 3-14</i> An intonation unit corresponding to a verbal clause; the Agent/subject is omitted.	131
<i>Figure 3-15</i> An equative clause, with 2 NPs.	132
<i>Figure 3-16</i> An ascriptive clause.	133
<i>Figure 3-17</i> An existential clause.	133
<i>Figure 3-18</i> A prosodic sentence consisting of 2 IUs: the first is a main clause, with a verbal predicate; the second is a finite subordinate clause which functions as a relative clause.	134
<i>Figure 3-19</i> A prosodic sentence with 2 clauses: the first is the main clause; the second is a non-finite subordinate clause, formed by a coverb and the dative suffix.	135

Figure 3-20 The second IU in this prosodic sentence is formed of an NP.	136
Figure 3-21 The interjection awu, at the beginning of a sentence.	137
Figure 3-22 A prosodic sentence containing two IUs, describing a succession of events in a narrative.	138
Figure 3-23 A prosodic sentence consisting of two IUs, the first corresponding to a clause containing a reporting verb followed by a second IU, which is a sequence of direct speech as complement to the first clause.	139
Figure 3-24 The annotation of a textgrid. The first tier is the syllable parsing, the boundaries of the other prosodic units are shown in the subsequent tiers.	141
Figure 3-25 The boxplot on the left show the average ratios per type of unit. The table on the right shows the actual values of the ratios.	148
Figure 3-26 Duration of pauses after each potential unit (ms).	149
Figure 3-27 Affinity index, a measure of the inter-onset interval for each potential unit.	150
Figure 3-28 The lowering of F0 shown for each type of unit.	155
Figure 3-29 Spectrogram showing breathiness and creakiness.	157
Figure 3-30 The onset of prosodic words, shown by the arrows, where the pitch reset is minimal, but still breaks the line of declination.	161
Figure 3-31 The hypothesised high tones in an utterance in Jaminjung. The pitch events under investigation are indicated by arrows.	162
Figure 3-32 Final velocity in the first syllables of each unit.	163
Figure 3-33 Mean F0 in the first 2 syllables of the units.	164
Figure 3-34 Excursion size, duration and intensity measurements of the first syllables of the units.	166
Figure 4-1 Example of IU consisting of a topic NP and a comment NP.	195
Figure 4-2 Example of IU consisting of an intransitive clause.	196
Figure 4-3 An IU consisting of a transitive clause: NP(top) + coverb V NP(comment).	196
Figure 4-4 Illustration of topic-predicate construction with a prominence on the first syllable of the predicate. This syllable corresponds to a coverb (the arrow points to it).	200
Figure 4-5 Examples with words of 1-, 2- and 3- syllables in comment-initial position which have a prominent first syllable, shown by an arrow.	202
Figure 4-6 The sentence contains a monosyllabic coverb, with a [fall] pitch target, followed by [mid] targets on the inflected verb.	204
Figure 4-7 A bisyllabic coverb bunburr with a pattern of a [fall] and [mid] targets on the first and second syllables.	205
Figure 4-8 A bisyllabic nominal juwud with a [high] on the first syllable, followed by a [fall].	206
Figure 4-9 A trisyllabic coverb yarraju displaying the pattern of pitch target [high] [fall] [mid].	207
Figure 4-10 A trisyllabic noun, mulurru, also displaying a pattern of [high] [fall] [mid].	207

Figure 4-11 Final velocities by word length.	209
Figure 4-12 The topic in the example is the noun <i>janyung</i> , and the comment is <i>buru gajgany</i> . The main prominence is heard on the first syllable of the comment, here the coverb <i>buru</i> . The velocity measurements for the initial syllables of <i>ja-</i> and <i>bu-</i> are also shown.	211
Figure 4-13 Mean F0 values in the monosyllabic coverbs, also showing the first syllables of the following word (top graph), in bisyllabic coverbs, nouns and inflected verbs (middle graph), and in tri-syllabic coverbs and nouns (bottom graph).	213
Figure 4-14 Mean F0 values for the syllables before and after the prominent syllables (<i>MOST_PROM</i> , boxed). Constituents are separated by a vertical line.....	214
Figure 4-15 Excursion size in the syllables in the words that precede the main prominence (left); and in the syllables that succeed the prominences, in the same word (labeled 2nd, 3rd, 4th, etc) and in the 2nd word, when they occur (<i>1st_postPROM</i> , etc). The most prominent syllable is boxed.....	215
Figure 4-16 Duration of the syllables in the first word of the comment, distinguishing syntactic categories and number of syllables. The prominent syllable is the first. Only non-IU final words are considered.....	219
Figure 4-17 Sentence focus construction in annuntiative function; all the constituents are new.	227
Figure 4-18 A thetic sentence where the speaker interrupts the telling of a story to comment on the recording equipment, thus used in its interruptive function.	227
Figure 4-19 A comparison of the words with perceived primary and secondary prominences.	230
Figure 4-20 A comparison of the final velocity values in the syllables of the constituents containing primary and secondary prominences.	230
Figure 4-21 A comparison of the excursion sizes in the syllables of the constituents containing primary and secondary prominences	231
Figure 4-22 A comparison of the duration values in the syllables of the constituents containing primary and secondary prominences.	231
Figure 4-23 Pitch traces and velocity measurements illustrating patterns of prominences in thetic sentences. The top pane shows a main prominence heard on the first syllable of the second word, <i>gardam</i> ; the bottom pane shows a prominence on the first syllable of the first word, <i>gugu</i>	234
Figure 4-24 A question-answer pair, illustrating a focused argument in the answer, <i>nguyung bunyag=ma</i>	235
Figure 4-25 Example sentences illustrating different types of argument focus constructions. In the top pane (64), the comment repeats the information in the previous IU, the only new information is the focused argument, <i>Barndi</i> . In the second pane (65), the argument in focus is contained in the answer, <i>waya</i>	238

<i>Figure 4-26 Final velocity measurements in the tree focus domains, according to syllable position.....</i>	<i>242</i>
<i>Figure 4-27 Mean F0 values of the syllables in the three focus domains. The words are separated by vertical lines.....</i>	<i>243</i>
<i>Figure 4-28 Mean excursion size values of the syllables in the three focus domains.</i>	<i>244</i>
<i>Figure 4-29 Durations of syllables in the prosodic words for each focus domain.....</i>	<i>245</i>
<i>Figure 4-30 The topic in this sentence is uttered in the first IU, wirib. It has been referred to many times in the course of the storytelling, and is thus a given topic.</i>	<i>253</i>
<i>Figure 4-31 A given topic at the left of the IU, this time occurring with the clitic =marlang.</i>	<i>254</i>
<i>Figure 4-32 A given topic occurring at the right edge of the IU, wirib-di.</i>	<i>254</i>
<i>Figure 4-33 The framesetting topic, yinju=biya is at the left edge of the sentence, and the given topic jarlig n-wirib is found at the right edge, after a pause.</i>	<i>255</i>
<i>Figure 4-34 A right dislocated detached topic with the clitic =marlang.</i>	<i>256</i>
<i>Figure 4-35 A new topic, buyud=biyang, at the left of the prosodic sentence.</i>	<i>257</i>
<i>Figure 4-36 A new topic, warladbari, separated by a pause from the rest of the prosodic sentence.</i>	<i>257</i>
<i>Figure 4-37 Average final velocity in the syllables of given topic/comment constructions.</i>	<i>260</i>
<i>Figure 4-38 Graph showing the average pitch values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang (dotted line), and those that are not (full line).</i>	<i>261</i>
<i>Figure 4-39 Graph showing the average pitch excursions in the syllables of given topic/comment constructions.</i>	<i>263</i>
<i>Figure 4-40 Average duration in the syllables of given topic/comment constructions.</i>	<i>264</i>
<i>Figure 4-41 A comparison of final velocity average values in all subtypes of given topics.</i>	<i>266</i>
<i>Figure 4-42 A comparison of mean F0 values between all subtypes of given topics.</i>	<i>267</i>
<i>Figure 4-43 A comparison of the excursion sizes in the syllables in all subtypes of given topics.</i>	<i>268</i>
<i>Figure 4-44 A comparison of the duration values in all subtypes of given topics.</i>	<i>269</i>
<i>Figure 4-45 Comparative values of final velocity in new topics, whether they are integrated in a clause or separated by a pause.</i>	<i>271</i>
<i>Figure 4-46 Comparative values of mean F0 in new topics, whether they are integrated in a clause or separated by a pause.</i>	<i>271</i>
<i>Figure 4-47 Comparative values of excursion size in new topics, whether they are integrated in a clause or separated by a pause in detached and non-detached new topics.</i>	<i>272</i>
<i>Figure 4-48 Comparative values of duration in new topics, whether they are integrated in a clause or separated by a pause in detached and non-detached new topics.</i>	<i>273</i>

Figure 4-49 A comparison between thetic sentence and new dislocated topics, including correlates of final velocity, mean F0, excursion size, and duration.....	275
Figure 4-50 A comparison of the final velocity measurements in all subtypes of topics....	277
Figure 4-51 A comparison of the mean F0 measurements in all subtypes of topics.	278
Figure 4-52 A comparison of the excursion size measurements in all subtypes of topics..	278
Figure 4-53 A comparison of the duration measurements in all subtypes of topics.....	279
Figure 4-54 The givenness scale in Jaminjung. On the left, the topics subtypes that are 'more given' and on the right those that are 'less given'.....	281
Figure 4-55 A contrastive argument focus is shown, julag=biji.....	282
Figure 4-56 A contrastive argument focus is shown, julag=biji. The bottom pane shows a corrective focus, the speaker providing a Jaminjung equivalent, langiny-bina, to a word first uttered in Kriol, log-bina.	283
Figure 4-57 The mean values of the measurements of the correlates of final velocity, F0, excursion size, and duration, distinguishing between contrastive and non-contrastive argument focus.	285
Figure 4-58 The contrastive topics jarriny-gi and balarrgu=mang.	287
Figure 4-59 The values of final velocity, mean F0, excursion size, and duration of contrastive topics and the following comments, compared to those of the given topic-comment constructions.	289
Figure 4-60 A sentence showing a contrastive argument followed by the clitic wirlga=gun.	291
Figure 4-61 Graphs of the values of mean F0, excursion size, final velocity and duration of the contrastive arguments in focus and topics which host the clitic =gun.	292
Figure 4-62 A comparison of the contrastive constructions, including the correlates of final velocity, mean F0, excursion size, and duration.	294
Figure 5-1 Simple declaratives, the first example is an IU with an explicit topic; the second consists of a topic and a nominal predicate.	304
Figure 5-2 Average mean F0 values of the syllables of simple declaratives.	305
Figure 5-3 Final velocity values in the IU-final words of varying length.	308
Figure 5-4 Sentences consisting of two clauses. The top pane shows a first IU with a falling contour; the middle panes show the 'fall-rise' contour, the first with the rise coinciding with the clitic =na 'now', and the second with a rise from the first syllable of the final word. The bottom pane illustrates a level contour.	313
Figure 5-5 A prosodic sentence made of a main verbal clause, followed by a coverb waga 'sit' as a non-finite clause.	314
Figure 5-6 A prosodic sentence with an NP functioning as an afterthought in the second IU.	315
Figure 5-7 A prosodic sentence with an NP functioning as a reactivated topic in the 2 nd IU.	316

<i>Figure 5-8 Iconic lengthening in the first IU followed, in the top pane, by an IU consisting of a non-finite predicate; in the middle pane by an NP; and in the bottom pane by a verbal clause.</i>	<i>319</i>
<i>Figure 5-9 The values of mean F0 in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.</i>	<i>321</i>
<i>Figure 5-10 The values of excursion size in the first IUS with a fall, fall-rise, level and lengthened contours. Only the first and last three syllables of the IUs are shown.</i>	<i>322</i>
<i>Figure 5-11 The values of final velocity in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.</i>	<i>323</i>
<i>Figure 5-12 The values of duration in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.</i>	<i>324</i>
<i>Figure 5-13 The mean F0 values in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.</i>	<i>327</i>
<i>Figure 5-14 The values of excursion size in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.</i>	<i>328</i>
<i>Figure 5-15 The values of final velocity in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.</i>	<i>329</i>
<i>Figure 5-16 The values of duration in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.</i>	<i>329</i>
<i>Figure 5-17 A sentence in which the second IU starts with the particle majani. The falling contour effectively starts at the beginning of the focused word burruriga.</i>	<i>331</i>
<i>Figure 5-18 The mean F0 values in 2-, and 4-syllables second IUs, as well as the average of the values for the syllables at the left and right periphery in all IUS.</i>	<i>334</i>
<i>Figure 5-19 A comparison of the excursion size of syllables at the boundaries of in the second IUs, shown for IUs of various length.</i>	<i>336</i>
<i>Figure 5-20 A comparison of the final velocity values in the second IUs, shown for IUs of various length.</i>	<i>338</i>
<i>Figure 5-21 The duration of the syllables in the second IUs, shown according to IU length. The bottom right graph shows the overall durations, including that of the preceding pause.</i>	<i>340</i>
<i>Figure 5-22 A comparison of the mean F0 values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.</i>	<i>342</i>
<i>Figure 5-23 A comparison of the excursion size values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.</i>	<i>343</i>
<i>Figure 5-24 A comparison of the final velocity values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.</i>	<i>343</i>
<i>Figure 5-25 A comparison of the duration values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.</i>	<i>344</i>
<i>Figure 5-26 A sentence showing two independent verbal clauses.</i>	<i>348</i>

<i>Figure 5-27 A prosodic sentence in which the second IU is NP, functioning as a reactivated topic.....</i>	<i>349</i>
<i>Figure 5-28 A prosodic sentence in which the second IU is an NP functioning as an afterthought.....</i>	<i>350</i>
<i>Figure 5-29 A prosodic sentence showing a non-verbal clause in the second IU, functioning as a non-finite predicate.....</i>	<i>351</i>
<i>Figure 5-30 Information questions, displaying a falling, a rising, and a fall-rise contour.....</i>	<i>355</i>
<i>Figure 5-31 Polar questions with a falling and a rising contour.....</i>	<i>356</i>
<i>Figure 5-32 Tag questions, the top pane an instance of ngih 'isn't it', and the bottom pane gurra 'right?'.....</i>	<i>358</i>
<i>Figure 5-33 The average values for the correlates of mean F0, pitch excursions, final velocity and duration in the penultimate and final syllables of the three observed contours.....</i>	<i>360</i>
<i>Figure 5-34 A comparison of the correlates of polar and information questions with a falling contour.....</i>	<i>363</i>
<i>Figure 5-35 The average of mean F0, excursion size, final velocity, and duration in the first, second, penultimate and final syllables of interrogatives with a rising contour, distinguishing between information and polar questions.....</i>	<i>364</i>
<i>Figure 5-36 The average of mean F0, excursion size, final velocity, and duration in the first, second, penultimate and final syllables of interrogatives with a fall-rise contour, distinguishing between information and polar questions.....</i>	<i>366</i>
<i>Figure 5-37 The mean F0 values in sentences ending with a tag. The first IUS have very similar default declarative patterns.</i>	<i>367</i>
<i>Figure 5-38 The average excursion size, final velocity, and duration in the syllables of the tags.....</i>	<i>368</i>
<i>Figure 5-39 An imperative sentence with a verb in the imperative form.</i>	<i>370</i>
<i>Figure 5-40 The mean F0 values of the syllables in imperative sentences. Words are separated by vertical lines.....</i>	<i>371</i>
<i>Figure 5-41 The average of mean F0, excursion size, final velocity, and duration in the 1st, second, penultimate and final syllables of imperative sentences.</i>	<i>372</i>
<i>Figure 5-42 A comparison of the mean F0 values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences, distinguished according to their contours.....</i>	<i>374</i>
<i>Figure 5-43 A comparison of the excursion size values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.....</i>	<i>374</i>
<i>Figure 5-44 A comparison of the final velocity values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.....</i>	<i>375</i>
<i>Figure 5-45 A comparison of the duration values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.....</i>	<i>375</i>

<i>Figure 5-46 A vocative in sentence-final position, used as an address.....</i>	<i>379</i>
<i>Figure 5-47 The mean pitch of the syllables in vocatives, showing a rise in values.</i>	<i>380</i>
<i>Figure 5-48 A comparison of the mean F0 values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).</i>	<i>382</i>
<i>Figure 5-49 A comparison of the excursion size values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).</i>	<i>382</i>
<i>Figure 5-50 A comparison of the final velocity values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).</i>	<i>383</i>
<i>Figure 5-51 A comparison of the duration values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).</i>	<i>384</i>
<i>Figure 5-52 A typical example of reported speech in the form of direct quotation. It is shown by a bracket, starting with maitbe and ending with ganinyba.</i>	<i>386</i>
<i>Figure 5-53 The values of mean F0 and excursion size in the IUs surrounding the quotative expressions. IUs are separated by a vertical line.....</i>	<i>388</i>
<i>Figure 5-54 The average values of mean F0, excursion size, final velocity and duration in the syllables of the IUs surrounding the quotation expressions.</i>	<i>391</i>

List of tables

<i>Table 3-1 The consonants of Jaminjung (from Schultze-Berndt 2000:41). The pointed brackets indicate the symbols used in the orthography to represent the sound. The numbers beneath represent the rates of occurrence of the phonemes in the lexical database of 2135 entries.....</i>	<i>97</i>
<i>Table 3-2 The rates of occurrences of consonant clusters in word final position.....</i>	<i>104</i>
<i>Table 3-3 The occurrence of open syllables in initial position, shown by the number of syllables in nominals, coverbs, the total for both, and the inflected verb forms.</i>	<i>105</i>
<i>Table 3-4 The occurrence of open syllables in final position, shown according to number of syllables in nominals, coverbs, the total for both, and the inflected verb forms.</i>	<i>106</i>
<i>Table 3-5 The occurrence of closed syllables in 2nd position of words of 3 syllables and more.</i>	<i>106</i>
<i>Table 3-6 Morpho-syntactic equivalence of the potential units.</i>	<i>140</i>
<i>Table 3-7 Summary of the potential units in each of the texts used in the dataset.</i>	<i>143</i>
<i>Table 3-8 Number of syllables in each of the potential units. N is the number of occurrences and % is the percentage of the total for this unit.</i>	<i>144</i>
<i>Table 3-9 Pitch range of the speakers in the corpus, expressed in Hz and in semitones. ..</i>	<i>145</i>
<i>Table 3-10 Mean duration of the last syllables of the potential units.</i>	<i>146</i>
<i>Table 3-11 Average duration/IOI of the last syllable of the prosodic words, phrasal constituents, IUs and prosodic sentences, with coverbs excluded from measurements in the prosodic words category.</i>	<i>151</i>
<i>Table 3-12 Mean F0 (in Hz) measures of the last syllables of each unit.</i>	<i>153</i>
<i>Table 3-13 The mean velocity measurement in the final syllable in each of the potential units.....</i>	<i>156</i>
<i>Table 3-14 Total number units that end with creakiness and breathiness in each text over the total number of units (1st line), and percentages (2nd line).....</i>	<i>157</i>
<i>Table 3-15 Value of pitch reset for each of the potential unit.</i>	<i>160</i>
<i>Table 4-1A typology of topic and focus constructions (from Neeleman et al. 2009:01)</i>	<i>185</i>
<i>Table 4-2 Rates of occurrences of first and second prominences</i>	<i>198</i>
<i>Table 4-3 The occurrences and positions of primary and secondary prominences and their correspondence with syntactic constituents, shown in verbal and nominal predicates.</i>	<i>199</i>
<i>Table 4-4 Final velocities by syntactic categories and by number of syllables.....</i>	<i>208</i>
<i>Table 4-5 Mean duration of the syllables in the IUs.</i>	<i>217</i>
<i>Table 4-6 Duration of the 1st prominent syllables shown according to the number of syllables in the word, and its syntactic category.....</i>	<i>218</i>
<i>Table 4-7 The structure of thethetic sentences in the dataset.....</i>	<i>226</i>
<i>Table 4-8 Position in the constituent of primary and secondary prominences in sentence focus.</i>	<i>228</i>
<i>Table 4-9 Rates of occurrences of primary and secondary prominences according to the syntactic structure of the sentence in focus.</i>	<i>229</i>

<i>Table 4-10 The dataset used for the analysis for the argument focus domain.</i>	<i>236</i>
<i>Table 4-11 The distribution of the subtypes of arguments in focus, related to their syntactic structure.</i>	<i>239</i>
<i>Table 4-12 Prominences shown by syntactic structure and argument position.</i>	<i>240</i>
<i>Table 4-13 Summary of correlates in the three focus domains tested in this section.</i>	<i>249</i>
<i>Table 4-14 The average final velocity values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.</i>	<i>260</i>
<i>Table 4-15 Average pitch values for IUS with given topics and comments, distinguishing topics that host the clitic =marlang, and those that do not.</i>	<i>262</i>
<i>Table 4-16 The average pitch excursion values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.</i>	<i>263</i>
<i>Table 4-17 The average duration values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.</i>	<i>264</i>
<i>Table 4-18 Subtypes of argument focus tested.</i>	<i>284</i>
<i>Table 5-1 Final velocity values of the syllable in simple declaratives, shown according to their position in the word and in the IU.</i>	<i>306</i>
<i>Table 5-2 The rates of occurrence of the falling, fall-rise and level contours of the first IUs in the prosodic sentences in this dataset, correlated with the types of second IUs.</i>	<i>320</i>
<i>Table 5-3 The mean F0 values in 2- and 4-syllables second IUs, as well as the average of the values for the syllables at the left and right periphery in all IUs. The values of the final syllables of the previous IUs are also shown, as they are useful to calculate pitch reset.</i>	<i>334</i>
<i>Table 5-4 The final velocity values of in the second IUs, shown for IUs of various length.</i>	<i>339</i>
<i>Table 5-5 The duration of the syllables in the second IUs, shown according to IU length.</i>	<i>341</i>
<i>Table 5-6 The duration of the syllables in the second IUs, shown according to IU length.</i>	<i>341</i>
<i>Table 5-7 The parameters ruling the prosodic encodings of second IUs in complex sentences.</i>	<i>347</i>
<i>Table 5-8 The rates of occurrences of the three types of contours in information and polar questions.</i>	<i>359</i>
<i>Table 5-9 Values of mean pitch, excursion size, final velocity and duration in the syllables of the vocatives.</i>	<i>381</i>
<i>Table 5-10 The location of the pitch minima and maxima in the sentences containing quotative expressions.</i>	<i>388</i>

Abstract

This thesis is a description of the prosodic patterns in Jaminjung, a language spoken in the Victoria River District in the Northern Territory of Australia. It is a quantitative and qualitative examination of the features associated with the intonational phenomena in Jaminjung. It is based on the idea that, while some aspects of prosody may be universal, each language has its unique characteristics. In this study I will make use of the PENTA model of intonation, a recent development that places communicative functions and articulatory constraints at the core of prosody, thus providing a clear explanation of prosodic phenomena, linking phonetics to semantics. The analyses are based on carefully selected representative tokens of the speech used in specific communicative situations by the Jaminjung speakers from recordings of spontaneous speech.

The features associated with the grouping function, that is, in the demarcation or organization of a string of words (or rather syllables) into chunks, are examined. Four main prosodic constituents are recognized: the prosodic word, the phrasal constituent, the intonation unit, and the prosodic sentence. They are distinguished at their left boundaries by pitch resets which increase from unit to unit. The larger constituents are cued at the right edge with F0 lowering and syllable lengthening, cues associated with finality in many languages.

The encoding parameters of some major information structural categories, topic and focus and contrast are investigated. A prominence is usually perceived on the first syllable in the focus domain. A [fall] pitch target is associated with this syllable; it is also marked by wider pitch excursions and longer durations. Topics, for their part, are marked by a [high] target on their initial syllables. The prosodic encoding of topics follows a scale of 'givenness', where more given topics are less marked than less given topics. Contrast in focused arguments and topics is encoded with a [fall] target on the initial syllable and thus share this feature with focus, but they also display a wider pitch excursion on all the syllables. This last feature marks contrast as an independent information structure category from focus and topic.

Declaratives, interrogatives and imperatives sentences are all predominantly uttered with a falling contour, however, they are clearly differentiated by pitch register – declaratives use lower reaches, imperatives higher reaches, and interrogatives somewhere in between.

DECLARATION

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on presentation of Theses.

Acknowledgments

My first thanks go to the Jaminjung and Ngaliwurru speakers who have welcomed me into their families, and taught me their language, with much patience, generosity and humour. I have many people to thank, starting with Dolly Bardbariya for her wonderful, unfailing keenness and resilience. And, of course, to all my dedicated teachers at Gilwi, Myatt, Bulla, and Bethnal camps. It is difficult to know who to thank first, so I'll move outwards from Timber Creek, starting with the people at Myatt. I am grateful to Jerry Jones and his commitment to language work. My thanks to all his family and to Judy Merchant and her husband. Josie and Debbie, however busy they were, could always make time for language work. Their help was precious for working out how to write things. Judy is not only a very patient teacher, she has been a great 'mate', accompanying me on many trips, showing me the country, and the plants and birds and everything. I am happy that she introduced me to Eileen Raymond at Fitzroy, who has contributed much in her quiet way. Down the road at Gilwi and Barak Barak are the Roberts. My heartfelt thanks to Doris Roberts and Laurie, her husband, and her sister Claire, who not only helped me learn language, but also shared much about their life and culture, and all the lovely, and lively, children. I wish to mention Laurie's mum, who welcomed me so warmly, she had such a big heart and was so much fun. I also wish to thank the ladies in Kununurra, Iza Pretlove, full of stories, Eileen Huddleston, so good at explaining things, and their sister Polly Warndanga, still giggling and full of knowledge when I met her. I wish to acknowledge their beautiful sister who has passed away since my first visit, I feel privileged to have met her. I am glad also have met with Bill Harney, who has such a lively spirit. There isn't enough space to thank individually all the people who have helped me. When I think of my time doing fieldwork, I don't have only pictures of language work in my mind, my head is full of images of fishing trips, afternoons talking with the grown-ups and playing with the children. I am grateful to the people in the communities in

Timber Creek for letting me share their lives. I have learned much, about language and many other things.

My time was made easier by the support of the Ngaliwurru-Wuli Association in Timber Creek who helped in many ways, particularly with providing accommodation. The staff and linguists at the Diwurrurru-Jaru Aboriginal Corporation in Katherine have also been most helpful, letting me use the facilities at the language Centre and providing much encouragement. Their vast experience has been invaluable. I am grateful to Salome Harris, for her kindness, friendship, and help with this work when she visited London. I also learned a lot by working with Glenn Wightman on his ethnobiological work. His knowledge, warmth and humour made fieldwork such a rewarding experience. I'd also like to thank Helen Harper and Michael Duffy in Darwin, whose warm hospitality helped in the transitions from one kind of life to another.

The person I feel most deeply indebted to is Eva Schultze-Berndt, the best supervisor in the world. She has been my linguistics teacher since the first day of the MA, and ever since has patiently, painstakingly guided my efforts. Thank you for your support, constant help and encouragements over these last years, months, weeks, days...up to the very last minute. I have been very lucky to be introduced to the Jaminjung and Ngaliwurru people by her, as she is loved and respected there. Her integrity, intellectual and personal, is a model to me.

I am also grateful for the advice and support I have received from my other supervisor, Yi Xu, at UCL. He has been enthusiastic from the moment I contacted him about this project, smoothing out difficulties, explaining and discussing problems, and making prosody a subject to be enjoyed rather than overcome.

All the people of the School of Languages, Linguistics and Cultures at the University of Manchester have been very supportive during my time there and I want to acknowledge their help. I wish to thank Louise Mycock for her insightful comments, and Martina Faller particularly for being a friendly supervisor/mentor.

The research reported here was also generously supported by the Dobes language documentation programme from the Volkswagen

foundation. I have also benefited from the expertise, technical and otherwise, of the team in Nijmegen. I am grateful to the other members of our documentation project also. To Patrick McConvell who shared freely his deep knowledge of the people of the area. I got to know Felicity Meakins in the course of this project, and it has been a great help to be able to discuss everything with her, from the most practical to the more philosophical, or linguistic! She is a very generous collaborator and friend.

My thanks also go to my fellow PhD students, Serge Sagna and Pete Budd, colleagues from the very start, who have been calmly encouraging through all these years; and to Dorothea Hoffmann who's been such a bright spark in Manchester. I also want to thank my fellow students in Manchester and the people in the linguistics department in SOAS who have provided stimulating discussions and camaraderie.

I have a debt of gratitude also to Peter Austin, Friederike Lüpke and the other people in the Endangered Language programme at SOAS. They gave me an opportunity to embark on these studies with the MA programme, and have warmly supported my endeavours since. My gratitude also goes to the researchers in the ARP (Laboratoire de la Recherche sur la Parole) at Université Paris 7 in France, to Philippe Martin, Hi-Yon Yoo and M. Boulakia who gave a great start to my prosodic studies. The colleagues at the Laboratoire, Marie-Pierre, Sophie, Catherine, In-Young, Jérémie, Yahia, and Henri helped greatly with my early attempts at prosodic analysis, with advice and encouragements and making me feel that it was all possible. I could never have been able to write this thesis without the help and support of the people around me.

I wish to thank Jacques Boizeau and Solange Lapierre (in alphabetical order!). They may have started me on this journey... I remember a conversation during a long walk in Muttrah. They both, in their own ways have great curiosity, never ceasing to try to understand things, people, life.

I am most grateful to Brigitte, little sister, for her gentle kindness through these years...never a moment of doubt or impatience, in all that time. *Merci ma très chère!* My thanks extend to all my family, who have remained steadfast in their encouragement. *Merci à ma mère et à toute la*

‘gang’. I wish to thank David Boucher especially. Who could dream of a better nephew than one who greets you with ‘what could I help you with?’. He has been very helpful with statistics, mathematics, software, programming difficulties and much more. Thanks to Martin Fisette also, my technical adviser, helping out with all things computer-related, often at very short notice. I am also grateful to Su, Una and Mark, Lewis, Beatrice and Amalie, for being there and reminding me that life exists outside linguistics.

Special thanks to Marie Morin, a great friend who understands everything. I am very lucky to have wonderful friends (who have put up with much neglect lately), thanks to Lise Côté, Véronique Vieyres, Séverine Drolet, and Bill.

My love goes to Philip, who has made it all possible, accompanying me to Timber Creek, and who has listened to countless sound clips, heard more about prosody than he ever cared for I am sure, and proofread all these chapters... I could never have a better mate.

Abbreviations in interlinear glosses

1, 2, 3	1st, 2nd, 3rd person
ABL	Ablative
ABS	Absolutive
ACC	Accusative
ALL	Allative
ASSOC	Associative
COLL	Collective
COMIT	Comitative
COND	Conditional
CONT	Continuous
CONTR	Contrastive focus
COTEMP	Cotemporaneous (“still”, “then”)
DAT	Dative
DEM	neutral demonstrative, usually ‘given’
DIR	Directional
DIST	Distal
DOUBT	‘-ever, I don’t know’
DU	Dual
EMPH	Emphatic clitic
ERG/INSTR	Ergative/Instrumental
EXCL	Exclusive
FIRST	“first”, already”
FOC	Focus
FUT	Potential/Future
GIVEN	‘Given’
IMP	Imperative
IMPF	(Past) Imperfective
INCL	Inclusive
IRR	Irrealis
KIN2	“your kin”
KIN3	“his/her kin”
LOC	Locative
MED	Medial
MOTIV	Motivative (“about”, “over”)
NEG	Negative particle
NOW	“now”, “then”
OBL	Oblique pronominal
ONLY	“only”
ORIG	Origin
PL	Plural
POSS	Possesser
PRIV	Privative (“without”)
PROPR	Proprietary (“having”)
PROX	Proximal
PRS	Present
PST	Past (perfective)
QU	Question marker

QUAL	Quality
RDP	Reduplication
REFL	Reflexive
SEMBL	Semblative (“like”)
SFOC 1	Sentence focus
SFOC2	Emphatic sentence focus
SG	Singular
SUBORD	Subordinator
TAG	Tag
TR	transitive

Abbreviations of Kinship terms

Br	brother
Fa	father
So	son
Si	sister
Mo	mother
Da	daughter
Wi	wife
Hu	husband
Ch	child

Conventions used in transcription and glossing

-	morpheme boundary
=	clitic boundary
.	separates categories encoded by a portmanteau morpheme
:	morpheme break not indicated in the text line
..	short pause
...	long pause
+	next/preceding line still in the same intonation unit
:(::)	lengthening
,	pause but non-sentence-final
[]	overlap (marks both overlapping strings)

Abbreviations used in textgrids

case	case
clitic	clitic
cov	coverb
dem	demonstrative
interj	interjection
KR	Kriol
n	nominal
nadj	adjectival nominal
prn	pronominal
v	inflected verb

Chapter 1 Introduction

People use speech to communicate. The verbal content of their messages is modulated by melodies and the study of how these contribute to meaning is the object of prosodic research in linguistics.

We use prosody for many purposes: to express emotion or surprise, to emphasise a word, to indicate where a sentence ends and so on. One can completely change the meaning of a sentence, for example ‘It’s OK’, by altering only the shape of its intonation. Ultimately, the prosodic aspect has precedence in the interpretation of the utterance over the verbal component. The central place of prosody in human language has been recognized and, while a fair amount of research has been done on the prosody of European and other major languages, prosodic descriptions of lesser-known languages are more rare. Jaminjung is one such language. It is spoken along the Victoria River in the Northern Territory of Australia.

This description of the prosodic patterns in Jaminjung is based on the idea that, while some aspects of prosody may be universal, each language has its unique characteristics. Moreover, the analysis and interpretation of these patterns in Jaminjung speech cannot be dissociated from the functions they serve. This approach is well served by the PENTA model (Xu 2005), which is used herein. This model is a recent development that places communicative functions and articulatory constraints at the core of prosody, thus providing a clear explanation of prosodic phenomena, linking phonetics to semantics. It is a powerful descriptive tool, particularly useful in the context of an as-yet undescribed prosodic system, such as that of Jaminjung, by anchoring the description of prosodic phenomena to communicative functions. A full explanation of this model is given in Chapter 2; briefly, it states that communicative functions are transmitted in parallel through encoding schemes, that are either language specific or universal, via a limited number of parameters which may be considered as the phonetic primitives (local pitch targets, pitch range, articulatory strength and duration). Importantly, the model does not stipulate the properties of the encoding schemes themselves, only providing a mechanistic framework for the schemes to be implementable. Indeed it is precisely the goal of this

investigation to define these schemes through a quantitative study of the measurements of the correlates of pitch and duration across datasets consisting of carefully selected tokens representative of the speech used in specific communicative situations by the Jaminjung speakers. A careful analysis of these measurements makes it possible to define the prosodic patterns of Jaminjung.

The thesis is structured as follows:

The remainder of this chapter presents an overview of Jaminjung and its speakers, as well as the context of the data collection for this research.

Chapter 2 is a presentation of the terminology used, and a review of the literature on prosody, with special attention given to the major theoretical models. Finally, methodological considerations are presented.

Chapters 3, 4, and 5 constitute the core of the analysis of the Jaminjung data. Each of the chapters is based on three well-recognized functions of intonation:

Firstly, Chapter 3 is concerned with the grouping function, that is, in the demarcation or organization of a string of words (or rather syllables) into chunks, often conventionally associated with prosodic structure. The chapter also provides an overview of the main phonological and morpho-syntactic features of Jaminjung, before evaluating its prosodic constituents and their encoding.

Chapter 4 is concerned with the marking of information structure of categories, notably topic and focus, and contrast.

Chapter 5 covers the function of distinguishing between sentence types by prosodic means, in this case, patterns associated with declaratives, interrogatives and imperatives. Other common constructions in Jaminjung are also examined: those with iconic lengthening, in which the durational aspect of an action or event is indicated by an elongation of the final vowel; vocatives; and finally quotatives.

Finally, Chapter 6 provides a summary of the preceding chapters, and concludes the thesis.

1.1 Jaminjung: speakers

Jaminjung and Ngaliwurru are two closely related language varieties spoken in the Northern Territory of Australia by approximately one hundred people. In this study, I will use ‘Jaminjung’ as a cover term for both varieties, for the sake of simplicity. Speakers of Jaminjung and Ngaliwurru historically occupied lands along both sides of the Victoria River (Figure 1-1, map from Meakins 2010:5). Today, they live in the townships of Kununurra (in Western Australia), Timber Creek, and Katherine. Formerly semi-nomadic hunter-gatherers, the Jaminjung and Ngaliwurru¹ still maintain close links to their traditional lands. There are now several small communities which live in outstations established in a family’s traditional country, usually in remote areas.

As is very common in Aboriginal communities, the speakers of the Jaminjungan languages are bilingual or multilingual, and it is unlikely that there would have been many monolingual people in the past. This multilingualism stems from a range of social and familial structures and is not an effect of the necessity of gaining at least a functional proficiency in the English-lexified Kriol, which has become the lingua franca of the region (Munro, 2000) and the language of interaction with non-Aboriginal people. Whilst Standard Australian English is the official language of schooling and most administration in the Northern Territory, the people are proud of their language and its cultural relevance is felt deeply. During my fieldtrips, I have witnessed older people using their traditional languages confidently and systematically in their daily life.

¹ See Schultze-Berndt (2000:7-14) for a detailed presentation of the history and social organization of the Jaminjung and Ngaliwurru people.

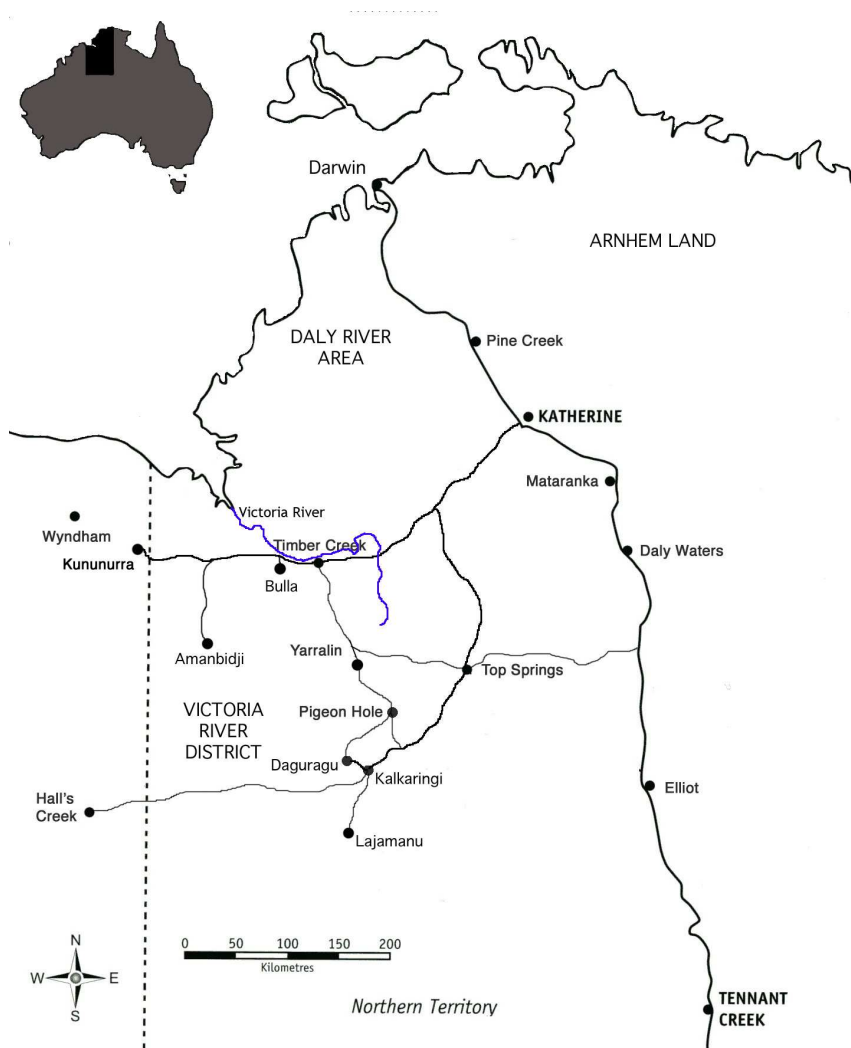


Figure 1-1 Map of part of the Northern Territory, Australia.

1.2 Genetic affiliation

The languages of Australia are usually divided into two main families, the Pama-Nyungan family, which occupies most of the Australian continent, and a group of distinct families usually grouped under the appellation ‘non-Pama-Nyungan’ or Northern, which are consequently much more heterogeneous typologically and comprise many language families between which the genetic relationships have not been established with certainty (Evans 2003).

Jaminjung and Ngaliwurru together with Nungali, which now has no speakers left, form the Jaminjungan family, also referred to as Djaminjungan, Yirram or Western Mirndi in the literature (Chadwick,

1997). Chadwick suggests a distant genetic relationship to the Barkly languages, or Eastern Mirndi, based on the form of the dual inclusive pronoun *mind*i which distinguishes the languages in this group from all other surrounding languages. The larger family, consisting of the Jaminjung and Barkly languages, is thus referred to as the Mindi (Mirndi) family. The Eastern and Western groups are now far apart geographically and differ greatly in their lexical and grammatical make-up. Details about the reconstruction and the prehistory of these language groups are presented in Harvey (2008).

1.3 Previous studies

By far the most extensive prior treatment of Jaminjung grammar is Schultze-Berndt's 2000 PhD Thesis, 'Simple and Complex Verbs in Jaminjung, A study of event categorisation in an Australian language', which also comprises a sketch grammar. Her research has led to a number of other publications, partly of a more comparative nature (Schultze-Berndt 1998, 2001, 2002a, b, 2003, 2007a, b, 2008). Since 1993, Schultze-Berndt has collected a substantial corpus of recordings and accompanying annotated texts in Jaminjung data, which have generously been made available for this research and make an extremely valuable contribution to the datasets analysed herein.

Prior to this a number of other linguists and anthropologists have collected Jaminjung data (in the form of field notes, audio-recordings, etc.). The earliest recorded wordlists are from Arthur Capell in 1938. Around the same time, a brief ethnographic survey was conducted by Stanner. The first grammar sketch was produced by John Cleverly in 1968. An unpublished grammar of Ngaliwurru was compiled by Janet Bolt, William Hoddinott and Frances Kofod (1971), presenting a fairly detailed coverage of the morphology and a syntactic sketch, and includes a word list of around 500 items and some texts. In 1971, Michael Walsh recorded extensive vocabulary and some grammatical information, which is available only in manuscript form (Walsh 1971). Mark Harvey also made recordings of Ngaliwurru during salvage work on Nungali, conducted mainly in 1996.

1.4 Fieldwork setting and data collection

The data collection for this research took place during three trips to the Northern Territory, between July 2005 and November 2007, amounting to 8 months in total. Fieldwork has been conducted as part of a project aiming at documenting the linguistic and cultural knowledge of the Victoria River district, funded by the DOBES programme (documentation of endangered languages) of the Volkswagen Foundation. The project resulted in the collection of audio and video recordings, the creation of lexical databases and annotated texts, as well as community-oriented materials.

During these trips, I was based in Timber Creek in accommodation made available by the Ngaliwurru-Wuli Aboriginal Corporation. This provided a good base to work with the Jaminjung and Ngaliwurru speakers in and around Timber Creek, as well as in Kununurra and Katherine (270-300 kms distant).

Most of the people whose speech is analysed here are experienced language workers, having been involved in 'language work' to various degrees through the years.

A normal working day involved collecting speakers at their camps, running various errands for and with them and finally sitting down in a quiet spot for the more formal language recording sessions. For these sessions, speakers were paid the rate set by the local Language Centres. Recordings were limited to at most two hours a day, given that most of the speakers are aged or had other commitments. The impossibility of working indoors has affected the quality of the recordings and is one of the major setbacks of this research.

Concerning the documentation of prosodic phenomena in endangered languages, Himmelmann (2006:163) states that "prosodic phenomena are highly variable and susceptible to contextual influences. This makes it difficult to recognize basic distinctive patterns. Prosodic pattern recognition is much facilitated by having the same utterance produced by a number of different speakers (or at least to have multiple versions of the same utterance".

In the Jaminjung context, such a strategy proved very difficult to apply. Firstly, the very concept of repeated sentences is an unlikely one in a language in which word order is regulated by information structure (hence information that is new the first time a sentence is uttered may not be judged to be so when a speaker repeats it, and may thus change in subsequent repetitions). Further, speakers do not relish participating in artificial communicative activities, and although they always tried to comply with my requests with using various elicitation tools, more culturally relevant situations (e.g. providing example sentences for the dictionary or explaining plant uses) were better suited to data collection. The datasets used in the following chapters are derived from three sources: narratives, picture-prompt narratives, and elicitations. The narratives consist of both personal anecdotes and mythological stories; the picture-prompt narratives are based on more widely used materials, such as the story 'Frog, where are you?' (Mayer, 1994 (1969)), and some of the tasks from the Questionnaire for Information Structure (QUIS) materials developed as part of the SFB 632 Information Structure (Skopeteas et al. 2006) research project. Finally, some of the data used for prosodic analysis was recorded in the course of the documentation of the ethnobiological knowledge of the speakers. Recordings usually involved more than one speaker. Two recording devices with a central microphone were used to avoid the risk of data loss if one device failed (a Marantz solid state recorder with a Rode Condenser microphone and a Sony MD recorder).

The source of each example sentence is given in square brackets below the translation line, consisting of the speaker's initials, the collector's initials, year of recording, and a code referencing the session from which the example sentence is extracted. To respect the speakers' anonymity, they are referred to only by their initials in the text. For the purposes of this research, the materials used are all from recordings in which speakers are aware that they are recorded for later linguistic analysis and have given permission for this recording to be presented to an audience.

The textgrids and sound files of the examples included in the following chapters are found on the accompanying CD.

Chapter 2 Prosodic research

Prosody is often described as the ‘music of speech’. The term can be traced back to Ancient Greek, where prosody is a musical term that means, loosely, ‘with the song’ or ‘accompaniment’ referring to the musical aspect of the words. It is used in versification to refer to rhythmical patterns, rhyming schemes and verse structure (Fox 2000:1-3).

In linguistic contexts, prosody refers to the properties that create the melody of speech, with the understanding that this melody is not random, that its properties form an organized, meaningful system that lends itself to investigation. Prosody’s main purpose is to aid the transmission of information in communication. This information is manifold. Crystal (1979:36-37), for instance, distinguishes many functions for prosodic patterns: grammatical (statement-questions, negative-positive, singular-plural etc), semantic (organisation of meaning in discourse, emphasis, parenthesis etc.), attitudinal (emotions), psychological (stressed elements help memory), and social (indication of age, sex, social status etc.). This dissertation will focus specifically on those functions that are usually recognized as more ‘linguistic’: highlighting (information structure, prominence, and emphasis), phrasing (chunking, stress and rhythm) (Grice and Bauman 2007), and distinguishing sentence types. The view taken here is that a meaning (or function) is expressed through a form which is encoded by articulatory means (Xu 2005).

This chapter provides an overview of existing research on prosody, a presentation of the theoretical framework adopted here, the PENTA model, and closes on some methodological considerations.

2.1 Basic notions in prosodic research

The terms *prosody* and *intonation* are often used interchangeably, but a distinction can be made: the former is a more general term, encompassing word and sentence phenomena, while the latter is usually restricted to the variation in pitch at sentence level (Di Cristo and Hirst 1998). Ladd’s (1996:6) often-quoted definition of intonation includes three

key elements, it is: “the use of suprasegmental phonetic features to convey postlexical meanings in a linguistically structured way”. The term *suprasegmental* describes the domains of application of the prosodic properties, in the sense that they apply to units bigger than the segments that define the verbal contents of an utterance (Lehiste 1970). Ladefoged (1993:15) points to the suprasegmental nature of prosodic features when he comments that they “are characterized by the fact that they must be described in relation to other items in the same utterance”.

2.1.1 Prosodic phenomena

Phoneticians distinguish between the articulatory, acoustic, auditory, and perceptive properties of prosodic phenomena. The articulatory properties of prosodic phenomena include “variations in the amplitude of articulatory movements, variations in air pressure, specific patterns of electric impulses in nerves leading to the articulatory musculature, especially those innervating the larynx” (Werner and Keller 1994:24). These articulatory movements result in the emission of sound waves which can be observed and measured with acoustic signal analysis tools and form the main material for the study of prosody.

The main acoustic properties of speech are: fundamental frequency (F0 hereafter), intensity, duration, pauses and position and form of the tongue and vocal tract. Their perceptive equivalents are pitch, loudness, length, pauses and phonation types (which results from the characteristics of the glottal source). They will be defined further below.

These properties can all be related to prosodic characteristics such as intonation, stress, rhythm, and speech rate; however, it is not possible to establish a one-to-one mapping of the acoustic properties to the prosodic characterisation and other phonological features (Di Cristo and Hirst 1998, Fox 2000, Ladefoged 1971). F0 is often related to intonation but does not constitute its sole feature; length and intensity are both linked to stress, but again, it is not possible to exclude F0 in the analysis of stress in languages. Ladefoged (1971:84) even states that ‘the more we look at prosodic features,

the more apparent it is that their physiological correlates are inextricably combined’.

Fundamental frequency (F0) is the lowest frequency produced by a vibrating object. In speech, F0 is sometimes referred to as the first harmonic². It is determined by the speed with which the vocal folds vibrate, measured in cycles per second [cps], or Hertz [Hz], which explains why measurements of the F0 in human speech can only be done on voiced segments. The variations in the vibrations of the vocal folds are perceived by listeners as changes in *pitch*. The faster the vocal chords vibrate, the higher the perceived pitch. In human languages, pitch is associated with the features of tone and intonation. It is important, however, to keep in mind that what is measured in speech analysis is not pitch but F0, and that there may be discrepancies between the production of the fundamental frequency and its perception.

Length is the acoustic counterpart of *duration* and is measured in milliseconds (ms). A segment of a certain length is produced by the speaker, while the same segment is perceived with a certain duration by the hearer. The terms are used quite interchangeably in the literature. The position of the segment within the word, as well as its relative prominence, can both influence its length. In intonation studies, the measured units of length are often syllables.

Intensity is also called *amplitude*. It is the production correlate of the perceived *loudness* of the utterance. Intensity is a measure of acoustic energy and is dependent on the amplitude of the sound wave, that is, on the variation in air pressure. It is measured in decibels (dB), 1 dB corresponding to the smallest change in loudness that can be perceived by human beings (Ladefoged 2003:90).

Voice quality is related to phonation events. These events are defined by the position and movements (vibrations) of the vocal folds and are expressed acoustically by the lower and higher harmonics. These harmonics are largely dependent on the size and structure of the vocal tracts of the

² The other harmonics are thus defined as multiples of the fundamental frequency.

speaker and are thus often the principal characteristic of that speaker's speech. Gordon and Ladefoged (2001) state that some types of phonation, such as creakiness, can signal syntactic or pragmatic completion or sentence final intonation (see also Laver 1994, Redi and Shattuck-Hufnagel 2001, Chavarria et al. 2004). Creaky voice is a type of phonation characterised by slow and irregular vibrations of the vocal folds. In creaky voice, only the front part of the vocal folds is vibrating, giving a very low frequency. Perceptually, its effect can be described, after Catford (1964:32), as a "rapid series of taps, like a stick being run along a railing". In breathy phonation, the "vocal folds vibrate ineffectively and never come fully together" (Ní Chasaide and Gobl 1995:447). The flow of air through the glottis is thus turbulent, which results in an auditory impression that has been described as "voice mixed in with breath" (Gordon and Ladefoged 2001:2). Figure 2-1 shows spectrograms with modal (normal), breathy, and creaky phonation in English.

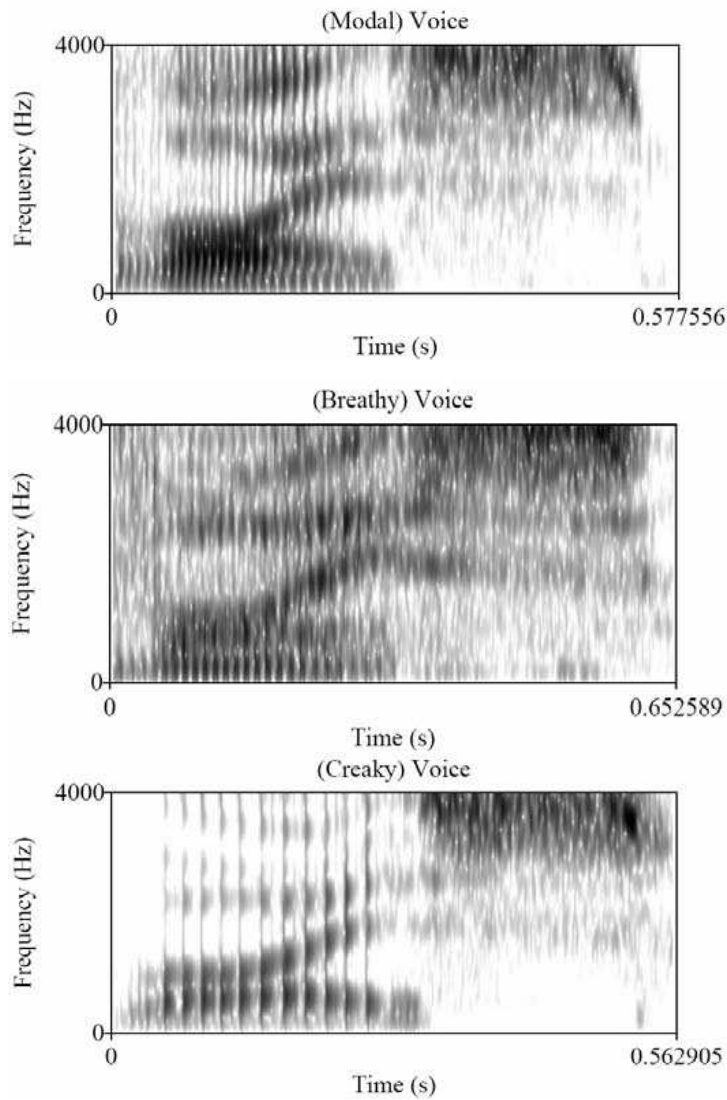


Figure 2-1 Spectrograms of the word *voice* produced with different phonations. From top to bottom, it is produced with modal voice, breathy voice, and creaky voice, respectively (adapted from Yoon et al. 2008:3).

Researchers have explored other characteristics associated with voice quality such as *spectral tilt* which measures the relative intensity at different harmonics rather than over the entire spectrum. It is quantified by comparing the amplitude of fundamental harmonics to that of higher frequencies, e.g. the second harmonic. In a creaky voice vowel, for example, the amplitude of the second harmonic is greater than that of the fundamental, leading to a greater positive value of the spectral tilt.

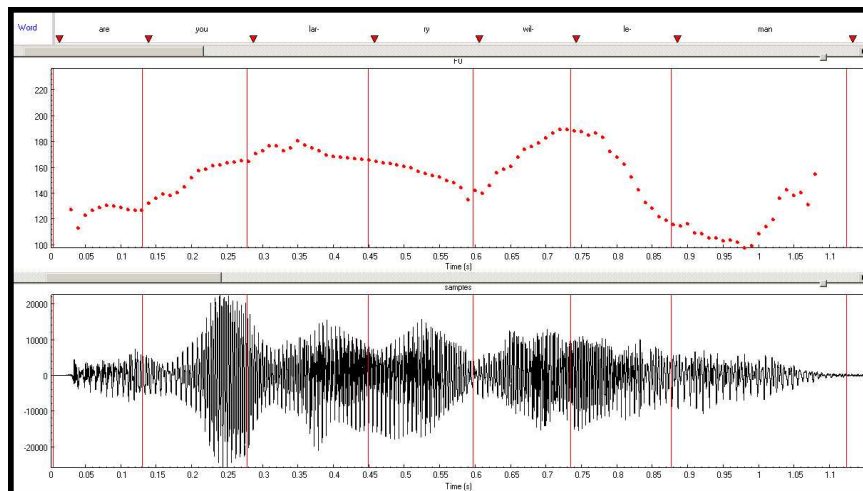
The features described so far are those usually associated with the linguistic functions of intonation. It is generally assumed that the more linguistics functions of intonation differ fundamentally in their encoding to

the paralinguistic functions. Paralinguistic phonetic features are defined in Crystal (1964:96, discussed in Gussenhoven 2004:69) as modifications by voice qualifiers (such as whisper, creak, falsetto etc.) or by voice qualifications (like a giggle, laugh, sob etc.) to express attitudes or the emotional state of the speaker. According to Roach (2000), linguistic prosody is strictly voluntary, while paralinguistic phenomena are either voluntary, such as with attitude and speaking style, or involuntary such as age, physical state, health and emotions. It is also often stated that the features associated with intonation's linguistic functions are categorical, while those associated with the linguistics functions are gradient in nature (Ladd 1996:33-38). The categorical/gradient nature of the encodings is discussed further below (§2.1.2).

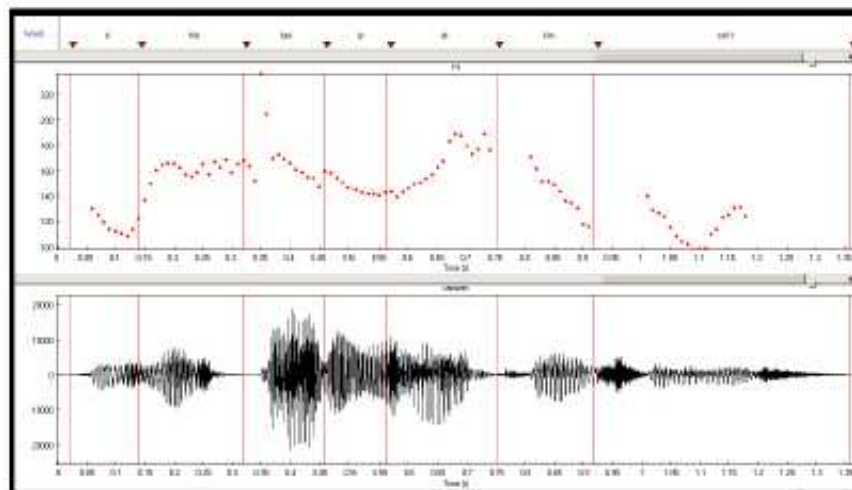
The term *microprosody* refers to the possible effects of the articulation of a string of segments on the suprasegmental level. These effects are not directly controlled by will nor do they convey primary phonological information. The phonetic details of microprosody are well documented (see Gussenhoven 2004:5-11, Ladefoged 2003:86-90 for a summary). Some of the features observed at the microprosodic level are:

- Voiceless gaps: F0 can only be calculated on voiced segments.
- Intrinsic pitch: the F0 of vowels [i, u] is higher than [a] for purely articulatory reasons.
- Obstruent perturbations: the F0 falls into an obstruent; rises out of a voiced obstruent, and falls out of a voiceless obstruent (Ladd 2008).

A pitch tracker shows these variations, as demonstrated in Figure 2-2 (from Ladd 2008), but perceptually, listeners tend to ignore them. The pitch track of the top pane is unbroken, illustrating the utterance consisting of voiced sonorants *Are you Larry Willemans?*; the lower pane is punctuated by gaps which correspond to the stops in *Is this Betty Atkinson's*.



Are you Larry Willeman?



Is this Betty Atkinson's?

Figure 2-2 A pitch track showing how the F0 is disturbed by consonants with an extreme constriction in the vocal tract (from Ladd 2008³).

Macroprosody can be equated with the temporal and intonational events at the suprasegmental level.

2.1.2 Intonational features

This section presents terminology used in describing phenomena related to intonation. The same terms are used in many intonational studies, often with slightly different meanings. An effort will be made here to

³ Online appendix to accompany Ladd (2008).

remain as theory neutral as possible or at least to point to various interpretations of the terminology.

In the study of intonation, the most basic unit of analysis is the *utterance*, defined as a natural unit of speech bounded by breaths or pauses (Aronoff and Rees-Miller 2001:xvi–823), rather than the sentence, which is associated with syntactic analyses often based on written materials. In this study, utterance will be used as a general term. I will make use of more specific expressions to refer to the units analysed, such as Intonation Unit (defined specifically in §3.1.1).

A *tune* is the basic pitch pattern of an utterance. It is expressed visually with an extracted pitch track as shown in Figure 2-3. Note that examples for which a picture representation is given are followed by an interlinear gloss for ease of reference. Kriol words are indicated by pointed brackets <>.

(1)

yawayi	ginarrga	burrina:		mung	yirrurra-ngayi-na
yawayi	ginarrg	burrina		mung	yirrurrangayina
yes	sharpen	3pl:3sg-chop.PST		watch	13pl:3pl-see-PST

<yuno>	gura-gurang
you.know	RDP-old.man

'Yes, they used to sharpen it, we used to watch them, you know, the old men'...
[EH:ES08_A04_02]

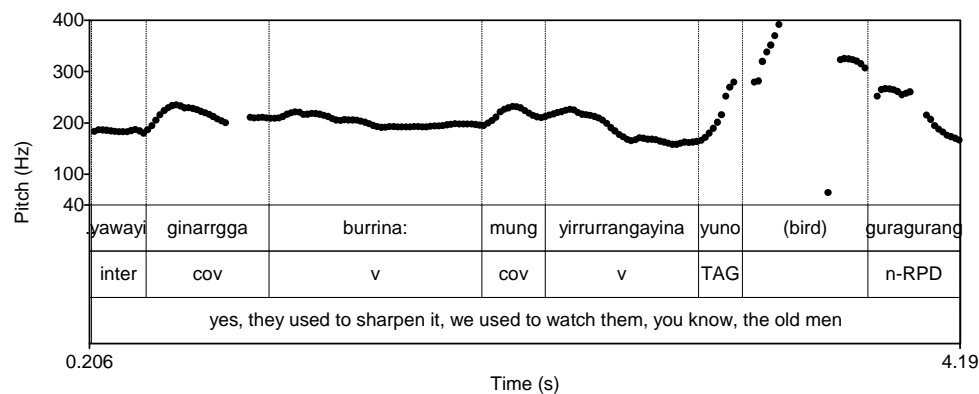


Figure 2-3 Pitch track.

In some of the early work on intonation, particularly that of the British school, the whole tune was the basic unit of analysis, with meaning directly associated to it (see also Delattre 1966, for French). According to these analyses, the tune starts at the *nucleus*, the most prominent syllable

(see below), and continues to the end of the phrase. Later studies have shown that a tune can be broken into parts which have some degree of independence (Fox 2000:291-292). In the Autosegmental Metrical theory (AM hereafter), the tune is the melody that is created by sequences of *pitch accents* and *boundary tones* over the course of an utterance; the intonational tune is manifested in the F0 contour. In this framework, it is still unclear whether the meaning is carried by the tune as a whole or whether the individual tones are the units which carry meaning (Clark et al. 2007:364).

The term *intonational contour* usually applies to the pitch movement of an utterance; it can also apply to F0 movements associated with various phonologically relevant domains.

Stress is a structural property of lexical words. In many languages, some syllables are pronounced more prominently, bringing forth the notion of *strong* and *weak* syllables. In some languages, stress is fixed (fixed-stress languages), determined by the position of the syllable in the word, notwithstanding the context or changing according to predictable rules (regular stress rules); and in other languages stress may fall on any syllable (free-stress languages). The functions of stress vary: in a free-stress language, such as English, stress distinguishes lexical meaning (transport: v and transport: n). In fixed-stress languages, it functions as a boundary marker. This is particularly relevant in the context of Jaminjung, and will be discussed in chapter 3 as part of the grouping function. Apart from its main lexical function, the stressed syllable is also the expected location of the intonational accent (below), if such an accent should occur.

Accent is a property of a word in the context of an utterance. It marks intonational prominence in order to highlight important words in the discourse. Traditionally, the term accent is restricted to a prominence that involves pitch. It is defined as a clear pitch movement that occurs on stressed syllables and its most common configuration is an F0 peak (Cruttenden 1997:16-17). It is called *tonic accent* in the British school of English intonation, and marks the part of the utterance that is in focus. It is the most important contour and the one by which tunes are classified. In the AM theory, the *nuclear pitch accent* or *nucleus* corresponds to the last

accented word in the intermediate phrase⁴; it is the most prominent accent of the utterance (Grice and Bauman 2007).

Boundary tone is a term originating in the AM theory, but it has become widely used in the literature. It is defined as a pitch movement associated with the edge of an intonational unit and gives indications of its relationship to an adjacent unit.

The exact nature of pitch accents and boundary tones is still disputed. In the British school they are viewed as pitch movements or contours, while the American structuralists and AM theorists see them as levels. In the latter, pitch accents and boundary tones can consist of a single tone (High or Low), or, in the case of pitch accents, they can be composites (High-Low, Low-High, High-High, Low-Low), giving the perceptual effect of a contour.

Declination (or *downdrift*) refers to the fact that the pitch at the end of an utterance is usually lower than at the beginning. It is said to be a universal phonetic effect, at least in declarative sentences⁵, caused by the diminishing air supply in the lungs which leads to a decrease in subglottal air pressure. However, it has been observed that the rate of F0 decline is typically much higher than the decline of air supply, and that there is a broadly inverse correlation between length of utterance and the degree of (negative) declination. In other words, short utterances have a steeper rate of F0 decline than long utterances. “If declination were purely a function of lung volume one might expect much the same rate of F0 declination no matter how long the utterance” (Ohala, Dunn, and Sprouse 2004:163). Thus an alternative hypothesis proposes that F0 declination is phonological rather than phonetic, imposed by speakers during the planning of their speech, presumably to make parsing easier for listeners.

First noted in tone languages, *downstep* refers to the effect of declination which causes successive tones to become lower and lower in pitch, with the result that, at the end of an utterance, a high tone can have a

⁴ A more complete account of the models is given in §2.2 below.

⁵ See Ladd (1996) for a discussion.

fundamental frequency (F0) which is as low as the low tone at the beginning of the utterance. In the AM theory, it is considered that the high element of a pitch accent may be lowered in the pitch range relative to a preceding high tone, hence downstep is seen as the grammaticalisation of declination (Gussenhoven 2004:98). Pierrehumbert and Beckman (1988) recognize the effect of declination, but point out that its magnitude is much smaller than that of downstep. Then again, Xu (1997) has found that downstep is gradient rather than categorical, making it less likely to be a phonological feature than a mechanical process. “A high pitch is lowered when following a low pitch because the pitch-controlling mechanism does not fully recover from the production of the low pitch”. Declination and downstep are illustrated in Figure 2-4 (adapted from Liu and Xu 2005) in which a gradual lowering of pitch is observed from the beginning of the utterance to the end, corresponding to declination, while the lowering of the successive high tones corresponds to downstep.

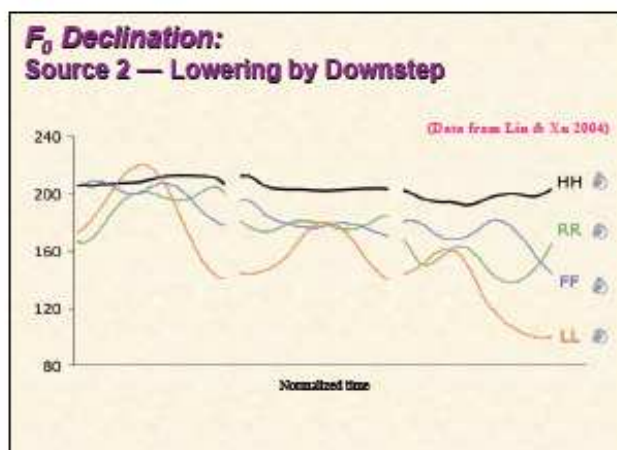


Figure 2-4 An illustration of declination and downstep (adapted from Liu and Xu 2005).

Pitch range is the relative variation in pitch produced by a human voice. Laver (1994:457) distinguishes three possible kinds of pitch range. The ‘organic pitch range’ of a speaker’s pitch, which is the maximum range a speaker’s voice can physically reach within the biological limitations imposed by the shape and size of his/her larynx. The ‘paralinguistic pitch range’, the range of pitch exploited for paralinguistic purposes, such as expressing attitudes, emotions, etc. And finally the ‘linguistic range’, the

range within which the phonologically relevant pitch usually varies in a paralinguistically neutral context⁶. Pitch range in this sense is language specific, inasmuch as speakers use its variations to convey meaning as specified in a language's grammar (Chen et al. 2004).

Alignment can be described as the interaction between the form of the F0 contour and the segmental content of the utterance. Of interest is the timing of the accent with respect to the accented syllable (Kohler 1991). Questions of alignment have generated a fair amount of research in the AM theory, particularly with the definition of association rules⁷, as the pitch accents are viewed as phonemes and the variations in their alignment as significant (early peaks and late peaks etc.). Alignment with the syllable is thus a cue to meaning.

Prosodic structure refers to the organization of a unit of language “in terms of the relationships between its parts” (Fox 2000:331). It is usually assumed that the domains of intonational phenomena are of different sizes, the larger including the smaller, and organized in a hierarchical manner. In intonation studies, it is described in terms of prosodic phrasing, in the sense that an utterance can have a prosodic phrase structure as well as a syntactic phrase structure.

Finally, traditional intonational typologies categorize languages as *intonation languages* (e.g., English), *tone languages* (e.g., Mandarin) or *lexical pitch-accent languages* (e.g., Japanese). In tone languages, each syllable is assigned a tone which is used to distinguish words. In pitch-accent languages, pitch can distinguish words, but all words do not have a pitch accent. In intonation languages, it is usually held that pitch is not used to make semantic distinctions between words (Cruttenden 1997:8-12). However, Fry's (1958) investigations of the prosodic cues associated with English lexical stress suggest otherwise.

⁶ It could be argued that ‘paralinguistic neutrality’ is never truly achieved, as there are no contexts that can be devoid of emotions or attitudes.

⁷ See Ladd (1996:53-55) for a discussion.

Much in prosodic analysis is based on the observation of prominent units in speech. *Prominence* is a broad term that can cover both stress and accent. The term will be used in this research to describe the salience of some units of speech, marked by variations in pitch, length or loudness⁸ (Cruttenden 1997:7). It is understood in the sense of Brazil (1997), as one of the places in the spoken utterance where speakers can make meaningful choices. It is thus assumed that prominence is functionally relevant, as pointed out in Cole et al. (2008), in their perception experiment involving naïve listeners: “Prominence was defined as a highlighting function used by the speaker for the listener’s benefit, to make some words stand out relative to others”. It is generally agreed that the three features of pitch, length and loudness form a scale of importance in bringing syllables into prominence, pitch being the most significant, followed by length and loudness (Cruttenden 1997:13). Prominence involves comparing syllables with each other and could therefore be seen as gradient rather than categorical. For instance, Poiré (2006) asked seven prosodic experts to label a 3-minute speech recording (165 syllables) in French for prominences. The proportion of syllables marked as prominent varied from 19% to 49%, showing how elusive the notion of prominence can be. Furthermore, it is known that listeners’ expectation of greater prominence in a foreign language is influenced by the prominence patterns of their native language and also by their degree of competence in the target language (Dellwo and Wagner 2005). These issues are acknowledged and considered, especially as prominence judgments are used as a basis for analysis in this study. The notions discussed above will all play a role in the discussion of Jaminjung. In the next section, the major theoretical models are presented, with a view to situating the present study in the wider frameworks of intonational research.

⁸ Prominence is discussed further in §4.2.

2.2 Theoretical Models

The PENTA model of intonation is used for the analysis of Jaminjung in this research. For our purposes, it best meets with Carmichael's (2003:18) description of an ideal model of intonation:

(...) would be able to identify the underlying intonational units of a language and their arrangement in speech while also explaining the surface variability of identical underlying patterns. It would be able to account for a broad range of speech, including different speech rates and conditions (...) It would clarify the degree and manner of covariance among intonation, syntactic, and discourse structure...

Before introducing the model in some detail, a brief review of other prevalent prosodic models/theories is presented in order to situate the model in current intonational research. The section opens with a discussion of the place of prosody within the wider study of language, and more specifically its place within phonetics and phonology.

2.2.1 Phonetics or phonology

To account for prosodic phenomena, many models have been proposed which distinguish themselves in various ways, "by their theoretical motivations, the primitive units and measures they consider relevant, and their purpose of application" (Carmichael 2003:16). One thing they all have in common is that they all recognize that prosody cannot be studied within the confines of a single system (understood here in the sense of the various sub-disciplines of linguistics: phonetics, phonology, morphology, syntax, semantics, and pragmatics).

The study of prosody straddles both phonetics and phonology and the debate as to its definitive place is not yet resolved. The difference in the approaches is simplified by Ladd and Cutler (1983:1) who state that researchers display their own preference for either "making measurements or constructing models". On a more serious note, Fox (2000:274) states that intonation can be defined phonetically as a continuous pitch pattern, and phonologically as a succession of significant rises and falls organized as a structure. Or at least, as pointed out by 't Hart and Collier (1975, in Couper-

Kuhlen 1986:83), as “potentially distinct pitch events [which are] grouped together into ‘meaningful’ categories”. Phonetic approaches focus on finding the acoustic cues to specific intonational phenomena, while phonological approaches assume an abstract level of representation for intonational features, trying to find discrete units to categorize phenomena. Whether a phonological level is at all necessary is a matter of discussion⁹. Phonetics-based models, in seeking to establish direct links between specific features and different kinds of communicative functions, situate prosody outside grammar, while phonology-based models describe an abstract system of intonational features and therefore an indirect relation to their functions. Phonological models, however, allow for interactions with other parts of grammar (von Heusinger 1999:54).

2.2.2 Characterizing phonological models

Many of the main current theories attempt to explain the linguistic component of intonation with a phonological model. Intonational phonology can be studied using the same principles as segmental phonology (see Gussenhoven 2004:61-69). Pitch events are represented at an abstract level; pitch contours should form an inventory of categorical, contrastive patterns, and the categories should reflect some linguistic meaning, while some of the variations should be non-distinctive, as in segmental phonology. What makes a phonemic analysis of intonation more difficult than one of segmental contrasts is that the significant distinctions they express are not related solely to lexical meanings but rather to discourse functions.

Phonological models have been characterized according to three sets of parameters. Firstly, according to the significant units they describe: either pitch targets as in the AM framework, or pitch configurations or tunes as in the British school. Secondly, to the organization of these units, whether

⁹ See Fox (2000:274-276) for an overview of the issues and Ladd (1996:20-41), von Heusinger (1999), Rossi (1997:13-54), DiCristo and Hirst (1998:1-45), Hirst, Di Cristo and Espresser (2000:51-57) for an in-depth discussion.

intonation components are linearly sequenced or superposed. And finally, to the meanings they encode, linguistic and paralinguistic, mentioned in §2.1.1 above.

The distinction between *superposition* and *tonal sequence models* (also called *linear models*) refers to the organization of the intonational components (Rossi 2000:142). The former are hierarchically organized models that interpret F0 contours as complex patterns resulting from the superposition of several components. In the latter, F0 contours are generated from the sequencing of phonologically distinctive tones, or categorically different pitch accents, that do not interact with each other. A typical superpositional model is that of Fujisaki (1983). Among linear or tone sequence models are those of the British school (Kingdon 1958, O'Connor and Arnold 1973, Cruttenden 1997), the AM model (Pierrehumbert 1980), and the Institute of Perception Research, or IPO, model ('t Hart, Collier and Cohen 1990).

The distinction between linguistic and paralinguistic meaning is useful, but it is also controversial. Firstly, linguistic/paralinguistic features are not always clearly demarcated in natural speech. Further, the categorical/gradient difference in their encoding is questioned (Grice and Bauman 2007:14). Taylor (2000:21) argues against this separation of roles. His view is that intonational phonology does not necessarily need to be categorical. He explains that making use of an abstract level of representation has made it possible for phonologists to describe sound patterns, and has thus been instrumental in the rapid advances of the last decades, without having to deal with pitch values directly. He stresses that the success of this approach has more to do with the adoption of an abstract level of representation in the broad sense, rather than a phonological one in the traditional discrete, categorical sense. He goes on to state that, although there are similarities between the structures of segmental phonology and those at the suprasegmental level, it does not follow that they are 'phonological in the same way'. At the segmental level, the phonological space is discrete, as shown in experiments in which people either perceived words such as /pill/ or /bill/ and not a new word, meaning something in between when voicing was artificially modified to stand just half-way

between a /p/ and a /b/. In intonation, it is possible to imagine a sound S_a associated with a meaning M_a , and a sound S_b associated with a meaning M_b . It is conceivable that a sound halfway between S_a and S_b could be given a meaning in between M_a and M_b (Taylor 2000:21). In his view, pitch accents occupy positions in a multi-dimensional sound space, H^* and L^* (etc.) representing points of particular importance in this space, with the possibility that accents occupying positions in between could be described as either. Although Taylor's Tilt model is not adopted in this analysis of Jaminjung and targets in the PENTA model are not viewed in the same way, Taylor's comments nevertheless address the important issue of the specificity of an intonational phonology and the association of form and meaning.

Grice and Bauman (2007) summarize the prosodic functions of intonation, and the means used to express them, through the graph in Figure 2-5 below. These authors, while not altogether espousing Taylor's argument, also point out that it is "not possible to state *either* that categorical means are used to express only linguistic functions, *or* that gradient means are used only for paralinguistic functions". They also advise keeping an open mind when relating form to function.

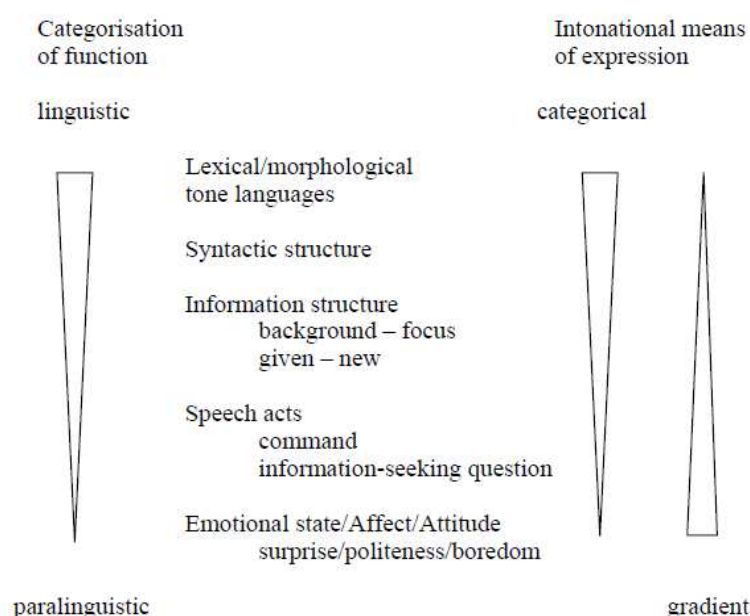


Figure 2-5 Functions of intonation and their intonational realization. On the left side is a scale from the paralinguistic to the linguistic, on the right is a scale from the gradient to the categorical. In the middle are the functions of intonation (from Grice and Bauman 2007:14).

Not only is the nature of the intonational phonemes a matter of discussion, a great deal of debate also concerns the dissociation of *form* and *function* in some prevalent models (Kohler 2004, Xu 2006, amongst others). For instance, while the AM model is widely used to describe intonational phonemes in languages, the original aim of linking intonational phonemes to meanings (Pierrehumbert 1980:59) has so far eluded its proponents.

The approach favoured in this study, the PENTA model, insists on maintaining a constant linkage between functions and forms. While it is anchored in phonetics, it is phonological in its interpretation.

What we need to recognize is that although the meaningful contrasts in speech are realized through form, the link between function and form is never broken, as otherwise successful communication would not have been possible (Xu 2006:1).

Before proceeding to describe this model in detail, the major phonological models which have become prevalent in the last two decades will be briefly introduced below, especially the AM model of intonational phonology.

2.2.3 Overview of the Models

2.2.3.1 The American Structuralists

The American Structuralists viewed intonation as composed of three subsystems: pitch, stress and juncture, combining to form intonational phonemes. They described contours as pitch levels, of which they usually count four (see Pike 1945, Trager and Smith 1951, Wells 1945). The American Structuralists with their interpretation of intonation in terms of levels were the precursors for the AM framework (von Heusinger 1999:56-61).

2.2.3.2 The British school

The description of intonation in the British School was originally pedagogical (or didactic) in nature, and not aimed at creating a theoretical framework. It is based on dynamic patterns such as ‘rises’, ‘falls’, or ‘rise-falls’. Speech is divided into tone groups marked by a tune, which consist

of, at most, four components, the most important of which is the nucleus, which coincides with the accented syllable; the tail, which follows the nucleus; the head, from the first accented syllable of the intonation phrase to the nucleus; and the pre-head which comes before the head¹⁰. The main proponents of the British school, having proposed variations of this structure, are Palmer (1922), Kingdon (1958), O'Connor and Arnold (1961, 1973), Crystal (1969) and Halliday (1967) and later Couper-Kuhlen (1986)¹¹. The tones proposed for English by Halliday (1967:16) are shown in Figure 2-6 below. Halliday's work has been very influential, as he not only provided a set of categories for the phonological description of intonation, but also "plac(ed) these categories within an overall model of language structure and meaning" (Fox 2000:280).

Term in the system	Visual Symbol	Tonic movement (= pitch movement)	Terminal tendency
1	\	falling	low
2	/	rising	high
2	∨	falling-rising	high
3	—/	rising	mid
4	↘↗	(rising-)falling-rising	mid
5	↗↘	(falling-)rising-falling	low

Figure 2-6 Halliday's tones (1967:16).

Similar models were developed for French by Delattre (1966) who described ten fundamental intonation contours on a grid of four levels, covering about one octave, and by Martin (1982) who defines intonation in terms of rhythmic groups, each containing no more than one stressed syllable, which carry the intonation contours. He distinguishes six sentence internal and four sentence final contours. These contours are governed by sets of 'dependency relations' (based on Hjelmslev's glossematics), forming an intonational grammar which makes it possible "to derive hierarchical descriptions of intonation from sequences of contours" (Dutoit 1997:142).

¹⁰ This structure was first developed by H. Klinghardt at the beginning of the 20th century (quoted in Hirst and DiCristo 1998:34).

¹¹ See Ladd (1996:209-210), Fox (2000:277-280), and Cruttenden (1997) for in-depth descriptions.

The British model rests mainly on auditory analysis which represents both its strength and its shortcoming. The terminology of the system is rather impressionistic, possibly because its originators did not have the technology to analyse F0 contours in detail (Taylor 2000:239). In fact, it is often difficult to relate tonetic or interlinear¹² transcriptions to F0 traces (Grice and Bauman 2007:18). This poses serious problems if one is to attempt an analysis of an as-yet undescribed intonational system, such as Jaminjung.

2.2.3.3 *Brazil and discourse intonation*

Brazil's Discourse Intonation (hereafter DI) theory is based on his examination of the relationships of intonation patterns to the structuring of information in spoken English. His description of intonation is based on the users' - not the linguists' - perspective, giving context and perception central roles. In DI, stress, tone, and pitch height are related to categories of meaning. DI does not aim to provide a link to categories of grammar, or between attitudinal and linguistic aspects of intonation (Brazil 1997). In this model, speech is divided into tone units which have either one or two prominences. These tone units may or may not be separated by pauses, and each unit must contain at least one tone. DI organizes intonation events in four systems: prominence, tone, key and termination. Prominence refers to the changes in stress, pitch or emphasis that allow for further meaning to be inferred in what is being spoken. Tones are pitch movements distinguished by their particular direction or contour. They are divided into five types: Falling, Rising, Rise-Fall, Fall-Rise, and Level. These tones either proclaim some new information, or refer to a common experience already shared by the speakers. The last prominent syllable in a tone unit is called the tonic syllable. Key refers to the pitch at the beginning of an utterance (relative to the last pitch on the previous utterance) and has three levels: High, Mid, and Low. The High Key (in English) conveys the meaning that something has

¹² Both terms are used to describe the transcription systems used by the proponents of the British school. It is also referred to as the 'tadpole' transcription, because of the shapes of the dots and lines used as symbols.

happened contrary to one's expectations; the Mid Key signals that a person has completed a series of tasks simultaneously; and finally the Low Key implies that an action resulted in something that was already expected. Termination also implies a low, middle, or high pitch-level choice, made by speakers at the beginning or end of a tone unit. The model provides a good framework for description, but is hampered by its theoretical underspecification.

2.2.3.4 *The Dutch school*

The Dutch School is based on the work of the researchers at the Institute for Perception Research in Eindhoven. Their approach is not based on finding categories but on identifying perceptually relevant features, that is to say, on identifying F0 movements which are salient for the listeners (Fox 2000:285). Phonologically, the smallest unit in this framework is a pitch movement, the smallest unit of perceptual analysis (t' Hart, Collier and Cohen (1990:72) referred to in Fox 2000:286). The methodology involves the stylization of F0 contours, and amounts to simplifying the F0 curve and keeping only what is relevant for speech communication, hence to a standardization process which leads to the identification of some basic patterns.

2.2.3.5 *Intonational phonology*

This model is presented in some length because of its prevalence in intonation research. Even though the study of Jaminjung presented here does not employ this framework, it is useful to understand its premises, principally because so much of its terminology has become widespread, and is useful if comparison are to be made between the phenomena described for Jaminjung and those of other languages, as so much current research makes use of the model.

Pierrehumbert proposed the first Autosegmental Metrical model (AM model hereafter) for American English in 1980. Although other models of intonation are based on the assumption that intonation can be analysed in phonological terms, the term *intonational phonology* is normally associated with the AM model. In it a distinction is made for the first time

between the phonological level and its phonetic implementation (Rossi 2000:26-29). It is 'autosegmental' as it does not present intonational phonemes in terms of their features but as sequences of elements represented on separate or autonomous levels (or tiers), linked with the segmental tier by association lines. The pitch contour thus consists of three components: pitch-accents, breaks, and boundary tones. Pitch accents and boundary tones are composed of one or more tones, H or L, which are phonological abstractions of a high and a low pitch level, respectively. Pitch accent tones and boundary (or edge) tones are seen as tone targets, while the pitch movements between separate intonation events are the result of interpolation.

Intonation is described in a grammar consisting of an inventory of contours which are organized by rules into more complex structures, and can, in turn, be used to define all the permissible tunes in a language. Figure 2-7 (from Pierrehumbert 1980, in Ladd 1996:80) illustrates such a grammar for English, with the possible pitch accents on the left, the potential boundary tones of the intermediate phrase in the middle section and the potential boundary tone of the intonation phrases on the right. The arrows indicate the allowed associations between each of the components. The grammar states that the tunes are made up of at least one pitch accent, followed by an obligatory phrase tone and boundary tone (Ladd 1996:81). The actual realization of these phonological abstractions as concrete pitch contours is referred to as phonetic implementation (see Ladd 1996, and Pierrehumbert and Hirschberg 1986, for a full description of the model).

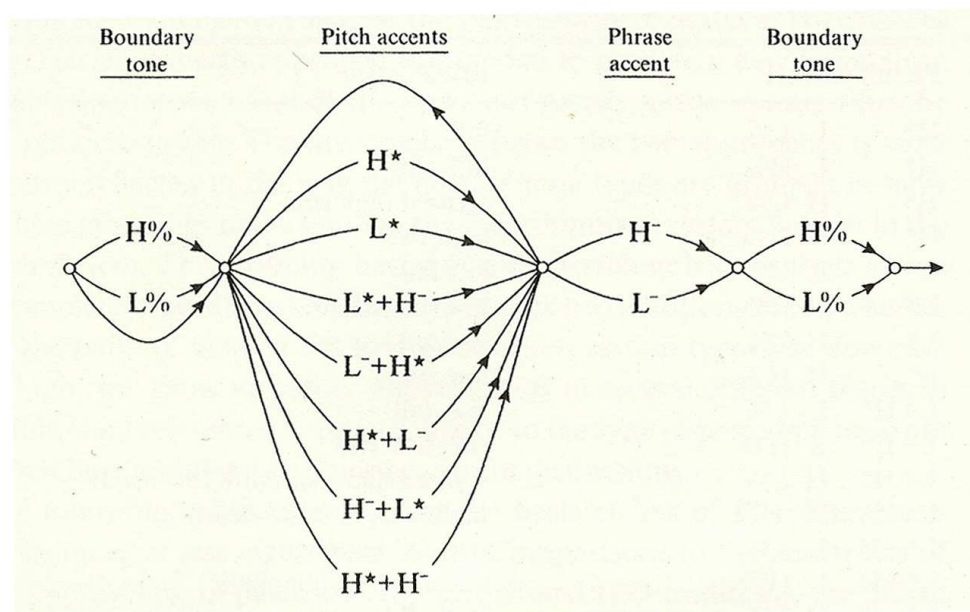


Figure 2-7 The grammar of English intonation patterns (Pierrehumbert 1980, figure from Ladd 1996:80) in which (A) Prosodic phrase boundaries coincide with boundaries of intonational tunes, and (B) Every intermediate phrase contains at least one accent.

The counterpart of this model is the ToBI¹³ (tones and break indices system) system for prosodic transcription (Silverman et al. 1992). It was devised to develop a common annotation system for large shared corpora for use in prosodic research and in speech technology. It is not a phonetic transcription system, inasmuch as the ToBI system needs to be adapted for each language, although the same annotation criteria apply. It is a tier-based annotation with four tiers: a tone tier, where the intonation of an utterance is decomposed into underlying H (high) and L (low) pitch targets, signaling two broad functions, phrasing and accentuation; an orthographic tier, consisting of an orthographic transcript; a break index tier, where the utterance is analyzed in terms of its prosodic constituency from the level of the word to the intonational phrase; and a miscellaneous tier, consisting of various relevant comments.

Figure 2-8 below presents the ToBI tones and breaks inventory (adapted from Clark et al. 2007:364-370). The model predicts that

¹³ See ToBI training materials (Beckman et al.) <http://www.ling.ohio-state.edu/tobi> and Ladd (1996:94-98)

intonation can be described by patterns of H and L tones. Pitch accents are either a single or a double tone and are indicated in the annotation by a starred tone (*). This accent is directly associated with the accented syllable. The inventory of pitch accents thus is H*, L*, H* + L, H + L*, L + H* and L* + H. The boundary tones are positioned at phrase boundaries, and are indicated by a percentage sign (%). Finally, phrase tones are used to show the contour after the last pitch accent to the end of that phrase and are marked with a dash (-). The break index determines the strength of the boundaries between the units of the utterance, expressed on a scale of 0 to 4. Figure 2.9 shows an annotation with the ToBI system.

<i>Pitch accent tones</i>	<i>H*</i>	<i>peak accent</i>
	L*	low accent
	L+H*	rising peak accent (contrastive)
	L*+H	scooped accent
	H+!H*	downstepped high
Boundary tones	L-L%	final low (Am Eng. Declarative contour)
	L-H%	continuation rise
	H-H%	yes-no question
Break indices	0:	clitics
	1	word boundaries
	2	short pause
	3	intermediate intonation phrase
	4	full intonation phrase/final boundary

Figure 2-8 Tone inventory for English, with the ToBI system of annotation inventory (adapted from Clark, et al. 2007:364-370).

EXAMPLE <<jam1>>: Will you have marmalade, or jam?

L* H- L* H-H%

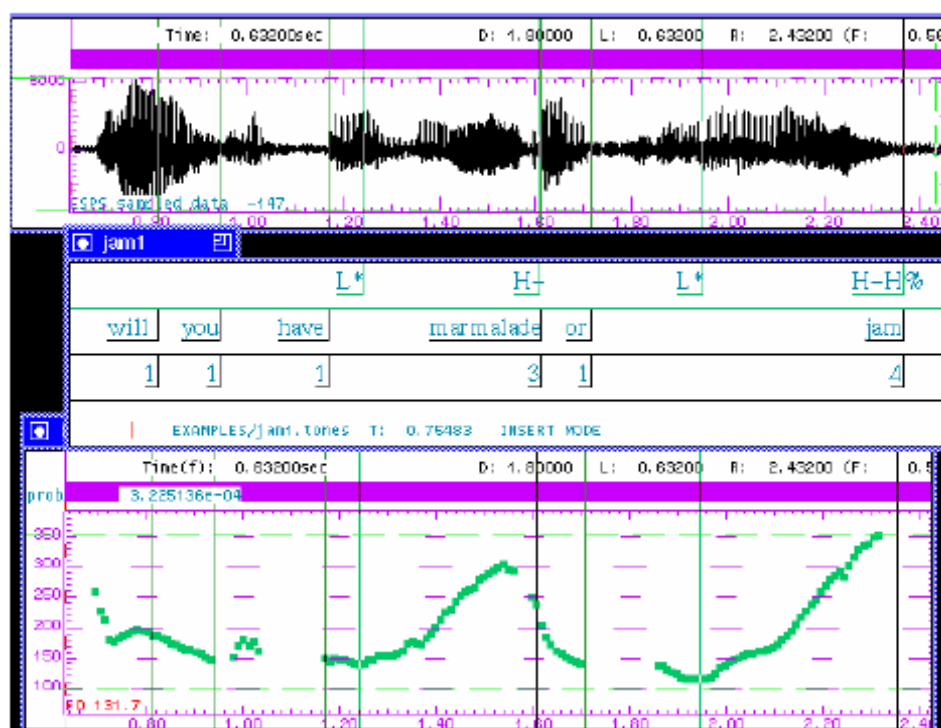


Figure 2-9 An example of the alignment of the pitch accents and boundary tones in English, annotated with the ToBI system. The upper part shows the speech signal, the middle part the annotation tiers (tone, syllable and break) and the lower part of the figure shows the associated pitch track (from Beckman and Ayers, 1997).

The AM model, with the ToBI annotation system, and its adaptation for other languages, is widely used. Ladd even proposes that it could provide “a genuinely universal descriptive framework” for prosody (Ladd 1996:4), but there are some concerns about the model, most importantly:

- There is little evidence of an emerging consensus on pitch accent meanings;
- The importance of relative height is overlooked in the model. (Cruttenden 1997:71-72). In fact, any variations which cannot be viewed as categorically distinctive, such as pitch range, are not taken into account, even if these also follow regular patterns (Carmichael 2003);
- There is low inter-annotator agreement on pitch accent types in ToBI (see Dilley et al. 2006 and Calhoun et al. 2005);
- The nature of the nuclear tone is underspecified. Nuclear accents and prenuclear accents are not distinguished by any means other than the

nuclear accent being followed by a boundary tone (see Fox 2000:297-307 for a discussion of these issues);

- With the ToBI annotation, prosodic (H, L) and functional representations (- * %) are concatenated, i.e. the latter symbols convey aspects of prosodic structure which expresses the organizational function of intonation, rather than its form (Hirst 2005:338).

2.2.3.6 Prosodic phonology and the Prosodic Hierarchy

A theoretical framework closely related to AM theory is the *Prosodic Hierarchy* theory (see Shattuck-Hufnagel and Turk 1996 for a review), which deals with the structure of prosodic units.

Prosodic hierarchy theory is concerned with the description of the prosodic constituents organized in a hierarchical structure, represented as branching trees with categorical nodes such as syllables (σ), feet (Σ), phonological words (ω), intonational phrases (φ), and utterances (U)¹⁴. The need for separate prosodic and syntactic hierarchies stems from the realization that the domains of phonological rules are not governed by syntax, and that mismatches between phonological rule domains and syntactic constituents are frequent (Inkelas and Zec 1995:537). It can be seen as the interface that connects morphology, syntax and semantics to phonology. It is considered universal, while the specific prosodic evidence for each level and the features that define their boundaries vary from language to language. Many versions of the hierarchy have been proposed¹⁵, the one in Figure 2-10 is suggested by Selkirk¹⁶ (1986:384).

¹⁴ See Ladd (1996:233-261) and Fox (2000: 335-355) for a discussion.

¹⁵ For examples, see Selkirk (1978, 1984, 1986), Nespor and Vogel (1986), Hayes (1989).

¹⁶ Pierrehumbert originally proposed the Intonational Phrase as the only constituent for American English, but in 1986, with Beckman, she suggested adding another constituent, the Intermediate Intonational Phrase.

Utterance (U)	(-----)
Intonational Phrase (I)	(-----) (----)
Phonological Phrase (φ)	(-----) (-----) (----)
Prosodic Word (ω)	(-----) (----) (-----) (----)
Foot (F)	(.) (..) (.) (.) (..) (.) (.) (..)
Syllable (σ)	00 () () () 00 () () 00

Figure 2-10 The hierarchy of units proposed by Selkirk (1986) in the Prosodic Phonology theory.

The relation between the components is defined by Selkirk (1984) in the Strict Layer Hypothesis (SLH): each prosodic constituent immediately dominates a prosodic constituent of a category level lower than itself, and a unit of a given level is exhaustively contained in the superordinate unit of which it is a part – prosodic recursion is not allowed (Nespor and Vogel 1986:7).

Although the idea that prosodic constituents are organized hierarchically is widely accepted, there are some difficulties. Criticisms relate to the status of the categories (the phonetic definition of different categories is not always clear), to the necessity of adding some new categories to account for the various processes observed in languages; as well as to the issue of recursivity. For instance, in a cross-linguistic study of phenomena associated with the word-level, Schiering et al. (2007:68) conclude that Vietnamese lacks a phonological word domain while Chukchi has more levels between the foot and the intonation phrase. The authors thus admonish that it is too early to call the Strict Layer hypothesis and prosodic hierarchy universal. Ladd (1996:237-251), for his part, finds that the Strict Layer hypothesis is incompatible with coordination as illustrated in the structures in Figure 2-11 which form recursive structures.

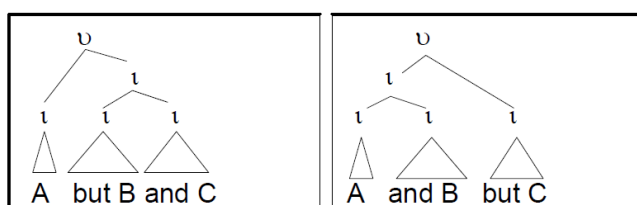


Figure 2-11 Ladd's examples of coordinated recursive structures (adapted from Ladd 1996:242).

Another issue that is widely discussed is whether and how prosodic constituents correspond to the units of syntactic analysis, often referred to as the ‘mapping’ between the two structures. In the AM and prosodic hierarchy theories, it is assumed that prosodic categories map to a certain extent to syntactic categories. The first three levels of the hierarchy (syllable, foot, word) are generally considered to be within the domain of word-level phonology, whereas the top levels cover domains beyond that of the word. The top levels of the hierarchy intersect with syntax, indeed syntactic information is considered crucial for the formation of these prosodic constituents. Selkirk (1986) goes on to propose to categorise the languages of the world according to how syntactic constituents attach either to the left or right edges of prosodic constituents. Much effort has been devoted to mapping the relationships between the prosodic and syntactic hierarchies, and to explain their mismatches (see Shattuck-Hufnagel 2000, Selkirk 1978, 1984, Nespor and Vogel 1986, and the articles in Inkelas and Zec 1990). A classical example of a mismatch is shown below. The bracketing in the first sentence shows the syntactic constituents, in the second, the prosodic constituents (adapted from Kreidler 2001:154).

- (2) [This is [[[NP the dog that chased] [NP the cat that bit] [NP the rat that was running away]]]].
 [This is the dog] [that chased the cat] [that bit the rat that was running away].

The lack of a direct correspondence between prosodic units and syntactic constituents is further illustrated by clitics: phonological words are summarily equated with grammatical words, whereas clitics stand as separate grammatical words, although prosodically they attach themselves to stressed words, either preceding or following them, and, in so doing, lose their independence as phonological words. Among the English clitics are the articles *a(n)* and *the*, which are bound to a following host noun. The same observation can be made for phonological phrases that are equated to syntactic phrases, such as NPs, and for intonational phrases which are linked to syntactic clauses.

Other models attempt to account for the relations between syntactic and prosodic structures. Martin (2000) proposes that prosodic structure is related to syntactic structure via a set of constraints, one of which is the ‘syntactic clash condition’, which forbids two minimal prosodic units (i.e. stress groups) to be reunited in the same higher prosodic group if they are dominated in the syntactic tree by two distinct nodes. In Optimality Theory, the alignment and ‘reranking’ constraints account for the interaction between the prosodic and syntactic structures (McCarthy and Prince 1993). More recently, Wagner (2005) has proposed that prosodic structure is built cyclically.

For Jaminjung, I will follow Fox (2000:330-331) in assuming that speech is organized into a structure which “in its most common linguistic usage (it) refers to the organization of a sentence or other unit of language in terms of relationships between its parts. Once this is assumed, the task is to “determine its nature and examine the role and place of prosodic features within it”. Details on how prosodic units are viewed in the PENTA model adopted in this study are given in §3.1.

2.2.3.7 INTSINT

INTSINT (International Transcription System for Intonation) is not strictly speaking a theoretical model but a transcription system for intonation, developed at the CNRS Laboratoire Parole et Langue in Aix-en-Provence (see Hirst and DiCristo 1998:14-18). In fact, its creators wished to provide an equivalent of the IPA for intonation, applicable to all languages. It is also intended to be theory neutral, however, this is obviously very difficult to achieve in practice.

In INTSINT, intonation is described as a sequence of labels representing target points. They can be defined either in reference to the speaker’s pitch range, Top (T), Mid (M) or Bottom (B), or relative to a previous target Higher (H), Same (S) or Lower (L). Accents are indicated with Upstepped (S) or Downstepped (D). The system has been used in studies of English (Hirst), Spanish (Alcoba and Murillo), European Portuguese (Cruz-Ferreira), Brazilian Portuguese (Moraes), French (Di Cristo), Romanian (Dascaľu-Jinga), Russian (Svetozarova), Bulgarian

(Misheva and Nikov), Moroccan Arabic (Benkirane), and Japanese (Abe) (in Di Cristo and Hirst 1998). An example of an English utterance (from Hirst 2005:343) annotated with this system is shown in Figure 2-12. In example 3, the tonal labels are assigned to the phonological constituents. The phonetic interpretation of the INTSINT tonal segments can be carried out using two speaker dependent (or even utterance dependent) parameters: key, which establishes an absolute point of reference defined by a fundamental frequency value (in Hertz); and range, which determines the interval between the highest and lowest pitches of the utterance. These can be converted into a sequence of tonal target points (mid panel) that can be used for a speech synthesis system, which in turn can be used to generate a continuous F0 curve by means of the F0 modeling algorithm, MOMEL (modeling melody). INTSINT is thus a mixture of a phonetic and a phonological model: from the low-level phonetic analysis a phonological description is derived.

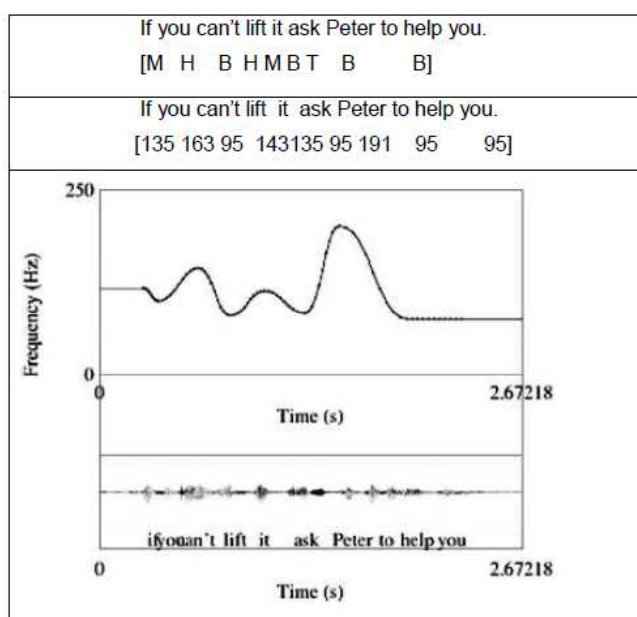


Figure 2-12 INTSINT annotation with an English sentence (from Hirst 2005:243).

2.2.3.8 Tilt

The Tilt Model was developed by Paul Taylor at the Centre for Speech Technology Research at the University of Edinburgh. It does not purport to model human behaviour but is strictly an engineering model, “designed to facilitate automatic analysis and synthesis” (Taylor 2000:4). It

is a phonetic model, attempting to form a representation of a natural F0 contour automatically from the raw signal and is a derivation of the AM model. Taylor's idea is that as no objective test can be developed to find the pitch accent categories of a particular language, it is preferable to bypass the definition of discrete numbers of categories. Accent and boundary tones are represented as a linear sequence of *intonational events* which constitute the basic intonational units. An event is characterized by continuous parameters representing amplitude, duration, and tilt (a measure of the shape of the event).

2.2.3.9 KIM

Kim (Kiel Intonation Model) is an intonation model for German, developed by Kohler and his team at the University of Kiel (Kohler 1991). It is a configurational, functional model that aims to integrate the syntactic, semantic, pragmatic and expressive functions of intonation, incorporating stress and intonation, timing and articulatory reduction. A system of distinctive prosodic features is used to specify abstract symbolic, phonological categories in these domains which then enter into sets of ordered symbolic rules. The prosodic features are attributed to phonological units, which are either segmental (vowels and consonants) or non-segmental (morphological and phrase boundaries).

2.2.3.10 Fujisaki

This model has been developed by Fujisaki since 1983 for Japanese intonation¹⁷. It is an acoustic-phonetic model which aims to link surface F0 directly to muscle commands, in effect, to simulate human production mechanisms. It is based on the idea that the production of the surface F0 contours is the result of accent and phrase commands superimposed on an original declining baseline. This has proven very fairly successful, particularly with synthetic speech; the model is often cited as a good example of the superpositional type (§2.2.2).

¹⁷ See also Möbius (1993) for German.

The models presented so far vary greatly in their assumptions and implementations. This is not surprising, considering that they have been developed to fulfill very different goals. The British school, for instance, evolved from an older tradition which was concerned with didactic uses. The AM model aimed at providing a theory of intonation that worked cross-linguistically. The Tilt model was designed solely for engineering purposes (Taylor 2000:229) and the Fujisaki model aims at creating a model that replicates known biological production mechanisms. The main difficulty resides in defining the level of ‘abstract representation’ between the phonetic (the production of the pitch variations) and the functional levels. However, it is not unreasonable to expect a theoretical model to serve all aims adequately.

The most obvious choice for the description of the intonation of Jaminjung could appear, at first glance, to be the AM model as it is so widely used in intonational research. There are some difficulties in using this model, some of which have been listed above (§2.2.3.5). They can be summed up by saying that the model assumes that intonational phonemes can be identified without direct recourse either to their phonetic realization, or the function they serve to express. Methodologically this is not feasible for Jaminjung. Firstly, it is unsound to assume that the intonational phonemes of all languages will resemble those of American English, on which the model originates, and therefore that all phonetic information can be bypassed. Secondly, describing components of intonation and their organization primarily by their form is only the first step in developing an intonational phonology, as recognized by Pierrehumbert herself (2000:59).

The following section will present the PENTA model, which is the theoretical framework adopted for this research. The model provides a clear explanation of prosodic phenomena linking phonetics to semantics. It is a powerful descriptive tool, particularly useful in the context of an as-yet undescribed prosodic system, such as that of Jaminjung, by anchoring the description onto prosodic functions.

2.3 The PENTA model

This description will use the Parallel Encoding and Target Approximation model (PENTA), developed by Yi Xu (2005), as a theoretical basis. This model is based on two assumptions: that speech conveys communicative meanings, so intonational components should be defined in terms of function rather than form; and that speech is produced by an articulatory system that has various biophysical properties which impose constraints on the way surface acoustic forms are generated. Very summarily, the model states that communicative functions are transmitted in parallel, through encoding schemes that are either language specific or universal, via a limited number of parameters that may be considered as the phonetic primitives. Four primitives can be recognized for speech melody: local pitch targets, pitch range, articulatory strength and duration. Xu emphasizes that

...the PENTA model by itself does not stipulate the properties of the encoding schemes. It only provides a mechanistic framework for the encoding schemes to be implementable. The detailed properties of the encoding schemes (...) can be discovered only through empirical investigations in which potential contributors to surface F0 contours are systematically controlled' (Xu 2005:246).

The encoding schemes define the targets that are only approximated during speech, resulting in surface F0 realisations. The model is presented in a schematic diagram in Figure 2-13.

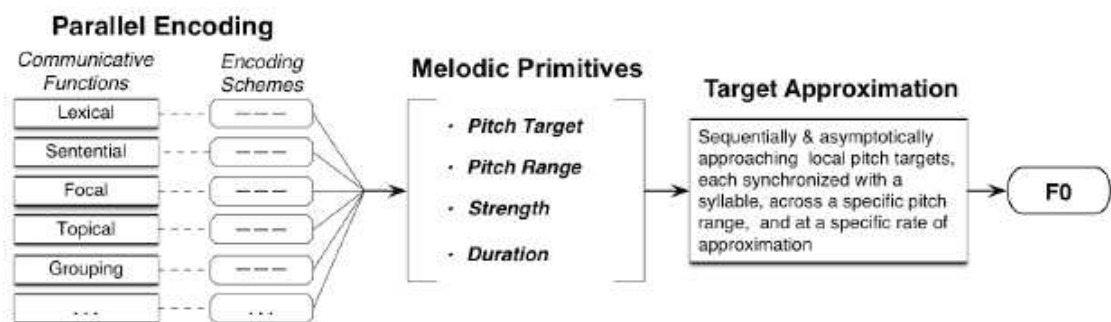


Figure 2-13 Intonation as articulatorily implemented encoding schemes defined and organized by communicative functions (from Xu 2005:243).

2.3.1 Parallel encoding

The specific F0 values produced for a given contour are the results of the need to encode multiple parallel communicative functions. Functions¹⁸ are thus transmitted concurrently through encoding schemes (hence Parallel Encoding) that assign values to the melodic primitives over specific temporal domains. These schemes are used as control parameters for the target approximation process and eventually lead to the realization of the surface F0. Figure 2-14 (adapted from Xu 2005:243) shows the possible values of the melodic primitives: local target, pitch range, articulatory strength and duration. Xu suggests a notation to distinguish each of the ‘primitives’ by using [], underline, boldface and italic, respectively.

Local target		[high], [low], [rise], [fall], [mid]
Pitch range	Height	<u>high</u> , <u>low</u> , <u>mid</u>
	Span	<u>wide</u> , <u>narrow</u> , <u>normal</u>
Strength		strong , weak , normal
Duration		<i>long</i> , <i>short</i> , <i>normal</i>

Figure 2-14 Possible symbolic values of the melodic primitives pitch target, pitch range, strength and duration (adapted from Xu 2005:243).

Figure 2-15 shows a hypothetical graphic representation of how the PENTA model can be applied to analyse the surface F0 contour of an English sentence (used in the study on English declaratives, Xu and Xu 2005). The thick black line shows the sentence spoken with a narrow focus on *mimic* and the thin black line shows the same F0 curve without contrastive focus. The following encoding schemes are shown: lexical stress, sentence type and focus. The pitch targets are determined by the first two, while pitch range is assigned by the latter. The representations of the local targets and pitch ranges correspond to specific numerical values. The straight dotted lines indicate the height and span of the local pitch targets. The block arrows show the pitch range adjustments provoked by focusing.

¹⁸ Amongst which are the marking of lexical tone, lexical stress, focus, sentence type (question, statement), new topic-turn shift, grouping, demarcation, relation marking, emotion, etc.

The unfilled block arrows indicate a wide pitch range and the filled block arrows indicate a narrow and low pitch range.

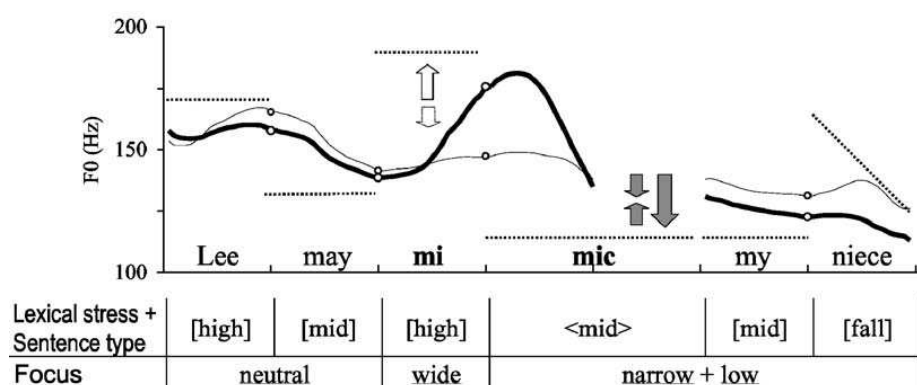


Figure 2-15 Decomposition of the F0 contours of “Lee may mimic my niece” according to the PENTA model (from Xu and Xu 2005).

2.3.2 Pitch target approximation

In the model, the most basic intonational units are discrete pitch targets which are synchronized with syllables. Pitch targets are simple linear functions that can be static, for instance [High], [Mid], and [Low], or dynamic [Rising] and [Falling], the inventory of which, for each language, should be empirically determined (Xu 2005, 2007, Xu and Xu 2005, Sun and Xu 2002, Xu and Wang 2001).

Targets are said to be underlying as the surface F0 is the result of their articulatory approximation by the speakers. They may often not be fully realized, that is, pitch targets are the intended goals rather than the realized F0 contours. A speaker thus produces F0 which approaches the target from the onset of the syllable and continues to approach it throughout the syllable. Figure 2-16 illustrates the Target Approximation model. Syllable boundaries are indicated by vertical lines. The underlying pitch targets are shown with dashed lines. The F0 contour results from asymptotic approximation of the pitch, represented here with the thick curve.

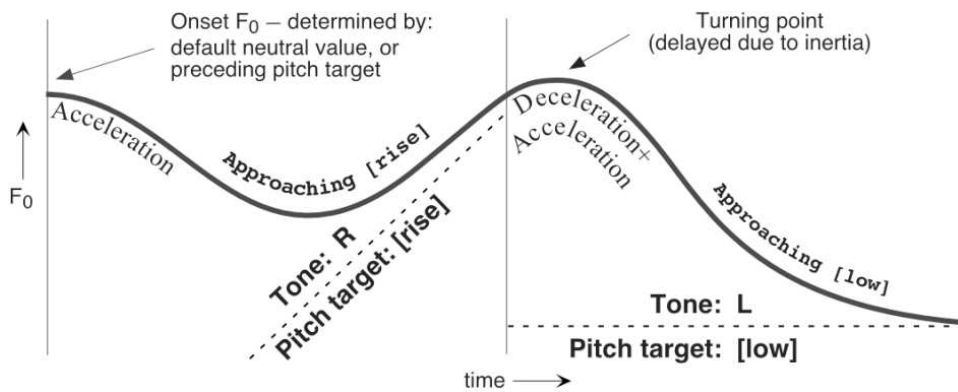


Figure 2-16 Illustration of the TA model (from Xu 2007:2).

2.3.3 Evaluating pitch targets

In the model, the syllable is treated as a synchronization unit, predicting that each syllable has its own specified target and that speech production unfolds one syllable at a time. Thus, the articulatory movement starts at the syllable onset and approaches the specified target at the end of the syllable; a good portion of the surface contours can be seen as transitions towards the target (Sun and Xu 2002). Accordingly, the best approximation of the target is reached in the later portion of the syllable. Examining the final F0 contour of the syllable seems the most obvious means to identify the property of the underlying pitch targets. However, it was found that the velocity of the F0 movement provides even more reliable evidence for their identification (Gauthier et al. 2007; Xu and Liu 2006, 2007)¹⁹. In this context, the *velocity of F0* is defined as the instantaneous rate of change of F0 which is mathematically its derivative (Figure 2-17). In non-mathematical terms, a derivative can be described as a measure of how much a quantity is changing at a given point.

¹⁹ See Xu and Liu (2006:18-21) for a theoretical presentation of the model, and Chen and Xu (2006), Liu and Xu (2005), Wu (to appear), for final velocity.

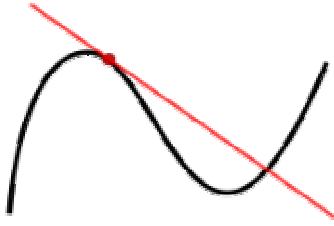


Figure 2-17 An illustration of a derivative. The function is drawn in black, and a tangent line to that function is drawn in red (Wikipedia, ‘derivative’, accessed 01/03/2010).

In this research, F0 velocity is computed by subtracting two adjacent F0 values, over time, as shown in the following equation:

$$F0_j' = (F0_{j+1} - F0_j) / (t_{j+1} - t_j)$$

where $F0_j'$ is the velocity value of F0, and t_j is the time of the F0.

(Xu and Liu 2006:18)

Figure 2-18 shows the prototypical velocity profiles (not to be confused with F0 pitch curves) of Mandarin tones obtained in a simulation in Gauthier et al. (2007). A black vertical line is drawn to indicate the points²⁰ (30 ms before syllable offset) where the *final velocity* measurements would be taken in our dataset.

The static tones, [High] (blue) and [Low] (yellow), mostly increase their speed from 0, reaching peak velocity (positive or negative) around the centre of the syllable and finally slowing down toward the initial speed of 0 near the end of the syllable. The dynamic tones, [Rise] (green) and [Fall] (red), increase their speed from 0 at syllable onset toward a negative/positive value, and quickly cross the zero speed line. In each case they change direction, and continue until a high velocity is reached near the end of the syllable. The simulation thus reveals that a measure taken before the syllable offset can be a reliable indication of the underlying target. In light of this, in the Jaminjung datasets, the measurements for final velocity (expressed in semitones per second) are taken at 30 ms before syllable offset. However, because the datasets are composed of spontaneous speech

²⁰ In that figure the values are the time-normalized date points, not ms. The actual measurement point, i.e., 30 ms before the conventional syllable offset is calculated this way: assuming mean syllable duration = 200 ms, $30/200 = 0.15$, $30 \times 0.15 = 4.5$. Thus the vertical black line is shown at 25.5 on the x-axis.

with no control over the segmental content, it would be unwise to rely solely on this measure to define the targets. It is matched with a visual inspection of the velocity curves and of the pitch tracks, as well as auditory impression.

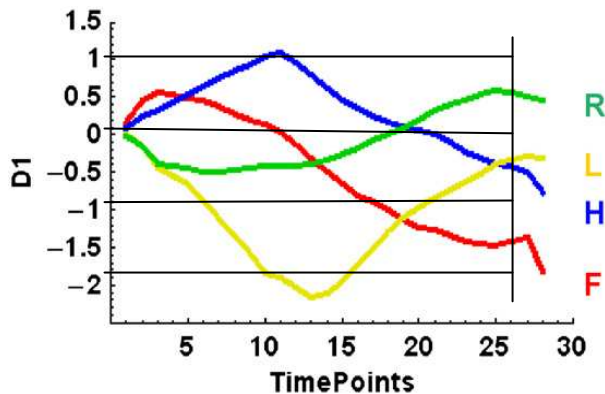


Figure 2-18 The profiles of Mandarin tones, obtained in a simulation (adapted from Gauthier et al. 2006:98). The black vertical line shows the point that is measured in the Jaminjung dataset.

The measurements help in defining the pitch targets, which are described in more detail here. Note that these descriptions apply to the targets in Mandarin and English, based on work by Xu (2005), Xu and Liu (2006), and Gauthier et al. (2007).

A High target (see velocity curve A in Figure 2-19 below, in which the velocity is shown by the thick blue line and the pitch by a thinner red line) reaches a peak in F0 before the end of the syllable and this is reflected in the velocity curve. As the velocity tends towards zero after reaching the peak, it often becomes negative by the end of a syllable with a High target.

The reverse is true of Low target (D), which reaches a trough early in the syllable, its velocity then tends towards zero or it can even have a low positive value.

A further static target is posited, the Mid target. It was found that the targeted F0 for the mid target is about halfway between the highest and lowest F0 values of the full targets. A further characteristic of the Mid target is that it is implemented with weaker strength, thus accounting for the decline in F0 encountered in a succession of syllables with Mid targets (Chen and Xu 2006:65-70). Indeed studies of English intonation suggest that an unstressed syllable may be represented by a static target [Mid], while a stressed syllable may have either a static or dynamic target depending on a number of lexical and postlexical factors (Xu and Xu 2005).

In the dynamic targets Rise (B) and Fall (C), the unidirectional movement reaches a desired velocity that is not zero. This means that the velocity of a dynamic tone should end at either a positive or a negative value when its execution terminates (Xu and Liu 2006). These high (positive or negative) values near the end of the syllable suggest that the high velocity itself is the final goal of the [Rise] and [Fall] targets (Gauthier et al. 2006:98). The targets are illustrated in a stylized fashion in Figure 2-19. Velocity curves from syllables in Jaminjung resembling the stylized targets are shown in Figure 2-20 with the corresponding pitch track of the IU from which the illustrative syllables are extracted.

Although final velocity is a good indicator of the underlying pitch target, it cannot be used without calling upon the pitch tracks and auditory perception to determine the pitch targets. This is particularly evident with [high] and [low] targets, which can both have velocity measurements between -5st/s and 5st/s in Jaminjung. Further, the pitch tracks of [mid] and [low] targets may resemble each other, as successions of [mid] targets often decline gently over the course of the word, but they are easily distinguished with the help of the final velocity curves and auditory perception. In fact, the combination of the final velocity curves, the pitch tracks and auditory perception is needed in order to evaluate the pitch targets. The examination of the Jaminjung data suggests the following ranges of final velocity for each target, these are indicative only, as deviations from these are observed in the data:

[rise]	10st/s and higher
[fall]	-10st/s and lower
[high]	between 0 and 5st/s
[low]	between 0 and -5st/s
[mid]	around 0st/s

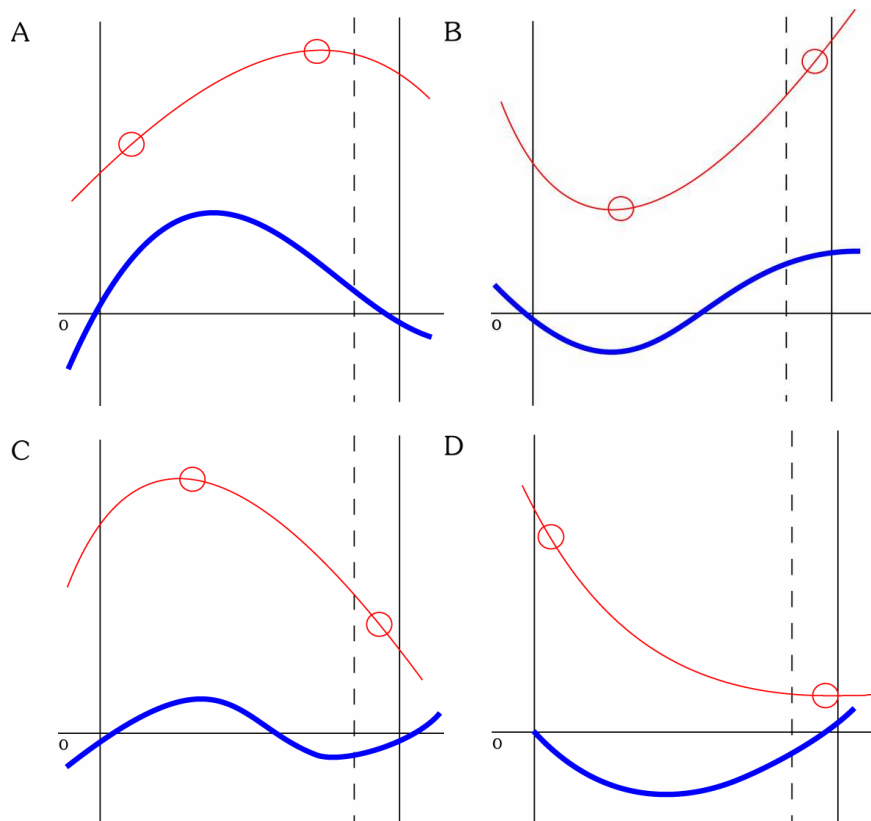
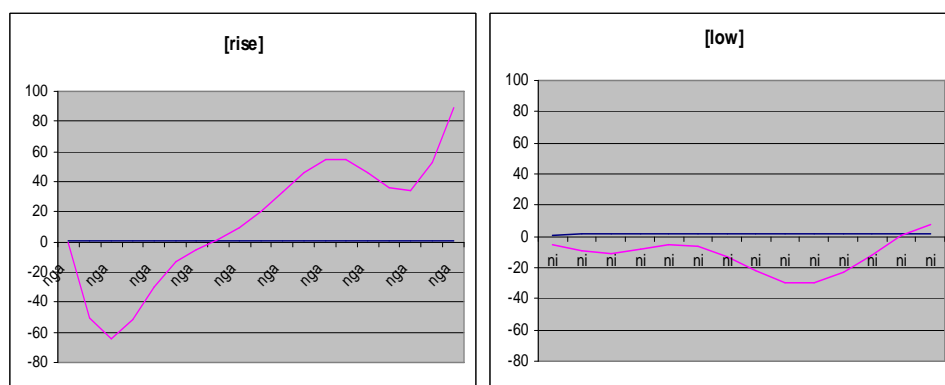
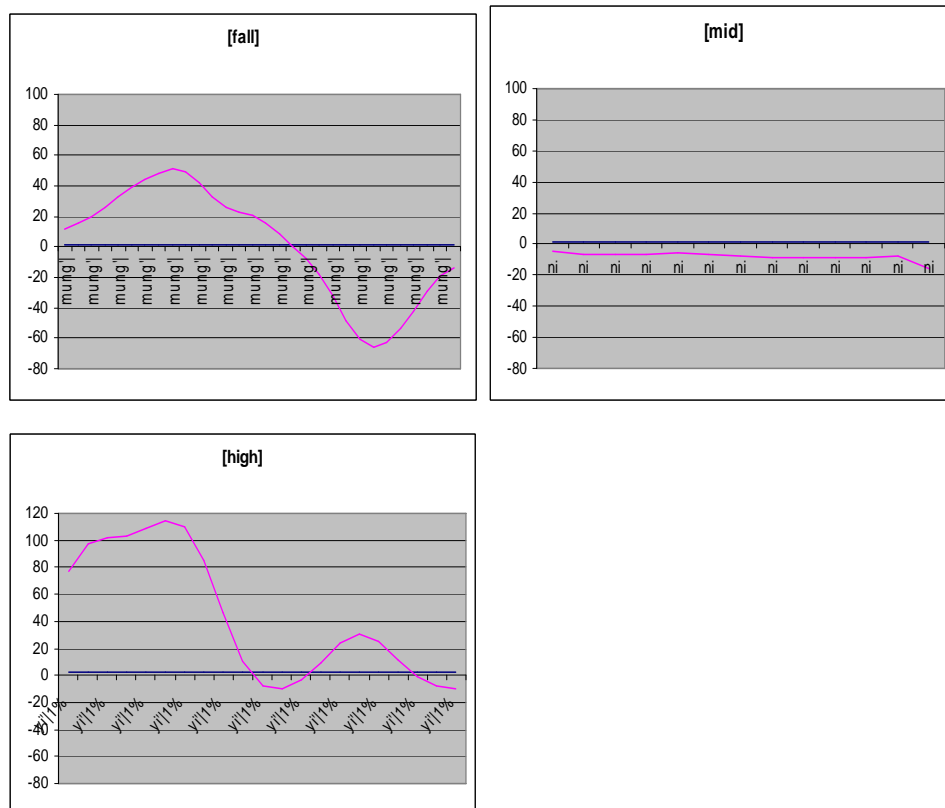


Figure 2-19 Stylised representation of the pitch targets. A [high], B [rise], C [fall], D [low]. The thinner line represents the pitch, the thicker line represents the velocity. The dots on the pitch line indicate the highest and lowest F0 points during the syllable.





(3)

ngarrung-ni mung ganignayim mayi
 woman-ERG/INSTR watch 3sg:3sg-see-PRS person

‘The woman is looking at the person.’

[JM:CS06_a036_01]

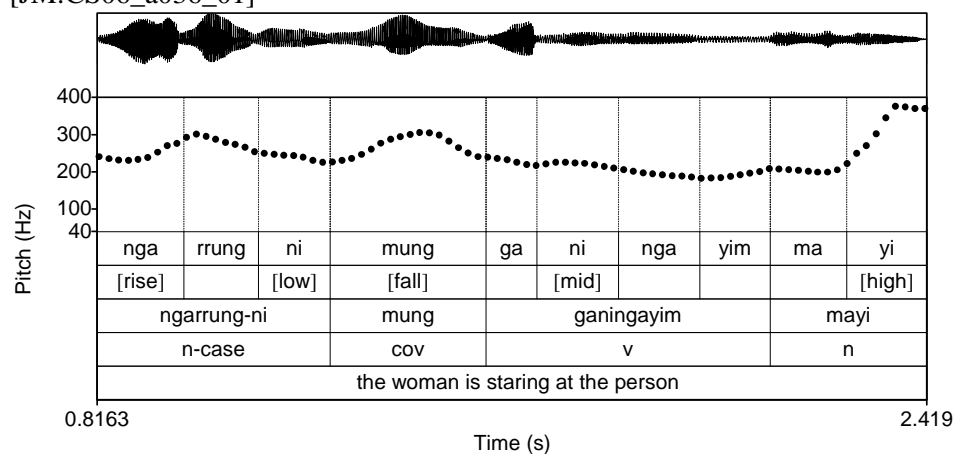


Figure 2-20 Examples of velocity curves for the targets in Jaminjung, over one syllable; and the pitch track of the IU from which they are extracted.

Finally, Xu and his colleagues (2009) describe a ‘carryover effect’ to explain the influence of the preceding target over a new target. That is, during each target approximation, the state of articulation depends not only

on the discrepancy between the current state and the target but also on the final velocity and acceleration of the preceding syllable. This effect would gradually decrease over time. Thus the state of articulation, as defined by F0 height, velocity, and acceleration, is transferred from one syllable to the next at the syllable boundary. Such transfer of articulatory state is assumed to explain the phenomenon of F0 peak delay in both tone and non-tone languages (Prom-On, Xu and Thipakorn 2009).

2.4 Methodological considerations

The description of the intonational patterns in a given language is often (e.g. according to the AM model) based primarily on the description of its pitch patterns, by determining which portions are significant, associating these with the segmental content of the utterances. Only as a final step are the meanings of the patterns sought. In our estimation, this methodology is not optimal. Firstly, it dissociates phonology (the labeling of the contours) from phonetics. Arguably, phonology can eventually make do without phonetics, but it can only do so after much verification has been conducted to ascertain the linkage between phonetic realizations and phonological representations. In a language that is described for the first time, this step can not yet be taken.

By adopting the PENTA model for this analysis of Jaminjung's intonation, a different methodology is put forward. The analysis starts with an intonational function. Here, the three most often quoted functions of intonation²¹ are investigated: dividing the stream of speech into units of various sizes; highlighting certain parts of an utterance; and distinguishing sentence types. Hypotheses are formulated as to their prosodic encoding. In testing specific hypotheses concerning links between intonational functions and prosodic features in a dataset built from spontaneous speech, efforts must be made to avoid circularity. Consequently, datasets are constructed

²¹ Although these are the most frequently cited uses of intonation, the PENTA model does not limit the functions which intonation can be used to express (Xu, p.c).

with carefully selected utterances best fitting the criterial descriptions for each function. To prevent circularity, the following measures have been applied, where feasible:

- the units of analysis are defined according to criteria that are as self-contained as possible (prosodic criteria should be avoided);
- the annotation and analysis remain independent of each other;
- the subsets analysed contain comparable tokens, inasmuch as such is possible in unrehearsed spontaneous speech data.

All utterances are transferred into PRAAT²² textgrids. The F0 contour of the tokens is calculated and visualized with the help of Praat. The speech signal and the spectrogram are used to confirm boundaries and parse the tokens into syllables. Syllables are measured from consonant onset to vowel/consonant end, from the beginning of the stop closure or the beginning of first formants in nasals and laterals. The end of a syllable, when there is no pause, is the onset of the next syllable (closure of a following stop, etc.). In the case of a following pause, the end of the coda is counted as the end of the speaker's oral noise. When the unit ends in a creak, the unit is measured to the end of audible phonation. These markings are based on the discussion of methodological issues associated with segmental duration in prosodic research in Turk et al. (2006).

Once the utterances are segmented into syllables, measurements are effected, including duration, pitch, and velocity. Intensity is no less interesting a correlate, but it must be excluded from the analysis because of the uncontrolled nature of the recording conditions during fieldwork. A script²³ is then run through the resulting Praat textgrids. The average values of F0 and duration are the results of ten measurements over each marked syllable (or pause), including the onset and any coda consonants.

²² For the analysis of the digitized audio data, PRAAT, a programme especially developed and designed for speech analysis by P. Boersma and D. Weenink at the Phonetic Sciences Department of the University of Amsterdam, was used (for further information, see: <http://www.fon.hum.uva.nl/praat/>).

²³ Downloaded from Yi Xu's website <http://www.phon.ucl.ac.uk/home/yi/tools.html>

The following measurements are used.

Mean F0 — Average of all F0 values in a syllable, in Hz.

Excursion size — Difference between the max F0 and min F0 expressed in semitones for each syllable.

Pitch reset — Difference in Hz between mean F0 of the first syllable of a unit and the last syllable of the preceding unit.

Final velocity — Velocity is a measure of the instantaneous rates of F0 change expressed in semitones per second, taken at a point earlier than the interval offset (here 30ms). It is an indicator of the slope of the underlying target of the interval.

Duration — Time interval between the onset and offset of the syllable, in ms. Affinity index or inter onset interval — Time interval between the onset of one syllable and the onset of the subsequent syllable, expressed in ms.

The Hertz scale is linear and the semitone scale is a logarithmic transformation of the physical Hertz scale, corresponding better to our perception of pitch. An octave is the interval between two tones whose frequencies are in the proportion of 2 to 1; a semitone is 1/12 octave, or a 6% change in frequency. Figure 2-21 is an illustration of the conversion table from Hertz into semitones, with a basis of 50Hz (see Nolan 2003, for an evaluation of different pitch scales).

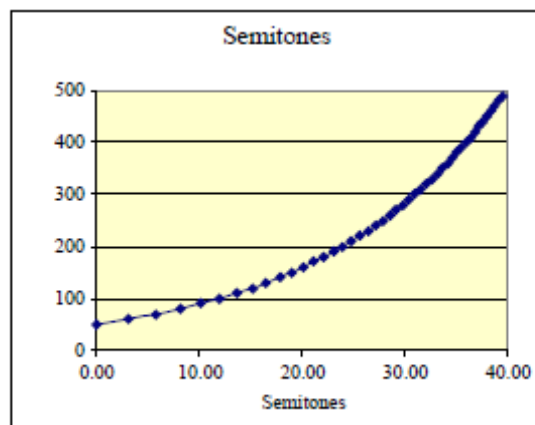


Figure 2-21 The relation between Hertz and semitones in the range 50 to 500 Hz (from Nolan 2003:771).

The raw data are then transferred into Excel and into the statistics analysis package SPSS, in order to make qualitative and quantitative

comparisons of their commonalities and differences. The descriptive statistics presented in this study are based on the raw measures. The same steps are taken at the beginning of each analysis. Firstly, the pitch range of each of the speakers in the dataset is calculated, that is to say, the interval between the lowest and highest pitch levels they reach over the duration of the text. The following formula is applied, where $RngF0_{sp}$ stands for a speaker's F0 range, and $F0_{max}$ and $F0_{min}$ stands for his/her maximum and minimum F0 respectively:

$$RngF0_{sp} = F0_{max} - F0_{min}$$

Secondly, a statistical test is conducted to check whether there are significant variations between the tokens in the datasets. In almost all situations, the variations between tokens were found not to be significant, thus showing that the tokens are fairly homogeneous and that they form a comparable set. In order to avoid unnecessary repetition, this part of the analysis is not mentioned, except when the results indicate significant variations between the tokens.

The analyses of the measurements allow us to evaluate the validity of the posited hypotheses. Only then can the specific correlates be associated with the functions they encode.

2.5 Conclusion

In this chapter, the basic notions necessary to conduct the analysis of the prosody of Jaminjung were introduced and the theoretical framework adopted in this research, the PENTA model, was presented, as well as a survey of methodological considerations. The next chapter will begin the analysis of the Jaminjung data with an examination of its prosodic constituents, under the grouping function of intonation.

Chapter 3 GROUPING

As stated by Ladd, one of the “most basic uses of intonation is to divide the stream of speech into chunks” (2001:1381). This chapter will look at how the function of grouping, or of demarcating significant units of speech, is expressed in Jaminjung. The working hypothesis is that Jaminjung has four levels of grouping cued prosodically. The potentially relevant phonetic parameters for the encoding of these levels will be tested. This chapter is organized as follows:

In section 1, matters related to prosodic phrasing are introduced: the prosodic units often described in the literature; the prosodic correlates usually associated with the encoding of prosodic units; the view of the grouping function in the PENTA model; as well as a brief survey of prosodic units described in other Australian languages.

In section 2, the units tested in this chapter are introduced together with their morpho-syntactic correlates. Before discussing criterial considerations, it is worth mentioning again that the speech data analysed in this study is from a language that has no written tradition and is not yet fully described, and that all materials issue from the transcription of natural speech. As a consequence, a basic unit of analysis is marked during the original transcription of the Jaminjung data. It is termed *intonation unit* (IU hereafter) and corresponds to a meaningful stretch of speech, usually bounded by pauses. This perceptual parsing and labeling of IUs is done before any acoustic measurements. The grammar of Jaminjung suggests four possible levels of grouping: prosodic words, phrasal constituents, intonation units, and prosodic sentences.

Finally, in section 3, the results of the analysis of the Jaminjung data are presented. The analysis will consist of a study of the phenomena associated with the posited units’ boundaries and comprises: at the right-edge: final syllable lengthening; final pitch target; phonation event (creak), pauses, and affinity index, a measure of the Inter Onset Interval, including syllable duration and following pause (if present), and at the left-edge: pitch reset and initial tone.

3.1 Prosodic units

In the PENTA model of intonation (see §2.3 for a full presentation of the model), grouping is considered as a communicative function independent of other functions (Xu 2008:17), and thus does not preclude a hierarchical organization of the constituents as it is understood in the prosodic hierarchy theory discussed earlier (§2.2.3.6). It is nevertheless possible to envisage an organizational ordering of prosodic units within the PENTA model, with units ranging from smaller to larger under the grouping function. Determining the relevant units and how they are encoded in a specific language is part of the descriptive purpose of the PENTA model. Ideally, boundary identification should be based on internal criteria, that is, on phonetic cues present at the actual boundary. For instance, Knowles (1996:96), for the SEC corpus, describes the boundary of a unit as a ‘cluster of discontinuities’. These discontinuities are related to prosodic cues such as pauses (both at the beginning and the end of a unit), anacrusis (beginning), tonal reset on unstressed syllables (beginning), or final lengthening (end) (Cruttenden 1997, DuBois et al. 1993, Di Cristo and Hirst 1998), they are discussed in some detail in §3.1.4 below. It is indeed the specific aim of this chapter to investigate what this cluster of discontinuities is made of for the prosodic units of Jaminjung. Until they are found, however, external criteria — morphological or semantic — are used to define which units to investigate. The analysis presented in this chapter is thus based on the following methodology: the potential units in a dataset consisting of narrative texts are labeled; these posited units are defined on morpho-syntactic grounds. A quantitative analysis of the prosodic cues at the left and right edges of the potential units is then conducted in order to determine which of the correlates (if any) can be associated with any of the potential units.

The structure of Jaminjung grammar suggests four potential units of grouping: two units smaller than IUs, established on morpho-syntactic criteria; *prosodic words*, and *phrasal constituents*; and one unit larger than the IU, referred to as *prosodic sentence*, established both on prosodic and syntactic criteria. They will be introduced at length in the following section, starting with intonation units.

3.1.1 Intonation Units

A basic unit of analysis is recognized in intonation research, here referred to as *intonation unit*, following Chafe (1994, see also Di Cristo and Hist 1998). A similar unit has been called *tone group* (Halliday 1967), *intonation group* (Cruttenden 1997, Brazil 1997), *tone unit* (Crystal 1969), *intonation phrase* (Beckman and Pierrehumbert 1986, Selkirk 1984), *breath group*, and *sense group* (O'Connor and Arnold 1973). Their respective definitions vary according to whether they are based on phonetic, semantic or syntactic properties.

As mentioned above, the criteria used to define the units to investigate in this chapter are mainly morpho-syntactic, but for IUs, it is difficult to abstract completely from prosodic criteria, a problem noted by many researchers (see Ladd 1986 and Swerts 1995, *inter alia*). Ladd (1986), for example, points to the risk of circularity with a general definition of IPs (equivalent to IUs in this study) as the largest phonological chunks into which utterances are divided, extending from one *phonetically definable boundary* to the next, which have a particular intonational structure and are assumed to be related to the syntactic or discourse level structure. He explains that using *boundaries* as a definitional characteristic introduces a phonetic element in the specification of the domains over which the phonological structure is specified. Hence, if something is structurally an IP then it is assumed to have boundaries, and if something has boundaries, it is assumed to be an IP. By taking into consideration the very basic nature of speech, this problem of circularity can be, if not circumvented, at least understood. IUs are units of speech, which has an intrinsic temporal dimension: speech occurs over time. Only once the temporal grounding of IUs is established is it possible to link them with other levels of analysis. Alternatively, some researchers suggest using *the contour* as definitional. DuBois and Schuetze-Coburn's (1992:229) often-quoted definition is a good illustration. An intonation unit is "a stretch of speech bounded by a single coherent intonation contour" linked to a functional, syntactic, semantic, or informational content. In practice, however, as Cruttenden (1997:29) points out, the segmentation of speech into IUs based on the

observation of contours is not always straightforward, and units are often detected primarily by their boundaries. This is also the case in the AM theory, where the intonational phrase is delimited by a boundary tone.

The works of Chafe (1980 and later) and DuBois et al. (1993) have been very influential in establishing the IU as a basic unit of speech analysis. Chafe's definition is anchored in cognition, which plays an important role in its construction. For him, people's minds contain a vast quantity of information or knowledge and only a small portion of this knowledge is called upon at any one time. This active information is stored in the short-term memory, which, according to George Miller (1956), can handle only seven items (plus or minus two) at a time. Next, Chafe associates IUs with clauses, making use of Halliday's idea (1994:295) that a clause usually coincides with one unit of information. He thus describes the relationship between IUs and clauses in the following manner: "Each clause verbalizes the idea of an event or state, and usually each intonation unit verbalizes a different event or state from the preceding" (1994:69).

Thus speakers choose how to organize information, and intonation (through phrasing in this context) reflects these choices (see also discussion in Miller and Weinert 1998:79). A fairly natural correspondence between clauses and IUs is thus established and many analyses are built on this assumption, e.g. for English, Mandarin Chinese, Japanese (Iwasaki and Tao 1993), and Wardaman (Croft 1995; see Di Cristo and Hirst 1998:36ff for a general discussion).

Another question, often debated in the literature, regards the degree of independence of IUs with respect to grammatical structure (Fox 2000:289). It has been observed — and this is particularly relevant to the study of Jaminjung — that IUs are not always coextensive with clauses. They may correspond to grammatical units smaller than a clause, such as clause modifiers, topicalised items, or parenthetical remarks (Cruttenden 1997:72). This observation is confirmed by the linguists involved in the preparation of the C-ORAL-ROM, a multilingual resource which aims to provide a comparable set of corpora of spontaneous spoken language for four Romance languages. Not only do they find the association of clause

and IU limiting, they are even weary of definitions which associate predicates to IUs, noting that on average 30% of IUs in their corpus are verbless (Cresti and Moneglia 2005:15). Their solution is to define IUs (utterances in the C-ORAL-ROM context) as the linguistic counterpart of *speech acts* understood in the sense of Austin (1962) as “utterances that serve a function in communication”, which are bounded by breaks or ‘silences’.

In this study, the first parsing of speech into units is done during the original transcription of the Jaminjung recordings and thus necessarily based on perception. Semantic criteria are also called upon in the initial marking, inasmuch as the perceived units corresponds to a unit of ‘sense’. Importantly, this perceptual parsing is done before, and devoid of, any acoustic measurements. For the purpose of this study, an IU is defined as a meaningful stretch of speech, often corresponding to a clause (but not exclusively), and bounded by a pause (but not always). The syntactic correlates of the units will be described further in §3.2 below.

3.1.2 Units smaller than the IU

Much effort in prosodic research has gone into defining the correspondences between phonology and grammar, more specifically, on defining how phonological processes map to prosodic domains of various length (phrasing), and how these domains may be organized, particularly in the AM and prosodic hierarchy frameworks. Taylor (2009:116) suggests that the problem can be described either with a ‘top-down’ or a ‘bottom-up’ approach, phrasing being dictated by the linguistic structure in the former, or by the prosodic cues heard in real speech (pausing, timing patterns, pitch) in the latter. He goes on to point out that no model has managed to do both satisfactorily and reconcile the results. I will apply a mixed-method in this study, as it is almost impossible to restrict initial definitional criteria for the posited units to prosody or to syntax only.

The first unit larger than a segment is the syllable. It is notoriously difficult to define, whilst at the same time deemed universal. It seems to have a psychological reality, in the sense that we speak with syllables, not

with sounds. Jaminjung speakers, for example, spontaneously break down words into syllables in slow speech. For our purpose, it is defined as an ordering of sounds (Fox 2000:51) which has a structure: a central peak of sonority (usually a vowel) and the consonants that cluster around this central peak.

Larger units are called *phonological word* or *prosodic word*, often used interchangeably; I will prefer *prosodic word* in this study.

Phonological/prosodic words are defined as being the domain of word stress, phonotactics and segmental rules (Peperkamp 1999:15). The terms need to be distinguished from the more general syntactic entity, *lexical word*, defined as an item “which is stored in the mental dictionary or lexicon” (Cruttenden 1997:23). Lexical and phonological/prosodic words may often be equivalent, but it should not be assumed that what is a word in syntax and what is a word in phonology are always equivalent. Morphology and phonology do not always coincide, and phenomena such as derivation, compounding, and cliticization involve phonological processes that can span domains larger than a lexical word.

Another domain is usually recognized between the prosodic word and the IU: the *phonological phrase* (*phrasal constituent* in this study). It accounts for some processes that cover more than one prosodic word but are still contained within an IU, for example, the *liaison* in French (Cruttenden 1997:24). It often corresponds with major syntactic constituents, although the reverse is not true.

Other degrees of phrasing have been reported to occur between the phonological/prosodic phrase and the IU to account for various processes in different languages, for instance major vs. minor phrase (Selkirk and Tateishi 1988), or intermediate vs. accentual phrase (Beckman and Pierrehumbert 1986). The term *accentual phrase* is used mainly in the AM and prosodic hierarchy theories (§2.2.3.5) to describe the domain of the pitch accent and does not figure in the prosodic hierarchy itself.

For Jaminjung, prosodic words and phrasal constituents are posited. Identifying prosodic words in Jaminjung as bearing stress would again entail a certain degree of circularity. Instead, in a first step, they are defined using morpho-syntax criteria: prosodic words are the units resulting from

the joining of a root and its affixes. Some phonological processes (excluding stress) also apply to these units, they are described in §3.2.2.3 below. There are many clitics in Jaminjung, their very definition and the place they should occupy in phrasing pose a specific difficulty. For the purpose of this study, units created with a root and a clitic are also treated as prosodic words; this question will be discussed further in §3.2.3.1.

Phrasal constituents (the term is preferred to phonological phrase) are also posited, also defined on morpho-syntactic grounds, referring specifically to syntactic phrases such as noun phrases (NPs) or complex predicates. They are described at length in §3.2.3.2 below.

3.1.3 Units larger than the IU

Phonological models of intonation usually restrict their analysis to that of the intonation unit, considering that phenomena involving larger stretches of speech are discourse related. However, many researchers working with spontaneous speech have found that connected speech exhibits patterns that span segments larger than the IU and describe a larger unit, often called the *utterance*.

In the British tradition, a larger unit, the *paratone*, is defined as “a coherent formal sequence of intonation units, analogous to the concept of ‘paragraph’ in writing” (Crystal 2003:336). It is, however, usually considered to be shorter than a written paragraph (Brown et al. 1980:26). It is delimited prosodically by a high onset and a very low pitch and a pause at the end. In Brazil’s (1997) discourse intonation, textual structure is signaled via the *key*, the baseline connecting the low pitch in successive IUs. Lehiste (1975) based her analysis on the *paragraph* in her investigation of the units found in speech production and perception. In a study of Mandarin, Tseng and Chou (1999) use the term *prosodic group* for such a larger unit. Wichmann refers to a *spoken sentence* (2000:128), a notion based on the work of Schuetze-Coburn et al. (1991), which showed that, in English, declination is not reset at each IU but rather over a set of IUs. Ewing (2005:171), in his study of Cirebon Javanese, refers to the *prosodic cluster*, which has an overall intonational trajectory, albeit not necessarily a down-

drift as found in English. He chooses this term as it does not imply “syntactic completeness, as [does] the term prosodic sentence”. He nonetheless likens prosodic clusters to Chafe’s *prosodic sentence* and Poedjosoeddarmo’s *utterances*. Genetti (in press, 8) notes how speakers of Dolakha Newar, a Himalayan language, use intonation to organize prosodic units into macro-units which she calls *prosodic sentences* which “function in narrative to produce prosodic cohesion over a number of independent prosodic units”. These may coincide with syntactic sentences, but not necessarily. Genetti points to similarities between the prosodic sentence and the paratone mentioned above.

In this study, the term *prosodic sentence* will be adopted for units larger than IUs, precisely because of the notion of ‘cohesion’ it implies. It is preferred to the very common ‘utterance’ as the latter is also used in a general way to refer to any stretch of speech, undermining its descriptive capacity. In the transcription process, a grouping of intonation units which forms a coherent semantic unit is marked as a prosodic sentence. It is worth stating again at this point that the definition of the posited sentences whose prosodic correlates are tested is based primarily on semantic and syntactic criteria.

3.1.4 Prosodic cues associated with the boundaries of prosodic units

Prosodic units are identified by prosodic events occurring at their boundaries. The cues that are commonly associated with boundaries will be briefly introduced here.

The most obvious indicators of boundaries between units are pauses. According to psycholinguists such as Goldman-Eisler (1968), pauses reflect the planning of forthcoming verbal output or the ‘generation of information’. Grosjean et al. (1979) have shown that the occurrence of pauses can be related to syntactic structure: the more complex the syntactic constituent is, the more likely a pause is to occur. Also, after reviewing current research in the variables determining the occurrence and duration of pauses, Krivokapić (2007:4) states that

the main conclusion to be drawn from these studies is that length of prosodic utterance plays a role in pause placement and pause duration and that syntactic structure has a large impact on pause placement and duration, but does not fully determine it.

Yang (2004) finds that pauses correlate fairly well with phrase and boundary marking in English and that the duration of the pause is significantly correlated with specific boundary status. Choi (2003) finds, similarly, that pause duration increases from non-boundary to intermediate to intonation boundary. Fletcher et al. (2004:438) suggest three levels of constituency in Dalabon, a language of Northern Australia, cued by the degree of juncture (pausing): a minimal juncture between adjacent words; an accentual phrase; and a full intonational phrase that is almost always followed by a silent pause. Pauses may yet serve other functions in spontaneous speech: in expressive speech, they are used for emphasis or dramatic effect; in conversation, they act as signals for turn-taking and possibly changes in topic direction, and are, as such, boundaries of some sort. For a review of the research on pauses in speech, see Zellner (1994).

Other cues have also been associated with prosodic units' boundaries, for instance, a more or less abrupt change in pitch, i.e. a jump up or down in pitch, which cannot be attributed to the highlighting function of intonation, or changes in duration, intensity, and voice quality. A change in rhythm can also be associated with boundaries, typically expressed as *anacrusis* (Cruttenden 1997:26), which relates to the patterns of acceleration and deceleration during the utterance of a unit, initial syllables being uttered faster than final syllables.

An important cue associated particularly with the left boundary is *pitch reset* which refers to the difference in pitch between two adjacent units. Pitch tends to decline in the course of a unit (§2.1.2), but is generally reset at the beginning of the adjoining unit. A reset is said to occur when the F0 values at the onset of the second unit are markedly higher than at the offset of the preceding unit (Cruttenden 1997:30-34). Research has shown that stronger pitch resets usually correspond to more significant boundaries, at least in English (Wichmann 2000) and in Swedish (Swerts 1997). Oliveira (2003:1) observes that “the melodic discontinuity that occurs

between information units — a consequence of the natural declination of pitch in the course of an utterance — is an important cue for discourse segmentation”. He shows that pitch reset is a very accurate indicator of discourse section boundaries, for example in procedural texts in Suyá, an endangered language of Central Brazil. At the level of the utterance, boundary strength is also associated with pitch reset. Cook and Muehlbauer (2005) report that in Nehiyawewin, an Algonquian language, the clearest acoustic signal to a constituent boundary is a pitch drop preceding the boundary and/or a pitch reset on the syllable following the boundary.

At the right boundaries, syllables at the end of units are usually longer in duration than comparable syllables that are not near a boundary, a phenomenon observed in many languages and possibly a linguistic universal. It is attributed to a slowing down during the production of the utterance, reflected in the signal as *final lengthening*²⁴: the larger the unit, the greater the degree of final lengthening²⁵. Wightman et al. (1992), in a large corpus study of English, found that lengthening differentiates at least three levels of boundaries: non-phrase final word boundaries, intermediate boundaries, and intonational phrase boundaries. Durational changes have been observed even at smaller boundaries (Wagner 2005), to help mark the relative strength of individual word boundaries. Final lengthening leads to longer duration of segments, usually of the vowel, which may result in the sounds being pronounced less loudly and clearly than in other syllables.

Final lowering signals the right boundary of a prosodic declarative utterance in many languages (Selkirk 2003). Karlsson et al. (2008:2), for instance, have identified this parameter in a non-tonal dialect of Kammu, a Mon-Khmer language, where the “default pattern of a neutral utterance uttered as one prosodic unit is a declining F0 course with a low (falling) terminal”.

²⁴ Vaissière (1983:60) reports that this is a phenomenon also observed in music, birdsong and insects chirps.

²⁵ See Tuttle and Lovick (2007) for a similar study in the Dena'ina language, Cambier-Langeveld (1997) for Dutch.

Phonation events, described as “controllable variations in the actions of the glottis”, such as creaky or breathy voice, are also cues for identifying prosodic units in many languages (Gordon and Ladefoged 2001) (see also §2.1.1).

In the specific context of the PENTA model, syllable onsets are seen as the locus of the co-onset of tone, consonant, and vowel, based on empirical evidence which continues to accumulate. A concept issuing from these observations is the ‘time marker hypothesis’ (Xu 2008:22) according to which syllable onsets could be seen as time markers in speech.

A time marker is an event, such as the tick of a clock that serves as a reference for the measurement of time and timing (...) Syllable onsets, where the unidirectional movements toward the initial C, the first V, the tone, and possibly the phonation register, all start simultaneously, seem to serve this purpose well.

In PENTA, it is considered that timing is of critical importance for speech and that duration patterns serve as an *affinity index*, signaling how closely constituents are related. The concept is based on previous observations of the lengthening of final syllables near boundaries (Lehiste 1973), of the subtle durational changes at smaller boundaries that help to mark the relative strength of individual word boundaries (Wagner 2005), and of the often-noted presence of pauses. The affinity index iconically encodes the relational distance between adjacent constituents. It is measured by adding the syllable duration to the pause duration, in other words, it is the temporal distance between the onset of the current syllable and the onset of the upcoming syllable (calculated as the inter-onset interval, or IOI).

3.1.5 Prosodic units in other Australian languages

In this section, research on prosodic units in Australian languages is presented. Many of the studies of Australian languages are developed within the AM theory and the constituents’ definitions refer to this framework.

For Dalabon, a language of Arnhem land (NT), studies by Ross (2003) and Bishop and Fletcher (2005) acknowledge three levels of constituency: words cued by minimal junctures; accentual phrases cued by a

falling or rising intonation contours at the end of a word, followed by a pitch restart on a following word; and intonational phrases which are almost always followed by a silent pause.

For Iwaija (spoken on Croker Island in the northernmost part of the Northern Territory) Birch (2002) states that the only prosodic constituent above the foot for which there is clear evidence is the IP which is marked by both tonal and durational features and contains a nucleus marked by the alignment of a pitch accent with the head of a foot. Birch questions the notion of a word unit, as a distinct constituent from a phrasal unit, noting that words uttered in isolation carry all the prosodic features associated with a well-formed IP, be they durational or tonal features demarcating boundaries, or the presence of a head or nucleus.

In Mawng, spoken in north-west Arnhem Land, Helmuth, Kügler and Singer (2007) propose three levels of constituency: the prosodic word, the phonological phrase, and the IP; phonological phrases being marked by local pitch resets. These IP phrase boundaries are found at relatively major (clause level) syntactic juncture points, such as after a topic or a fronted argument. There are also junctures at more minor syntactic juncture points, for instance after the subject in a subject + verb sequence. The authors suggest that the tonal marking in this case could be an indication that the subject has been (vacuously) fronted, or could be topicalised. They promise further investigation to look into the way many subjects and predicates are separated by prosodic juncture.

Bishop (2003) defines three prosodic constituent levels: the phonological phrase, the IP, and the utterance, for the Kuninjku dialect of Bininj Gun Wok in central Arnhem Land. She notes that the phonological phrase in Kuninjku could correspond to the intermediate phrase in English, as defined in Beckman and Pierrehumbert (1986). There are some substantial differences, however, one being that the unit which defines the relative prominence relationships between accents is the intermediate phrase in English and the phonological phrase in Kuninjku. She goes on to say that the phonological phrase in Kuninjku resembles more closely the Japanese accentual phrase, as they both represent the minimal tonally demarcated unit. She remarks, furthermore, that the phonological phrase in Kuninjku is

most frequently isomorphic with the phonological word (Bishop 2003:387-388).

For Alawa, a neighbouring Non-Pama-Nyungan language, Sharpe (1972:34ff) describes a word, phrase and clause level, although the differences between the last two levels are not clear.

For Warlpiri, a language of Central Australia, Pentland and Laughren (2004:6) propose to distinguish between prosodic words and phonological phrases. For them the first is the domain of stress and the latter the domain of phonological rules but in this case the distinction is one aiming at clarifying the processes observed in Warlpiri word prosody, not necessarily a matter pertaining to prosodic constituency.

Other studies of Australian languages usually distinguish at least a word and a clause level. Patz (2002:35ff) thus describes the intonation of Kuku Yalanji in southern Queensland in terms of word stress and clausal patterns.

This first section has presented an overview of important notions related to the analysis of the grouping function in Jaminjung that will constitute the remainder of this chapter. The definitions of prosodic units were reviewed, the prosodic parameters associated with prosodic units were introduced, the view of the grouping function in the PENTA model was discussed, and prosodic units in other Australian languages were presented. This quick overview helps to position the research presented hereafter in the wider field of intonation studies. The next section will introduce the phonology and grammar of Jaminjung, to motivate the choices of potential units tested in the remainder of the chapter.

3.2 Outline of the grammar of Jaminjung, its (potential) prosodic units and their morpho-syntactic correspondence

Firstly, a short overview of the segmental phonology and phonotactics of Jaminjung and of the structure of its syllables is presented. Secondly, the potential prosodic units are introduced, together with the corresponding morpho-syntactic constituents in the grammar of Jaminjung.

3.2.1 Introduction to the grammar of Jaminjung

Jaminjung shares many of its main characteristics with other Australian languages (see Gaby 2008 for a brief overview). It is said to have free word order²⁶ in the sense that there isn't a more frequent or more basic pattern in the ordering of the arguments in relation with the verb (Schultze-Berndt 2000:108). It also lacks a category *verb phrase*, and lexical arguments can be freely omitted in a clause.

Many of the characteristics of Jaminjung are also found in other non-Pama-Nyungan languages of Northern Australia. Argument roles, for example, are shown by bound pronominals which attach to the verbs as prefixes and by case markers which attach to constituents of noun phrases as suffixes. Jaminjung has complex predicates, a distinctive characteristic of the linguistic area where it is spoken. They consist of inflected verbs, forming a closed category of around thirty members, which associate with members of an uninflected class of words, distinct from nominals, referred to as *coverbs* (Schultze-Berndt 2002:39-40). Reduplication is frequent in nominals and in coverbs. There are very few derivational suffixes and when they do occur, they usually do not change the class of the word to which they attach. Finally, there is a lack of distinction between nouns and adjectives.

²⁶ But see chapter 4, which investigates the possibility that word order is conditioned by information structure on a discourse pragmatic level.

In the next section, the terms *nominal*, *coverb* and *verb* can refer either to the lexical category (roots and stems) or to the word forms consisting of a stem and its inflections.

3.2.2 Phonology and syllable structure in Jaminjung

This is a presentation of the segmental phonology of Jaminjung, describing its phoneme inventory, allophonic variation, and phonotactics. The Jaminjung-Ngaliwurru dictionary database, collated by Schultze-Berndt, is used to obtain the quantitative data and examples.

Generally, the main features of Jaminjung's phonology are similar to that of other Australian languages: it has "a typically small inventory of vowel phonemes, and a relatively large number of place distinctions for consonants" (Gaby 2008).

3.2.2.1 The phonemic inventory of Jaminjung

3.2.2.1.1 Vowels

Jaminjung has three vowels with allophonic variations as indicated in Figure 3-1. The front vowels /i/ and /u/ are the most frequent, followed by the back /a/ and finally by the very rare /e/.

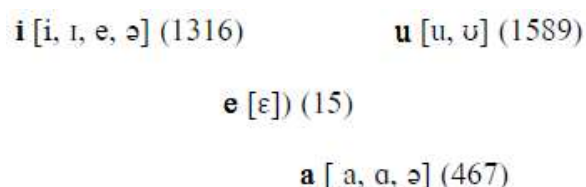


Figure 3-1 The vowels of Jaminjung and the rates of their occurrences (in brackets) in the lexical entries of the database - as of 2009: the database is continuously updated.

The vowel contrasts are illustrated in example (4) in monosyllabic coverbs.

(4)	bal	[bal]	'flat'
	bil	[bil]	'bite (of louse?)'
	bul	[bul]	'emerge, come out, come up, appear'
	yarr	[jar]	'be in one line'
	yirr	[jir]	'move along'
	yurr	[jur]	'rub something on'

A fourth vowel, /e/ is found in a few words. In a review of the distribution of phonemes in Australian languages, Busby (1980:97) classifies Jaminjung as a four-vowel system but notes that the status of the /e/ vowel is uncertain.

Chadwick (1977) describes a three-way vocalic opposition, as does Schultze-Berndt (2000), who notes that the vowel /e/ occurs in a few coverbs only, which, she goes on to suggest, could be either loans from neighbouring languages with four- or five-vowel systems or actually sound symbolic (Schultze-Berndt 2000:41, and 2001:355).

It has been suggested that languages with a smaller set of vowel phonemes display a “greater variation in the pronunciation of vowels” (Pensalfini 2002:21). Jaminjung conforms to this pattern, with a fairly wide range of allophonic variation in its vowels.

- /a/ The low central-back unrounded /a/ can be realised as [a], or as [ɑ], or as [ə].
- /i/ The high front unrounded vowel /i/ has allophones [i] (high close fronted unrounded) found in open syllables and [ɪ] (high open front unrounded) which occurs in closed syllables and as the second element in the long vowel sequence /iyi/, as [e], and as [ə].
- /u/ The high back rounded [u] is sometimes realised as [ʊ] which occurs in open syllables; as [ʊ] (high open rounded) in closed syllables; or as [ɔ] (mid open back rounded) in some rare occurrences.

Long vowels are not common and can be said to be non-distinctive, except in the last example, which is inconclusive, as the name for birds are often derived from their songs and the spelling given here is highly onomatopoeic. In fact, most instances of longer vowels could be interpreted as onomatopoeic, as is shown in the examples given in example (5).

(5)	aa	[a:]	'interjection: ah !'
	yaa yaab	[ja:] [ja:b]	'stroke, touch lightly'
	yaag	[ya:g]	'fish'
	biib bib	[bi:b bi:b]	'blow horn, noise of car horn'
	liiny	[li:ɲ]	'speech, word, language, talk'
	niini	[ni:ni]	'finch, any small bird'
	didid	[didid]	'roll'
	diidiid	[di:d di:d]	'pee-wee bird'
	buu	[bu:]	'dive, enter water'

3.2.2.1.2 Consonants

The consonant inventory is also characteristic of other Australian languages (Dixon 2002). There are six places of articulation: bilabial, apico-alveolar, retroflex, lamino-dental²⁷, lamino-palatal, and velar; and five manners of articulation: occlusives, nasals, laterals, trills, and glides. The nasals and the occlusives have five places of articulation, whilst the laterals have three. There are four liquids: the three laterals, and a trill, the apico-alveolar [r]. The inventory is completed by three glides: one labio-velar, one retroflex, and one lamino-palatal. Again, rather characteristically, there are no fricatives and the voicing of the occlusives is not distinctive. The inventory is shown in Table 3-1 below.

	Bilabial	Apico-alveolar	Retroflex	Lamino-dental	Lamino-palatal	Velar
occlusives	p 833	t <d> 345	ʈ <rd> 169	ʈ <th> 91	c <j> 678	k <g/k> 896
nasals	m <m> 595	n <n> 441	ɳ <rn> 135		ɲ <ny> 307	ŋ <ng> 598
laterals		l <l> 608	ɭ <rl> 215		ʎ <dy> 38	
trill		r <rr> 659				
glides	w <w> 506		ɻ <r> 169		j <y> 286	

Table 3-1 The consonants of Jaminjung (from Schultze-Berndt 2000:41). The pointed brackets indicate the symbols used in the orthography to represent the sound. The numbers beneath represent the rates of occurrence of the phonemes in the lexical database of 2135 entries.

The lamino-dental occlusive /th/²⁸ is very uncommon, with only 91 occurrences²⁹, as is the lamino-palatal lateral /ly/, with 38. Other consonants, whose occurrences are restricted to certain positions, are also less frequent, notably the retroflexes /rn/, /rl/, /r/, with 135, 215 and 169

²⁷ This place of articulation is only found in Jaminjung, not in Ngaliwurru.

²⁸ The phonemes will now be represented to in their orthographic notation.

²⁹ Furthermore, it is restricted to the Jaminjung, as opposed to the Ngaliwurru, dialect.

occurrences, respectively. A closer look at phonotactics (§3.2.2.2 below) explores these restrictions further.

Occlusives (stops) in Jaminjung lack a voiced/voiceless distinction. Usually, an occlusive in initial and intervocalic position is voiced, but is unvoiced in word final position. Occlusives are shown as voiced in the orthographic representation for convenience. The examples in (6) below show some minimal pairs with respect to place of articulation in initial position.

(6)	burd	[bur]	'burn country'
	durd	[dur]	'hold one'
	jurd	[cʊr]	'piled up'
	gurd	[gʊr]	'get up'
	bard	[bar]	'cover up'
	dard	[dar]	'stuck in throat'
	gad	[gad]	'cut'
	thard	[tar]	'start' (loan)

Nasals are very frequent. The contrasts are illustrated in (7) with some minimal pairs below, with nasals in word initial and word final positions.

(7)	marda	[maɾa]	'antbed'
	ngarda	[ŋaɾa]	'anguish'
	malany	[malan]	'taste'
	nyalany	[nalan]	'sexy'
	nad	[nad]	'leaning on elbow'
	ngad	[ŋad]	'bogged'
	warn	[wan]	'get hooked'
	wang	[wan]	'look in vain'

The contrasts in the apico-alveolar, retroflex and lamino-palatal liquids in final position are shown in (8).

(8)	bil	[bil]	'bite'
	birl	[biɭ]	'blow'
	bily	[biɻ]	'burst'

Glides in initial position are illustrated with the near minimal set in (9).

- (9) rang [ɾaŋ] 'have its ears up, prick up its ears, be alert'
 wang [waŋ] 'look in vain'
 yanggi [jaŋgi] 'ask'

3.2.2.2 *Phonotactics*

In this section the phonotactics of Jaminjung i.e. the set of permissible arrangements or sequences of sounds, are presented. It is assumed that phonotactic constraints provide a cue for speakers to identify word or syllable boundaries.

The frequencies of occurrences of the phonemes in initial and final positions in the lexical units of Jaminjung are examined more closely. The database comprises 692 coverbs; 1288 nominals (free pronouns, noun-adjectives, locatives, demonstratives, etc), 53 interjections, particles and clitics, while the remaining items form a group of 42 elements that cross categories. Inflected verb forms are not broken into their constituent morphemes, because they often resyllabify and because morpheme boundaries are not always transparent. They are thus treated separately. There are 742 items in the paradigms for the 32 inflected verbs in Jaminjung, but this is the list as collected so far, and should not be construed as a full set of paradigms (see §3.2.3.1 below for a description of these word classes).

3.2.2.2.1 *Initial position*

All consonants, except retroflex consonants and palatal lateral, occur in initial position albeit with differences in frequency. /ɾ/ in initial position is only found marginally in a few coverbs. The rates of occurrences for the segments in initial position are shown in Figure 3-2 below. The initial segment in all word classes is always a consonant, with the single exception of the interjection *aa* corresponding to English 'ah!'. The most frequent initial segments are the occlusives /b/, /d/, /g/, /j/, the bilabial nasal /m/, and the bilabial glide /w/. The apico-alveolar occlusive /d/ is more frequent in coverbs than in nominals, representing 10.4% to 4.6% of the initial segments in their respective categories. On the other hand, the velar

occlusive /g/ is much more frequent in nominals, 19.7% to 8.7% in coverbs, and so is the bilabial nasal /m/ with 14.2% to 9.8%. Glides occur frequently in word initial position in all categories. The lamino-dental occlusive /th/ is rare but still prefers coverbs, 5.1% to 1.7% of nominals.

Unsurprisingly, the inflected verbs have a much less varied segmental inventory in initial position, with only five consonants and one glide, reflecting the initial consonants of the bound pronominals which are always prefixed (§3.2.3.1.3). The initial segments are not different from that of the other categories, except that the lamino-palatal glide /y/ is relatively more frequent.

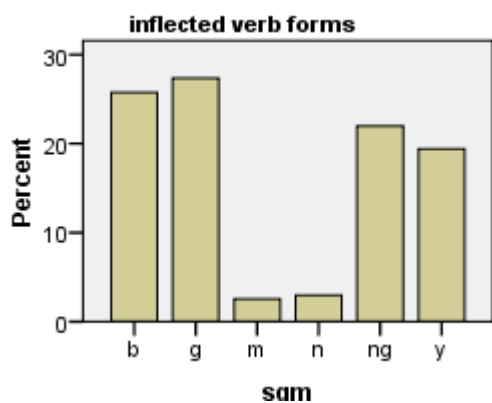
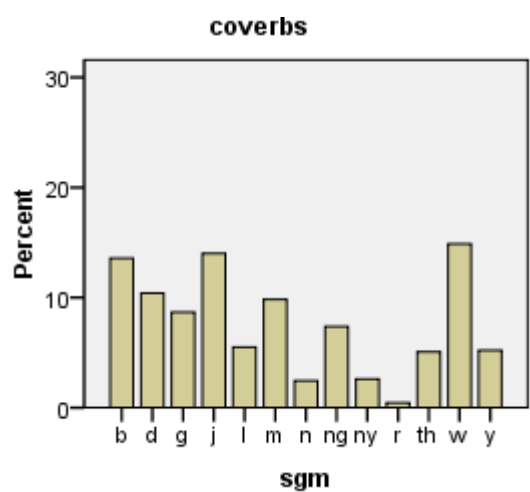
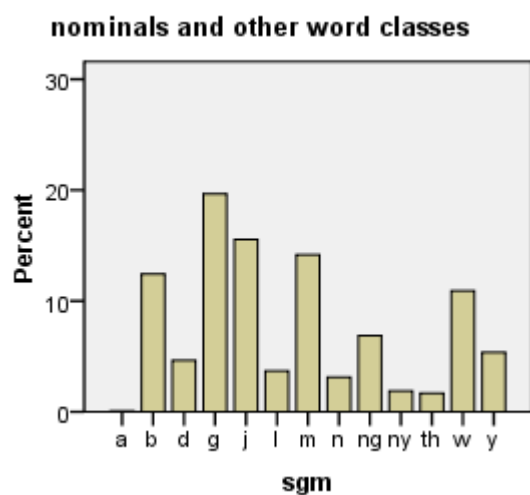


Figure 3-2 The rates of occurrences of segments in initial position in nominals and others word classes (interjections, particles, etc), coverbs and inflected verb forms expressed as percentages.

3.2.2.2.2 *Final position*

Segments in final position are examined next. This time, vowels are part of the inventory, in fact, in total, 40.9% of all mono-morphemic words in Jaminjung end in a vowel (Figure 3-3). In final position, the lamino-dental occlusive (/th/), lamino-palatal lateral (/ly/), and glides (/w/, /r/, and /y/) do not occur.

There is a marked difference between the segments ending nominals (and other word classes) and coverbs. Vowels end 50% of nominals, but only 23% of coverbs. Occlusives account for 16.3% and 42.8% of nominals and coverbs, respectively. The nasals are fairly close at 27.9 and 21.8%. The laterals and the trill, even if altogether much less frequent, still display a difference in their rates of occurrence with 3.1% and 5.2% for the laterals, and 2.6% and 6.1% for the trill. Glides do not usually occur in word final position, except in a few cases where they are followed by a vowel.

In short, nominals are more likely to end in vowels, while coverbs tend to end with consonants, more often occlusives or nasals.

As to the inflected verb forms, they are most likely to end with the vowel /a/, accounting for 45.4% of all items, or with either of the two other vowels /i/ or /u/, in 13.1% and 11.7% of all cases. Altogether, vowels account for more than 70% of all final segments in all inflected verb forms. Otherwise, the final segments are either the bilabial or lamino-palatal nasals /m/, /ny/, or the occlusive /j/.

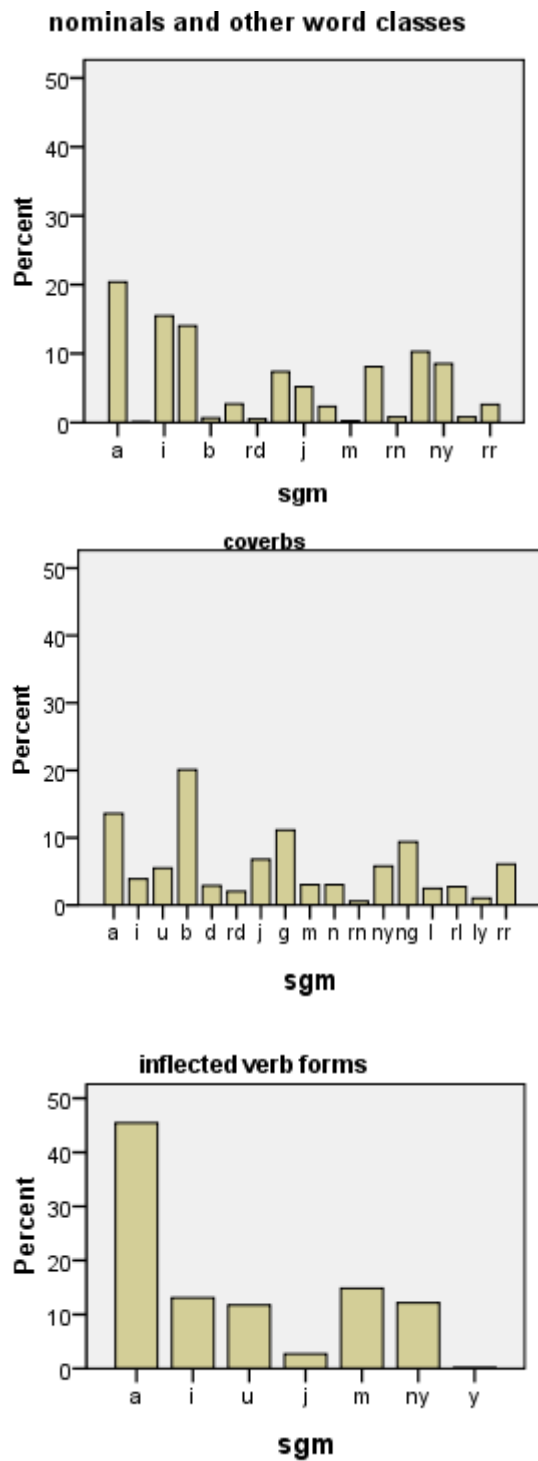


Figure 3-3 The rates of occurrences of segments in final position in nominals and others word classes (interjections, particles, etc.), coverbs and inflected verbs expressed as percentages.

3.2.2.2.3 Consonant clusters

Jaminjung allows a limited number of consonant clusters, which appear predominantly in syllable codas. Clusters can be composed of a lateral (/l/, /rl/), or the trill (/rr/), followed by the bilabial or velar occlusive (/b/, /g/k/), or the velar nasal (/ng/) (Schultze-Berndt 2000 and 2001), thus:

- (10)
- | | |
|--------|----------------------------------|
| <lb> | balb 'struck flat' cv |
| <lg> | bulg 'take out guts' cv |
| <rlg> | garlg 'push' cv |
| <rlng> | jabarlng 'frog' n |
| <rrg> | dirrg 'tied up, bound' cv |
| <rrb> | jaburrb 'wade, walk in water' cv |
| <rrng> | dirrng 'fart' cv |

In word initial position, there is only one permitted cluster, the velar occlusive /g/k/ followed by the lateral /l/ which occurs only once in:

- (11) <gl> gladbug 'fruit of *grewia breviflora*' n

It is very rare in all positions in any case, occurring in 8 nominals and 1 coverb in the database.

The rate of occurrences of the consonant clusters in word final position is shown in Table 3-2. Although heavy syllables are not very frequent overall, with only 5.5% of all the entries in the database, it is interesting to note that coverbs are much more likely to end with consonant clusters: in 12.5% of cases, with a much higher rate of occurrence in one-syllable coverbs (7.5%), compared to only 1.9 % in other word classes. Heavy syllables do not occur in verb forms.

Nb of syll	nominals		coverbs		totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	2	0.1	52	7.5	54	2.6
2	19	1.4	32	4.6	51	2.5
3	6	0.4	2	0.3	8	0.4
4			1	0.1	1	0.0
total	27	1.9	87	12.5	104	5.5

Table 3-2 The rates of occurrences of consonant clusters in word final position.

Verb forms, being never monomorphemic, have the largest number of syllables with various structures, the most frequent being a quadrisyllabic succession of open syllables, CV CV CV CV, followed by the trisyllabic

CV CV CV. Other frequent structures in verb forms are the four syllables CVC CV CV CV CV and the trisyllabic CV CVC CV.

Coverbs are special in that they account for 95 % of all mono-syllabic words and mostly consist of a closed syllable CVC. They are also the only ones to permit heavy syllables in the coda (CVCC). They are otherwise mostly bi-syllabic, with the structure CV CVC, CVC CVC or CV CV. Tri-syllabic coverbs are also frequent, but longer coverbs are rare.

Nominals are mostly two or three syllables long. Their make-up is quite comparable to that of the coverbs of similar length. Apart from the structures noted above in bi-syllabic coverbs, other common structures are three open syllables CV CV CV, or two open syllables followed by a closed one, CV CV CVC. Five-, six-, seven- and eight-syllable nominals are marginally represented in the database and could be said to be the results of lexicalised derivational processes or reduplication.

To look more closely at the types of syllables and the position they occupy in the words, a filter is applied to distinguish between open and closed syllables. In initial position, open syllables are much favoured, accounting for 64.6% of all entries in the lexical items, although they are more common in nominals than in coverbs, accounting for 71.8% and 50.8% of their respective entries. The count is similar in inflected verbs, with 66.31%. The counts are shown in Table 3-3.

Nb of syll	nominals		coverbs		totals		inflected verbs	
	Freq	%	Freq	%	Freq	%	Freq	%
1								
2	329	23.8	211	30.5	540	26	21	2.83
3	418	30.2	111	16	529	25.5	166	22.37
4	208	15	28	4	236	11.4	194	26.15
5	30	2.2	2	0.3	32	1.5		
6	4	0.3			4	0.2		
7	1	0.1			1	0		
8	1	0.1			1	0		
misc	1	0.1			1	0		
Total	992	71.8	352	50.8	1344	64.6	492	66.31

Table 3-3 The occurrence of open syllables in initial position, shown by the number of syllables in nominals, coverbs, the total for both, and the inflected verb forms.

In word final position, however, there is a preference for closed syllables in lexical items: 59% of all the words end with a consonant. Coverbs are much more likely to end in a closed syllable with an average of 77%, while only half of the nominals do (49.9%). Inflected verb forms are much less likely to end with a closed syllable, with a count of only 33.02%. Table 3-4 summarises these findings.

Nb of syll	nominals		coverbs		totals		inflected verbs	
	Freq	%	Freq	%		Freq	%	Freq
1	6	0.4	183	26.4	189	9.1	4	.54
2	334	24.2	253	36.6	587	28.3	17	2.29
3	266	19.2	79	11.4	345	16.6	106	14.29
4	78	5.6	18	2.6	96	4.6	93	12.53
5	6	0.4			6	0.3	25	3.37
6	2	0.1			2	0.1	0	0
Total	692	49.9	533	77	1225	59	245	33.02

Table 3-4 The occurrence of open syllables in final position, shown according to number of syllables in nominals, coverbs, the total for both, and the inflected verb forms.

Word-internal closed syllables are found to occur more often in nominals and in verb forms than in coverbs. Table 3-5 shows these results. For ease of reference, the total number of 3-, 4-, 5-, 6-syllable words in the database is given in the last line of the table.

Nb of syll	nominals		coverbs		totals		inflected verbs	
	Freq	%	Freq	%		Freq	%	Freq
3	111	8.0	23	3.3	134	6.5	87	11.73
4	80	5.8	6	0.9	86	4.1	87	11.73
5	7	0.5	0	0.0	7	0.3	37	4.99
6	1	0.1			1	0.0	0	0
misc	1	0.1			1	0.0	0	0
total	200	14.5	29	4.2	229	10.9	211	28.45
Overall total	852	61.6	178	25.7	1030	49.6	640	86.26

Table 3-5 The occurrence of closed syllables in 2nd position of words of 3 syllables and more.

3.2.2.2.4 Reduplication

Partial or complete reduplication is very common in Jaminjung and very productive. Monomorphemic nominals (including nouns, adjectives, temporals, locatives and demonstratives), coverbs, and some verb roots can

be reduplicated. Reduplication serves to augment or intensify the semantic property of the word. In nominals, it has the function of indicating plural or a multiplicity of referents, whilst in coverbs; it expresses extended duration, repetition or intensity of events. The resulting forms can become lexicalized and are counted as lexical items.

As mentioned above, in nominals they may emphasize the quality or state of the referent; reduplicated forms also include many plant and bird names which are often onomatopoeic. Reduplicated nominals can count up to 7 syllables, especially if a derivational suffix is also added. Some examples are provided in (12).

- (12)
- | | |
|--------------|-------------------------|
| mawurduwurdu | 'small club' |
| barragbarrag | 'diver duck, cormorant' |
| birdibirdij | 'gum tree species' |
| birdugbirdug | 'jacana, bird species' |
| ngurungurung | 'whisper, moustaches' |
| wiriwirimawu | 'glider possum' |

In coverbs, reduplicated forms often describe essentially repetitive or continuous actions. Full reduplication often (but not always) involves monosyllabic coverb roots:

- (13)
- | | |
|----------------|--------------------------------------|
| balabbalab | 'make dog tracks' |
| balthabalthang | 'fall in small lumps' |
| durdurdub | 'roar, thundering noise' |
| girrgirmib | 'remember:CONT' |
| jurduyurdu | 'straight' |
| burngburmg | 'bubble, boil up' |
| warmgwarmg | 'walk, go along by moving wings' |
| balabbalab | 'make footprints like that of a dog' |
| dirrbdirrb | 'make dots in paintings' |

Schultze-Berndt (2000:78) notes that “reduplication has to be distinguished from repetition of a coverb to iconically represent a repeated action; reduplicated coverbs only carry a single word stress, while each reiterated coverb receives its own word stress”. Repeated forms are therefore labeled as two separate lexical units in the dataset when they are not lexicalized.

Reduplication is of interest because of the perception of the correlate of stress in its forms, and thus as an indicator of a potential prosodic unit. When reduplication is complete, the resulting words follow the general

pattern of stress falling on the first syllable (§3.2.2.4), with a secondary stress on the reduplicated item. A more detailed analysis of the phonological processes involved in reduplication and their interplay with stress placement is beyond the scope of this study, but would be of interest in describing potential prosodic phenomena occurring within prosodic words.

To summarise syllable and word phonotactics, lexical units must begin with consonants, or with a glide, never with a vowel. The retroflex /rd/, /rn>, /rl/ (except for a few rare occurrences of the glide /r/), the trill /rr/ and the lamino-palatal lateral /ly/ are not found in word initial position, but they occur in syllable onsets. Lexical units in Jaminjung are more likely to start with the occlusives /b/, /d/, /g/, /j/, the bilabial nasal /m/, and the bilabial glide /w/. Glides do not appear in syllable or word codas.

In final position, there is a marked difference between the segments ending nominals (and other word classes) and inflected verb forms on the one hand, and coverbs on the other. The former are more likely to end in vowels and the latter with consonants (or a closed syllable), most often occlusives or nasals.

Consonant clusters occur very rarely in syllable onsets, but cluster from a small set of permissible combinations are common in syllable codas. The only and very infrequent cluster in word-initial position is the occlusive velar /g/k/ and the apico-alveolar lateral /l/. Consonant clusters mostly occur in codas of mono- or bi-syllabic coverbs. This serves to distinguish coverbs in Jaminjung which appear to have distinctive phonological forms as opposed to other main word classes (nominals and verbs).

Basic constraints on syllable and word structure can be thus expressed as follows: word forms are at least bisyllabic with the exception of coverbs, and a few bimoraic nominals. Permitted syllables are CV, CVC, or CVCC while the structure CVCC is only found in coverbs. There are, finally, some rare instances of CCV, or CCVC structures.

After this review of syllable and word structures, the next section will examine the processes that can cue word boundaries.

3.2.2.3 Word-level phonological processes

3.2.2.3.1 Deletion

In syllable-initial position, the glides /j/ and /w/ can be deleted in fast speech when they precede the vowels /i/ and /u/, so that the syllables /ji/ and /wu/ are heard as [i] and [u]. In their analysis of two geographically fairly distant languages, Djaru and Jingulu, Tsunoda (1981:26) and Pensalfini (2002:21-22) note that in the sequences /iyi/, and /uwu/, the semi-vowel is not realised phonetically. The glide in intervocalic position can also be reduced in Jaminjung, e.g. *juwud* ‘eye, seed’ can be realized as [ju:d] or [juʊd]. The reduction is most obvious in some of the forms of the nominal demonstratives which have been noted as *ngi(yi)ya*.

Deletion also occurs in the first dual exclusive and first plural exclusive pronominal forms and in the bound pronominals that prefix the verb stems that are transparently linked to these:

- (14) *yirrinyi* [yirɪɲi] ‘prn 12DU ABS’
 yirrag [yirag] ‘prn 13PL.OBL’

Deletion of the initial glide results in a syllable consisting of a sole vowel sound, which can nonetheless bear word stress.

3.2.2.3.2 Lenition

Lenition, a process by which consonants in medial or final position are produced with less articulatory effort (Ladefoged 1971:29) and often occurs in unstressed syllables, has been observed in many Australian languages (Butcher in press). Perceptually, lenition has the effect of decreasing the contrastiveness of a sound.

Most relevantly in the context of this chapter, Fougeron (1999:12) has shown that articulatory variations observed on the segmental level can be conditioned by prosody, according to the particular position of the segment in the utterance. She finds that the initial position in prosodic constituents is generally marked in terms of articulatory strength, in comparison with non-initial positions. She goes on to suggest that, while most of these effects do not change the segments themselves, only their realizations, it is possible that some of “these positional characteristics may

have been phonologized” (45). Functionally, the variations associated with certain positions help the hearer deduce the presence of a boundary. These cues combine with other cues (F0, duration, etc) to define prosodic boundaries. Following Fougeron, Ingram and Laughren (2007:14) found, for Warlpiri, that “left edge prominence marking was indicated by suppression of consonantal lenitions in word/phrase initial and post-vocalic posttonic position”. The distributional analysis of phonetic variants they present, however, does not yield conclusive results that would allow them to distinguish levels and positions in the prosodic hierarchy; results which might be due, they suggest, to the shortcomings of the software they used for the research (Ingram and Laughren 2007:16). They are optimistic that further work will lead to more decisive results.

In Jaminjung, the segments /b/ and /g/ lenify to /w/, and /j/ lenifies to /y/. In nominals and coverbs forms, lenition only affects affixes (suffixes) as in examples (15) and reduplicated forms as with *jurduyudu* ‘straight’ in example (13).

- (15)
- | | |
|---------------|---|
| -bari ~wari | deriv. 'having the quality of...' |
| -burru ~wurru | deriv. 'having s.th., possessing s.th.,propriative' |
| -garni ~warni | case 'over, because of, about' |
| -gina ~wina | case 'for, for the benefit of, possessive, propriative' |
| -gu ~wu | case 'to, for' |

For clitics, the lenition process is much less evident. Out of the eighteen clitics listed in the database, nine have occlusives in onset position, among which the only reduced form is:

- (16) =jirram ~yirram clitic 'two/dual'

Lenition also occurs in verb forms and accounts for the alternations found in pronominal and modal verb prefixes. Some stems undergo lenition from an initial bilabial stop /b/ or a lamino-palatal stop /j/ to a glide, /w/ and /y/ respectively, intervocalically (Schultze-Berndt 2000:98-99). For example, the third person singular verb form *gani-wardagarra-m* 3sg:3sg-FOLLOW-PRS is a result of this process from the stem *-bardagarra* FOLLOW.

syllables. In a general presentation of the subject, Kager (1999:143) lists four cross-linguistics properties regarding word stress in so-called stress languages:

- The culminative property: there is only one possible peak in a word. In many languages, this applies to content words only, and function words are prosodically dependent on content words.
- The demarcative property: stress is often placed near the edge of a constituent. It helps to identify boundaries in speech. The favoured positions for primary word stress, cross-linguistically, are (a) the initial syllable, (b) the prefinal syllable, and (c) the final syllable (in decreasing order of popularity among the world's languages).
- The rhythmic property: stress is organized in rhythmic patterns, regularly alternating strong and weak syllables. The smallest units of linguistic rhythm are metrical feet. Trochees are preferred but iambs are also found among the world's languages.
- Quantity-sensitivity: An element that already has some intrinsic prominence is more likely to attract stress. For example, stress tends to be attracted by long vowels rather than by short ones; or stressed vowels tend to get lengthened, thus increasing syllable weight. "Mutually reinforcing relations of prominence and quantity are highly typical for stress systems".

An interpretation of these properties based on the PENTA model is that they are all possible encoding schemes available in languages to demarcate a word unit.

Impressionistically, Jaminjung conforms to a very common pattern in Australian languages³⁰ where stress is heard on the first syllables, with some variation as to the allocation of the secondary stress. By and large, stress placement is assumed to follow a trochaic pattern (McGregor 2004:94). It is worth noting again that most accounts of word stress and its placement in Australian languages are based on the perceptions of the

³⁰ Some languages are reported to have stress on the penultimate syllable (see Pensalfini 1999, Bishop 2003, Pentland et al. 1999, inter alia).

researchers, with very few acoustic studies of word stress, with the notable exception of Warlpiri (Pentland et al.1999).

A few additional comments can be made regarding the perception of stress in Jaminjung. Firstly, each nominal stem, verb form, and coverb receives at least one stress. Secondly, all word classes (nominals, coverbs and verbs) can take suffixes which can be a single phone or can contain up to 3 syllables. Some of these suffixes may receive stress, at least secondary stress. And thirdly, verbs are morphologically complex, and the pattern of stress on the initial syllables is not always as clear as in other word classes. Jaminjung inflected verbs are composed of at least two morphemes: an obligatory pronominal prefix and a root. They can also have mood prefixes and T/A suffixes. Verbs are minimally bisyllabic, but can have up to 6 syllables. In verbs, it is not always clear whether stress falls on an initial syllable or on the stem following the pronominal prefix.

3.2.2.5 Summary

This section has presented an overview of the phonology and phonotactics of Jaminjung. It was shown that the phonemic inventory is not dissimilar to that observed in other Australian languages, with a limited number of vowels and a large number of place and manner distinctions for consonants. Lexical items in Jaminjung are mostly two, three, or four-syllable long; inflected verb forms tend to be longer, counting up to seven syllables. One-syllable words are very rare in nominals, but not in coverbs. Nominals and inflected verb forms are more likely to end in open syllables, and coverbs with closed syllables. Clusters are not frequent; they are found mostly in syllable coda position and almost exclusively in coverbs.

Lenition is the main phonological process observed in Jaminjung. The domain of lenition is the prosodic word which consists of a root and its affixes or reduplicated forms. Clitics are also treated as prosodic words in their own right; however, further testing should help in determining whether they form a unit of their own, as suggested by their resistance to lenition.

Stress is perceived on the first syllable of content words, another widespread pattern in Australian languages. In inflected verb forms, the

main stress may be heard not on the word-initial syllable, which is the bound pronoun, but on the first syllable of the stem.

3.2.3 Prosodic units in Jaminjung

One of the basic assumptions in this chapter is that utterances are built up from smaller building blocks related to one another in a coherent way. Four levels of phrasing are posited for Jaminjung: prosodic words, phrasal constituents, intonation units (IUs), and prosodic sentences. In this section, their morpho-syntactic properties are described in more detail.

3.2.3.1 Prosodic words

The label *prosodic word* is given to units consisting of content words, or free morphemes (nominals, free pronouns, coverbs, verbs, interjections) and their associated bound morphemes (case markers, pronominal prefixes, TAM suffixes, derivational markers, and clitics). The prosodic word is also considered to be the domain of the phonological processes of deletion, lenition and fortition, described in the previous section.

The morphemes of Jaminjung, whether free or bound, are briefly introduced below. Their possible combinations are then introduced.

3.2.3.1.1 Nominals

Nominals in Jaminjung can be nouns, free pronouns, adjectival nominals, adverbial nominals, demonstratives, and interrogatives, which function mainly as constituents of noun phrases. Bound morphemes in the form of derivational morphology, case markers, and clitics attach to nominals.

Nouns can be further subdivided into proper nouns, kinship terms, common nouns and numerals. Example (18) shows two NPs in apposition, the first formed of the Kriol demonstrative *that* with a proper noun *Nawurla*, and the second of a possessive pronoun *ngarrgina*, an adjectival nominal *gujugu*, and a common noun *jayiny*. The speech signal and pitch track – speckled line – are also displayed.

Free pronouns can be subdivided into three categories: possessive, absolutive, and oblique pronouns. The first two are not common, tend to be emphatic, and are restricted to animate referents. The oblique pronouns, however, are relatively frequent, have a tendency to cliticise to the verb (or another constituent), to be unstressed, and to cross-reference lexical arguments. The first person dual inclusive form, *mindì*, is an exception. Interestingly, Schultze-Berndt (2000:67) observes that in some cases it is not related to the semantic role of any participant in the event in question, but refers to the speech act participants. In this context, it functions as an ‘evidential’, when the evidence is shared by both speaker and addressee and the speaker wishes to emphasize the shared nature of the information. This use of the pronoun is attested in the corpus and when used in this function, *mindì* is translated as ‘on you and me’ or ‘you and me watching’.

Adjectival nominals cover the semantic areas of dimension, physical property, age and value. The other semantic areas usually associated with the adjectival class are covered by coverbs in Jaminjung (Schultze-Berndt 2000:47). Adverbial nominals can be subdivided into time nominals and locational nominals. In example (19), the subordinate clause is formed of an adverbial nominal bearing the subordinating clitic *=ma*, *balarrgu=mang*, and the adjectival nominal *garrij* ‘cold’.

Finally, there are six demonstratives in Jaminjung, distinguishing distal and proximal (based on distance from the speaker), and a third form which is used to refer to an entity that is ‘given’ in the context of the utterance. Each of the three demonstratives has two forms, adnominal or adverbial.

(18)

<that> Nawurla ngarrgina gujugu jayiny
that proper.noun 1SG:POSS big MoMo/DaCh

...'that Nawurla, my big granddaughter,'
[IP:ES97_03_02]

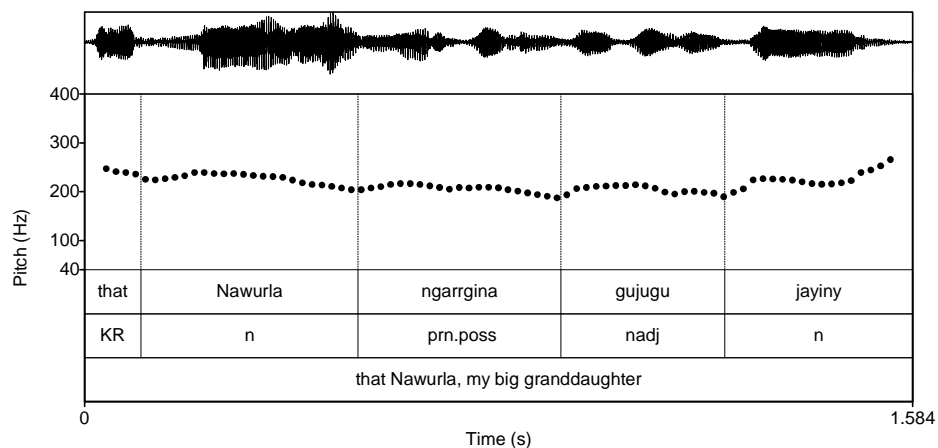


Figure 3-4 An example with 2 NPs in apposition, the first formed of the Kriol demonstrative with a proper noun, and the second of a possessive pronoun, an adjectival nominal, and a common noun. The speech signal and pitch track – speckled line – are also displayed.

(19)

balarrgu=mang garrij
outside =SUBORD cold

'Outside on the other hand (it was) cold.'
[DB:ES97_01_02]

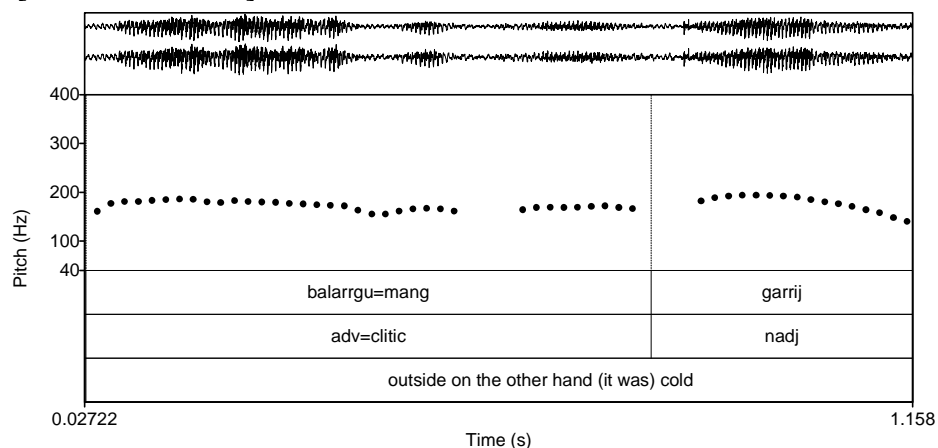


Figure 3-5 The adverbial nominal *balarrgu*

3.2.3.1.2 Case markers

Jaminjung has a fairly elaborate system of case markers which are treated as inflectional suffixes in Schultze-Berndt's analysis. Some of them could be treated as clitics on phonological grounds, in that they can be

separated from the nominals by pauses, or because they have scope over the whole noun phrase. Functionally, “they all serve to relate one constituent to another, but may operate on different syntactic levels. In their prototypical function, case markers operate on the clause level, relating arguments to their predicates: (Schultze-Berndt 2000:52).

An example is given in (20), with the locative case marker *-gi* in a stative predicate.

(20)

jarriny-gi ga-gba
hole-LOC 3sg-be.PST

'...it was in the hole.'

[IP:ES97_03_02]

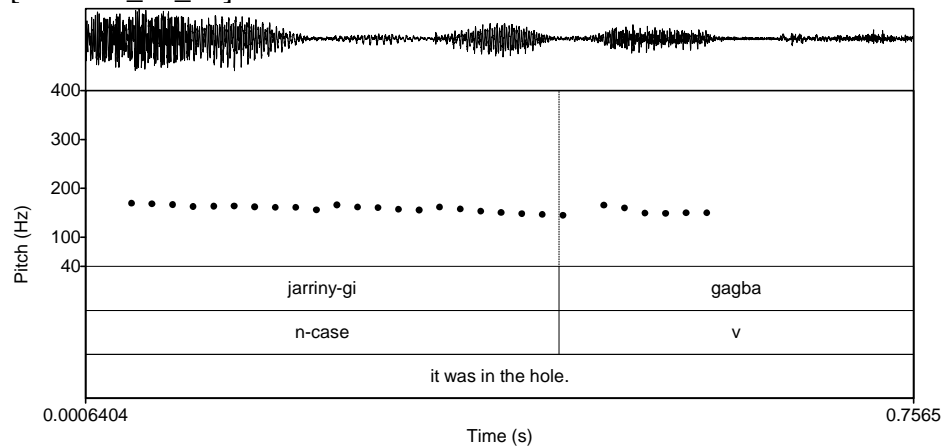


Figure 3-6 The noun *jarriny* followed by the locative case marker *-gi*.

3.2.3.1.3 Verb forms and bound pronominals

A crucial feature of Jaminjung is the division of predicative lexemes into two distinct lexical categories. Verbs, that is, those lexemes carrying verbal inflections, constitute a closed class with around 30 members. Members of an open class of uninflecting elements, termed coverbs, cover a semantic area which is more usually covered by verbs, but also by adverbs and adpositions, in other languages. Consequently, predicative expressions are complex rather than simple, consisting of a coverb and an inflecting verb. The inflecting verbs are introduced here, the coverb category immediately after.

The structure of the inflected verbs can be represented schematically as:

(Mood1-)Bound Pronominals-(Mood2-)V Stem(-Tense/Aspect)
 (Schultze-Berndt 2000:84)

In example (21), the sentence contains the intransitive BE, *bunyjuyu*, which can be analysed as comprising the dual bound pronoun *buny-* and the present stem *-juyu*.

(21)

garlagarla	buny-juyu	jarlig	...	jirram
playing	3du-be.PRS	child		two

'The two kids are playing,'
 [JoJ:CS07_065_01]

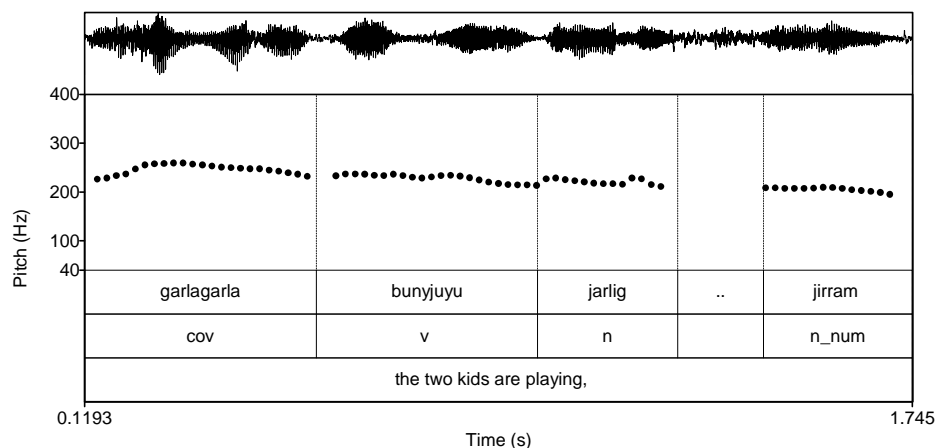


Figure 3-7 An example sentence containing the intransitive verb form *bunyjuyu*.

The affixes comprise pronominal and mood prefixes, and tense/aspect suffixes. For a number of verbs, some tense/aspect categories are expressed by stem suppletion rather than suffixation. Pronominal prefixes (see below) obligatorily occur in all verb forms (except in some imperative forms with a singular addressee). Tense and aspect is only marked in indicative mood (with the exception of imperfective future/potential forms).

Verbs do not have non-finite forms; in Jaminjung, coverbs may fulfill the functions that are held by non-finite verb forms in other languages (Schultze-Berndt 2000:84).

Four mood categories are distinguished in the tense/aspect/mood system: indicative, potential/future, irrealis, and imperative. The last three categories are marked by prefixes to the verb stem, while the indicative is unmarked. Tense or aspect distinctions are only made in indicative and

potential mood. Imperfective and perfective aspect are only distinguished in past indicative and potential/future forms.

3.2.3.1.3.1 Bound pronominals

All verbs take bound pronominals which attach as prefixes. These pronominals express the person/number of at least one argument. The number distinction, however, only applies to higher animates while for lower animates and inanimates, the singular forms are used, though some exceptions have been found to occur in the data (Schultze-Berndt 2000:85). Transitive and intransitive verb stems associate with two different paradigms of pronominal prefixes. Morphologically, the transitive pronominals can be understood as being formed of two parts: the first being the Actor (subject prefix) and the second the Undergoer (object prefix). The intransitive pronominal prefixes and the Actor prefix in the transitive pronominals are mostly formally similar, i.e. pronominal prefixes (unlike NPs) display nominal alignment (Schultze-Berndt 2000:86).

For the purpose of this study, verb forms are treated as unanalysable morphemes, because the affixes that constitute them, if segmentable at all, are often not syllabic. Furthermore, they have never been observed separated by pauses, in slow speech for example, as opposed to case markers and clitics. For ease of identification in the figures illustrating the examples in this section, they are shown with a dash ('-') between the bound pronominal prefixes and the stem, but they are usually represented without this break.

3.2.3.1.4 Coverbs

The second lexical category in complex verbs is the 'coverb'. The term was selected because it shows both the dependent nature of members of this class, and the fact that these lexemes are equivalent in meaning to verbs and adverbs of many other languages (Schultze-Berndt 2000:71). A similar class of words is found in other Australian languages, even if their appellation differs. The same criteria as in other Australian languages can be used to distinguish them from nominals and inflecting verbs: distinctive

morphological marking, differences in syntactic functions, and phonological and phonotactic peculiarities (§3.2.2.2 Phonotactics).

Morphologically, coverbs are distinguished in that they do not take inflections for any verbal category. The only derivations on coverbs which do not result in nominalization are reduplication and marking for continuous aspectual character. Verbs cannot be derived from coverbs.

Coverbs may take a sub-set of case inflections because they may function syntactically as the main predicate in non-finite adverbial clauses, whose relationship to the main clause is encoded by a case marker in ‘complementising’ function (§3.2.3.4.4) (Schultze-Berndt 2000:110). However, they appear as part of a complex predicate together with an inflecting verb.

Coverbs are also distinguished by their phonological and phonotactical characteristics. They are the only lexemes which may form monosyllabic words (apart from rare nominals). They may have certain clusters in word-final syllable codas (e.g. /rrb#/ as in *garrb* ‘gather, pick up many things’), and may contain the mid vowel /e/ (as in *deb* ‘knock down’). A subset of coverbs can be argued to be sound-symbolic (Schultze-Berndt 2000:71, and 2001).

Another piece of evidence for the status of coverbs as a distinct word class is the existence of pro-forms consisting of a demonstrative coverb, *maja* ‘thus, do like that’ and an interrogative coverb, *warndug* ‘how, do what’.

3.2.3.1.5 Particles

Particles are free morphemes. They can be distinguished from the major lexical categories of nominals, verbs, and coverbs, in that they do not inflect, and unlike clitics they are free forms which can occur clause-initially; however, particles such as the *bugu* ‘only’ or *barraj* ‘further’ may also be cliticised to a preceding word. Prosodically, they are interesting in that they do not receive the prominence associated with focus (§5.2.3.3.1) (22) is an example, here the particle *bugu*, showing some iconic lengthening.

(22)

diwu nga-yu bugu garrngan=binji=wung
fly 1sg:3sg-SAY/DO.PST just blood=ONLY=JUST

‘I just chucked it away, just blood (i.e. just bleeding)’

[IP:ES97_03_02]

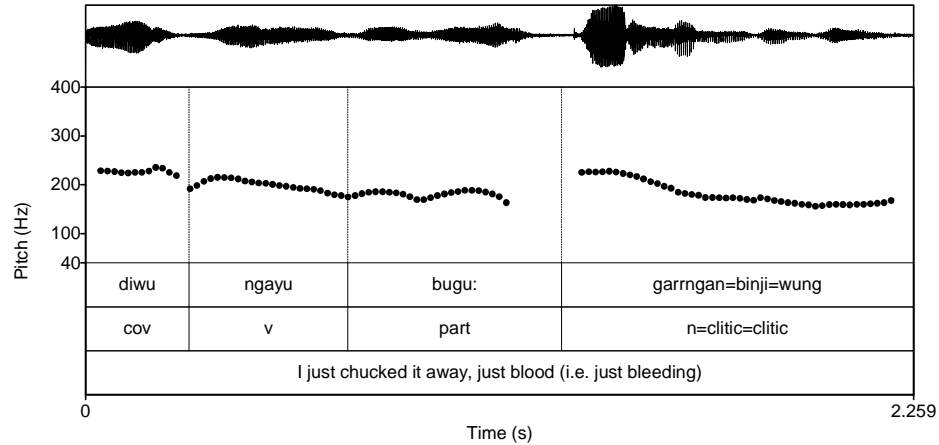


Figure 3-8 The particle *bugu*.

3.2.3.1.6 Clitics

For the purpose of this study, clitics are defined in a morpho-syntactic sense, and in a prosodic sense only secondarily, that is to say, in terms of morpho-syntax, they are considered to be clitics and not affixes because they can attach to different parts of speech. In Jaminjung they are enclitics in that they always follow another constituent. Prosodically, they are integrated with their host and it is assumed that they are no different from affixes. Whether a specific clitic domain exists and, indeed, whether all clitics behave in the same way, will require further research.

For greater clarity in the notation, clitics are distinguished from suffixes by ‘=’ rather than ‘-’ as a boundary symbol.

The next set of examples show the clitic =*biyang*, which is frequent in the corpus. It indicates temporal succession or topic shift, and can attach to verbs (23), coverbs (24), or nominals (25). Its Kriol equivalent, =*na*, is also very common.

(23)

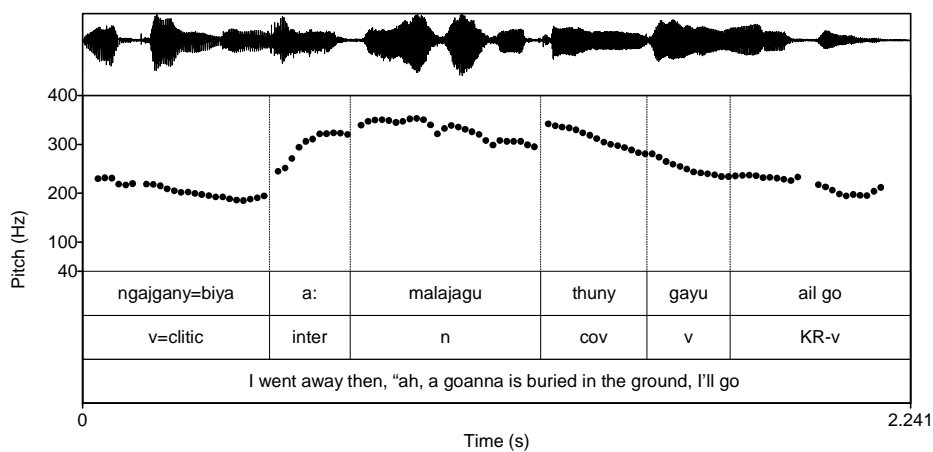
nga-jga-ny=biyang ah marlajagu thuny ga-yu
1sg-be.PST=now ah goanna buried.in.ground 3sg-be.PRS

<ail go>

I'll go

I went away then, "ah, a goanna is buried in the ground, I'll go.

[IP:ES97_03_02]

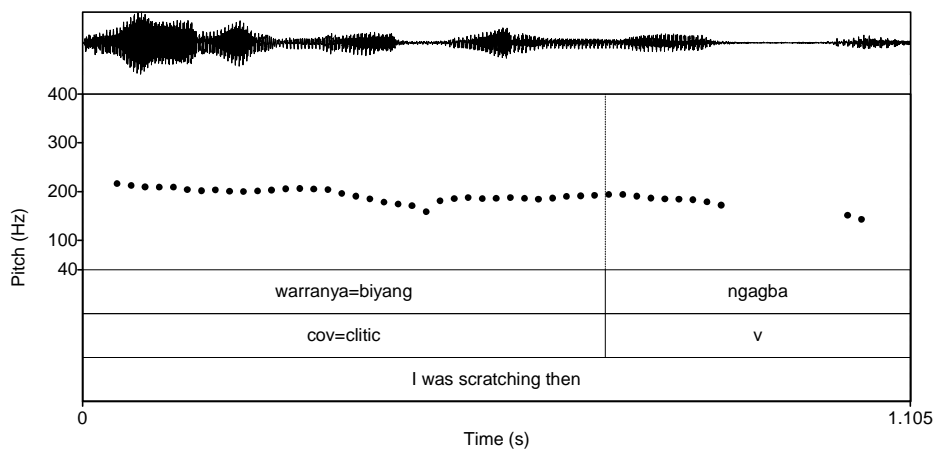


(24)

warranya=biyang nga-gba
remove.cover:CONT=now 1sg-be.PST

'I was scratching then.'

[IP:ES97_03_02]



(25)

ngayug=biyang nga-jga-ny guyawud
1sg=now 1sg-be.PST hungry

'And me, I was hungry.'

[IP:ES97_03_02]

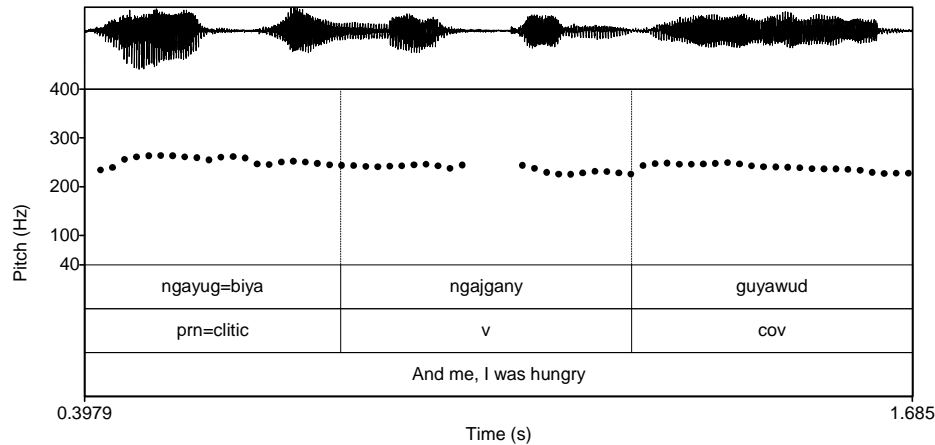


Figure 3-9 The clitic =biyang attaching to a finite verb, a nominal, and a coverb.

3.2.3.1.7 Summary prosodic words

In the previous section, the free and bound morphemes in Jaminjung have been introduced. To understand how the label prosodic word is applied, it is useful to summarize the permitted combinations of free and bound morphemes, which can be represented schematically as:

Prefix	Root	Suffix	Clitics
	nominal	derivational suffix case marker	clitics
(bound pronominal)	verb form	(TAM)	clitics
	coverb	derivational suffix case marker	clitics
	particle		
	interjection		

An illustration of the labeling used for the analysis presented in §3.3 is shown in example (26) and Figure 3-10. The first prosodic word is a nominal *jarlig* ‘child’; followed by *janju-ni=marlang*, a demonstrative with two bound morphemes, a case marker and a clitic; the third prosodic word is the verb form *ganurrungawu*; the fourth is *warrb-bina*, a coverb with a case marker; the fifth is *wirib-ni-mij*, a nominal with two case markers.

(26) (also 102))

[jarlig]_{prosW} [janju-ni=marlang]_{prosW} [ganurru-ngawu]_{prosW}
 child DEM-ERG=GIVEN 3sg:3pl-see.PST

[warrb-bina]_{prosW...} [wirib-ni-mij]_{prosW}
 be.together-ALL dog-ERG-COMIT

'That child saw them sitting together, together with the dog.'

[CP:ES96_18_02]

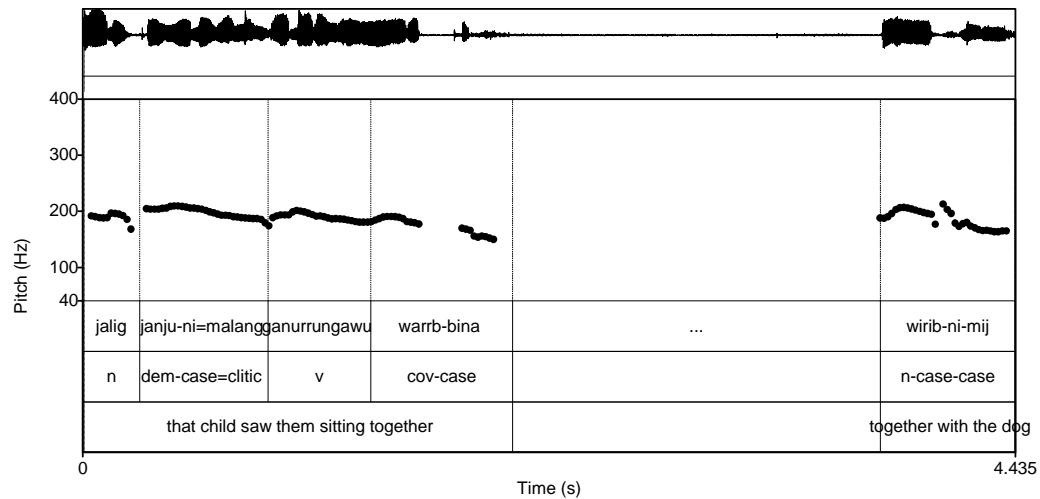


Figure 3-10 A sentence segmented into prosodic words.

3.2.3.2 Phrasal constituents

The second syntactic unit which is tested for prosodic correlates is that of *phrasal constituent*, also defined in terms of morpho-syntactic criteria.

The constituents that receive this label can be:

- Noun Phrases (NP),
- Complex predicates (CP), which can consist either a coverb and an inflected verb, or an inflected verb on its own,
- Non-finite predicates (nfP), consisting of a coverb without an inflected verb,
- Particle group (PG), made of a particle and an NP, or a particle and a CP, etc.

It is important to note that the label CP is extended to include simple inflected verbs, even if they are not, grammatically speaking, complex predicates. This is in order to avoid the label 'verb phrase' since this might be taken to include a predicate and its direct arguments, a constituent for

which there is no evidence in Jaminjung on syntactic grounds. Each of these is described further below.

3.2.3.2.1 Noun phrases

Some researchers have proposed that some Australian languages do not have noun phrases, interpreting co-referential nominals as always being in apposition, “thus accounting for ‘discontinuous noun phrases’, and the lack of a distinction between nouns and adjectives” (Schultze-Berndt 2000:43). This position is not upheld for Jaminjung, for which Schultze-Berndt defines a category, *noun phrases*, with these properties:

- the noun phrase is the domain of case marking; the position of the case marker with respect to the noun phrase is free, that is, it may follow any of its constituents.
- the *apposition* interpretation mentioned above presupposes that all nominals are functionally equivalent, whereas in Jaminjung some nominals are restricted to functioning as heads or modifiers (including determiners), and some others may function as both.

Schultze-Berndt (2000:43) also uses prosodic criteria in her definition of noun phrases as “nominal constituents under a single intonation contour, which are not separated by pauses or other constituents”. This is mentioned here, but is not used as a defining criterion for the phrasal units, with the obvious aim of avoiding circularity.

Structurally, a noun phrase consists minimally of a referential head, which can be accompanied by one or more modifiers, and, optionally, by a demonstrative which functions as a determiner. The order of the constituents in a noun phrase is free, except for the determiner: “a demonstrative can only occur once in a noun phrase, and always precedes any modifier (if present). That is, the determiner either precedes both modifier(s) and head noun (in either order), or it separates the head noun and a following modifier” (Schultze-Berndt 2000:44).

- a) (Det) (Modifier*) Head (Modifier*)
- b) Head (Det) (Modifier*)

3.2.3.2.2 Complex Predicates

In Jaminjung, the combination of a coverb and an inflected verb forms a complex predicate (Schultze-Berndt 2000:118). “Complex predicates can be defined as predicates which are composed of more than one grammatical element (either morphemes or words), each of which contributes a non-trivial part of the information of the complex predicate” (Alsina et al.1997:1). This type of complex verb construction is commonly found in the languages of Northern Australia. The order of the constituents in such a construction is almost always coverb-verb. Figure 3-11 shows an example of such a complex predicate, formed of the coverb *gurrija* ‘dig’ and the finite verb *ngagba* ‘I was’.

(27)

gurrija *nga-gba* *malajagu=gun*
digging 1sg-be.PST goanna=CONTR

‘I was digging for goanna.’

[IP:ES97_03_02]

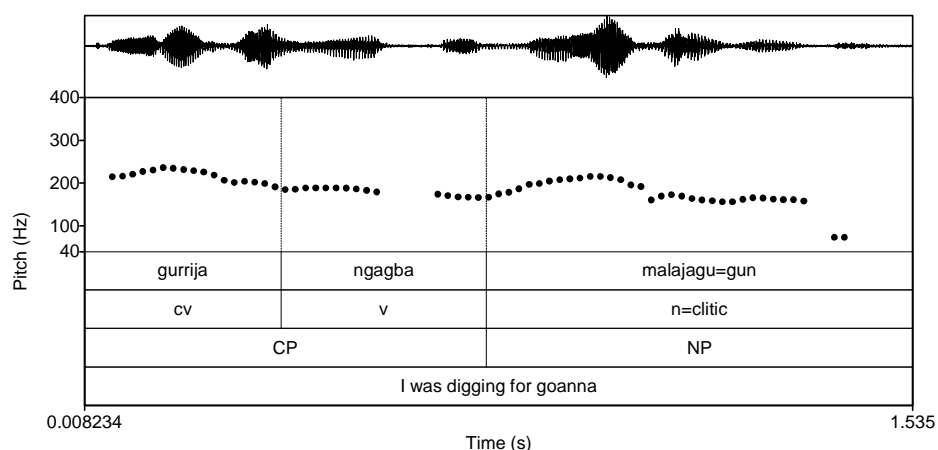


Figure 3-11 The first constituent is a complex predicate, formed by a coverb and an inflected verb.

Predicates can also consist of simple inflected verbs in Jaminjung. Considering the limited inventory of inflected verbs, they are quite common, accounting for 40% of verbal predicates in narrative texts (Schultze-Berndt 2000:118). In the dataset here, as they share the phrasal category as the complex predicates, they are labeled as such to avoid multiplying categories. Measurements will be made to test whether complex predicates consisting of a coverb and a verb, and simple inflected verbs receive the same encoding.

3.2.3.2.3 Non-finite Predicates

Coverbs may function as the only predicate in non-finite clauses, which are either subordinated by means of a case marker or non-finite (§5.2.3.4). Instances where coverbs appear as non-finite predicates are labeled as phrasal constituents in this dataset. An example of such a construction is given in Figure 3-12, where the coverb *mubayib* ‘digging’ is labeled as a phrasal constituent.

(28)

warranya=biyang	nga-gba...	mubayib
remove.cover:CONT=now	1sg-be.PST	dig.out

‘I was scratching then, digging.’

[IP:ES97_03_02]

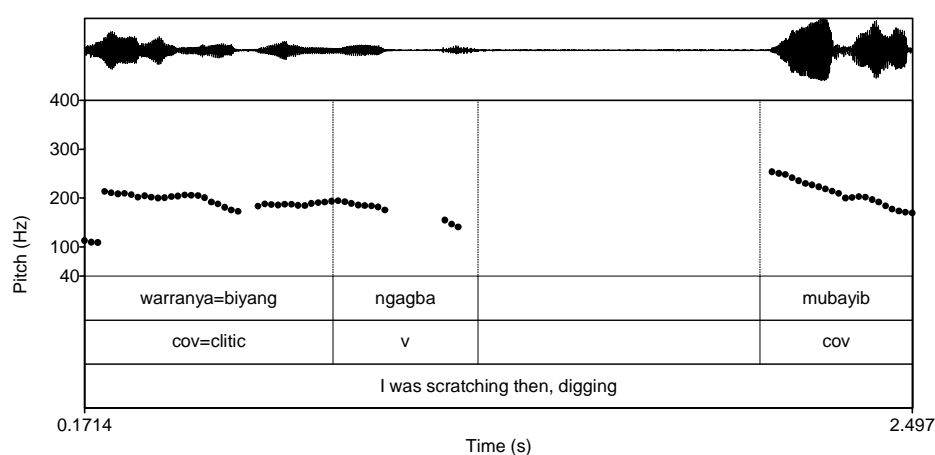


Figure 3-12 The non-finite predicate *mubayib* at the end of the sentence is labeled as a phrasal constituent.

3.2.3.2.4 Particle group

This category is created to account for the groups formed by a particle and an associated noun phrase or complex predicate where the particle precedes the other constituent. Figure 3-13 shows an example in which the particle *majani* ‘maybe’ forms a group with an NP consisting of a demonstrative and its clitic, *ngiyi=na* ‘here now’.

(29)

<go on then> majani ngiyina
go.on.then maybe DIST

“go on then, maybe it is here!”

[IP:ES97_03_02]

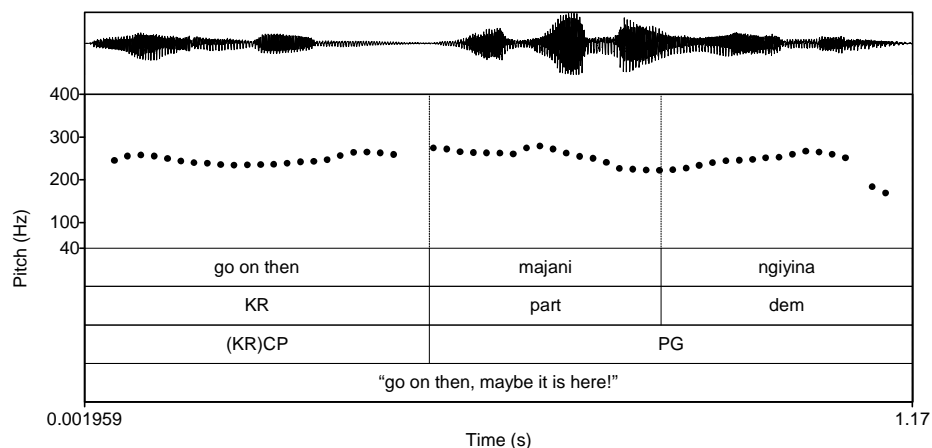


Figure 3-13 The group formed by the particle and a noun phrase is labeled ‘particle group’.

This concludes the presentation of the units that are smaller than intonation units. The next section introduces intonation units, and the prosodic sentence, a unit larger than the IU.

3.2.3.3 Intonation units and prosodic sentences

Intonation units and prosodic sentences will be introduced together in this section. It is assumed that these units are marked by intonational means. There are no morpho-syntactic markers that would identify an intonation unit in Jaminjung, such as the occurrence of a given particle in final position, as reported for Dolakha Newar (Genetti 2007). Prosodic sentences on the other hand are defined in grammatical terms as a succession of verbal or non-verbal clauses, or NPs functioning as afterthoughts or reactivated topics.

3.2.3.3.1 Intonation units

In this research, IUs are the basic unit of analysis. As a working definition, I will posit that IUs are a stretch of speech containing one piece of information uttered under a single intonation contour, marked off by pauses, changes in tempo, and other prosodic cues, which usually (but not always) corresponds to a clause³¹ (Chafe 1994:58, see also McGregor 2004:95). In Jaminjung IUs may correspond to clauses with verbal or non-verbal predicates, to noun phrases, or to interjections. This is also in accordance with Schultze-Berndt (2000:107), who uses the clause as the basic unit of syntactic analysis in Jaminjung, which she defines as “an intonation unit [that] contains a (verbal or nonverbal) predicate, and no more than one predicate”.

3.2.3.3.2 Prosodic sentences

IUs are grouped together in a meaningful way to create larger units, and it is assumed that these larger units are marked by intonational means. The *prosodic sentence* is defined as a sequence of IUs bearing a semantic relation to each other, a relation which is apparent in the structure of the clauses or non-clausal constituents corresponding to its IUs. Prosodic sentences may correspond to a succession of independent clauses, a main clause and a subordinate clause, or to a main clause and a dislocated element. In some instances, they also correspond to direct speech, as an argument of a reporting verb, or to an interjection (forming its own IU) and a main clause.

The next section will introduce the main types of clauses found in Jaminjung.

³¹ The notion of clause and its relation to that of Intonation Unit has already been discussed in §3.1.1.

3.2.3.4 An inventory of clause types in Jaminjung

Independent clauses will be presented first, followed by a review of subordinate clauses. A few general comments can be made about clauses in Jaminjung. Firstly, there are very few subordinate constructions: general finite subordinate clauses (with a verbal predicate) which functionally often correspond to relative clauses; conditional temporal clauses formed with the clitic =*wunthu*; and non-finite subordinate clauses which can be embedded in the main clause through case marking or placed at the right-edge of the sentence, separated by a pause. The latter either have an adverbial or a secondary predicate interpretation.

There are no non-finite complements with obligatory control, such as those introduced by the English complement-taking verbs *tell*, *want* or *try* in Jaminjung. In Jaminjung, this is expressed by a juxtaposed, finite clause, or by using the future form of the verb which can have a desiderative reading, or by clause-level particles (like *birri* ‘TRY’). Perception verbs do not take non-finite complements either.

The more common types of clauses are described and exemplified below. The first tier in the figures shows the prosodic words, the second their part of speech, and the third a free translation. The same examples are shown with interlinear translations underneath.

3.2.3.4.1 Independent clauses: Verbal clauses

Verbal clauses are the most common. They always contain an inflected verb, either as a simple verb or as a constituent of a complex verb. The grammatical order of the arguments is free (it is suggested later that it is conditioned by information structure (see Ch. 4) and lexical arguments can be omitted (Schultze-Berndt 2000:107-108).

I will consider that the syntactic role of subject in Jaminjung can be defined as the NP in intransitive clauses (S), and the Agent (A) in transitive clauses, which can also correspond to the 1st prefix on the inflected verbs. NPs in transitive clauses may also be Objects (O).

An example is given in (30) and Figure 3-14. The sentence is extracted from a narrative in which the speaker recounts how she was once bitten by a centipede. The subject/agent (the speaker herself) is omitted. The

clause consists of a complex verb *dag ngarna* ‘I got warm’ and an NP *guyug-di=biyang* ‘by the fire’.

(30)

guyug-ni=biyang dag nga-rna
 fire-ERG/INSTR=NOW warm.self 1sg-burn-PST

‘I got warmed by the fire,’
 [IP:ES97_03_02]

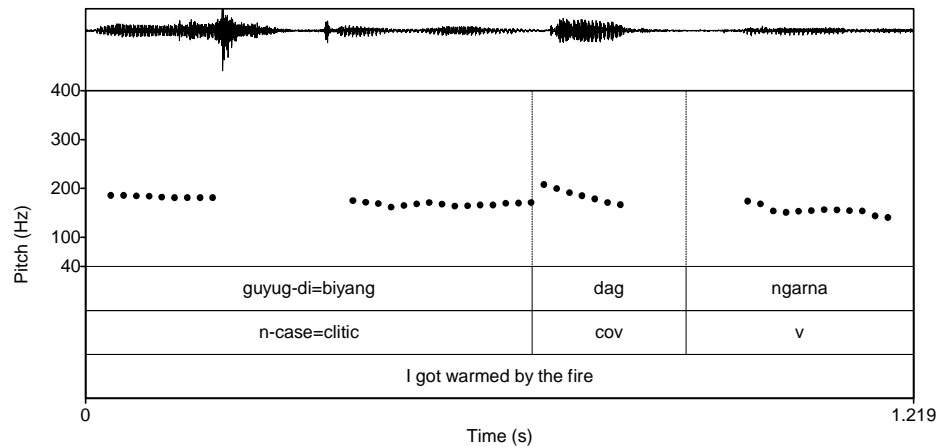


Figure 3-14 An intonation unit corresponding to a verbal clause; the Agent/subject is omitted.

3.2.3.4.2 Independent clauses: Verbless clauses

Verbless clauses are fairly common in Jaminjung; the short list presented here does not include all subtypes and possibilities.

Equative clauses assert or negate the identity of the referents of the subject noun phrase and the noun phrase functioning as predicate. Figure 3-15 shows example (31), extracted from an elicitation session, in which the speaker helps labeling the parts of the body on a poster.

(31)
 nginju... juwiya
 PROX nose

'This (is) the nose.'
 [JM: CS06_014_01]

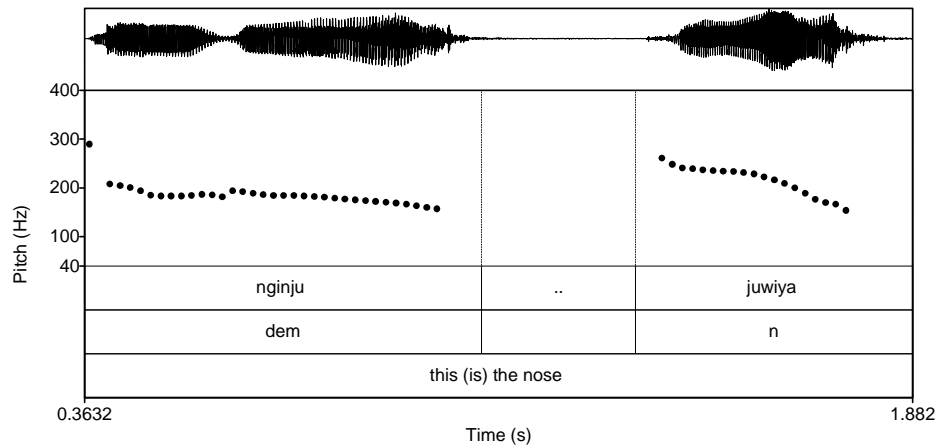


Figure 3-15 An equative clause, with 2 NPs.

Ascriptive clauses characterise the referent of the subject NP. The nominal predicate can be an unmarked nominal from the adjective subclass, or a nominal marked with the proprietive (*-burru* ~ *-wurru*) or privative suffix (*-marnany* (*Jam.*) / *-miyardi* (*Ngali*)) or the 'HABITAT' suffix (*-mawu*). It can also be a case-marked noun phrase. Ascriptive nominal clauses have a further interesting property: if the predication base is a first or second person, it is cross-referenced with an oblique pronominal. In example (32) (Figure 3-16), the speaker explains how each part of the body can be related to a kin.

(32)
gurdbu <rait hand saide>... ngaba-wu
lower.leg right.hand.side older.brother-DAT

'The lower leg, on the right hand side, (is) for the older brother.'

[DR:CS06_021_01]

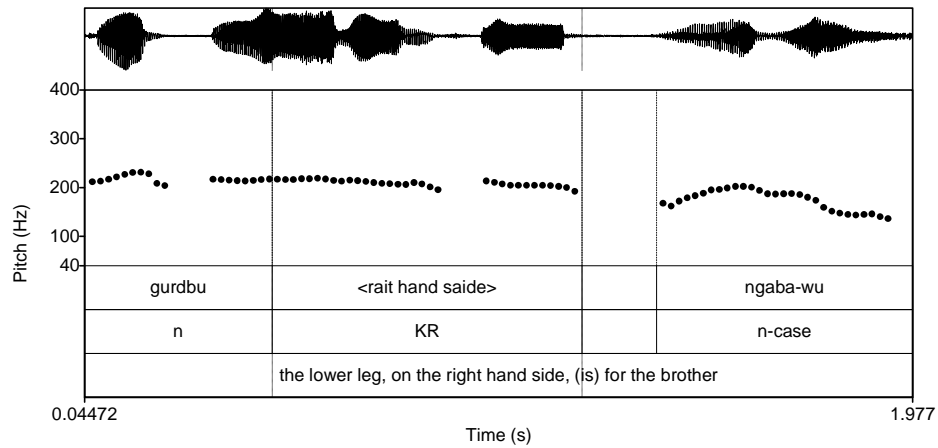


Figure 3-16 An ascriptive clause.

Existential clauses are used to draw attention to the existence of an entity, usually in a particular location (Schultze-Berndt 2000:109-110).

Example (33) (Figure 3-17) is from an elicitation task where speakers were asked to describe pictures, in this case a picture that showed a crocodile underneath a house (on stilts).

(33)
warrij walthub ga-yu jalbud-gi
freshwatercrocodile inside 3SG-be house-LOC

'There's crocodile inside the house.'

[JoJ:CS07_065_01]

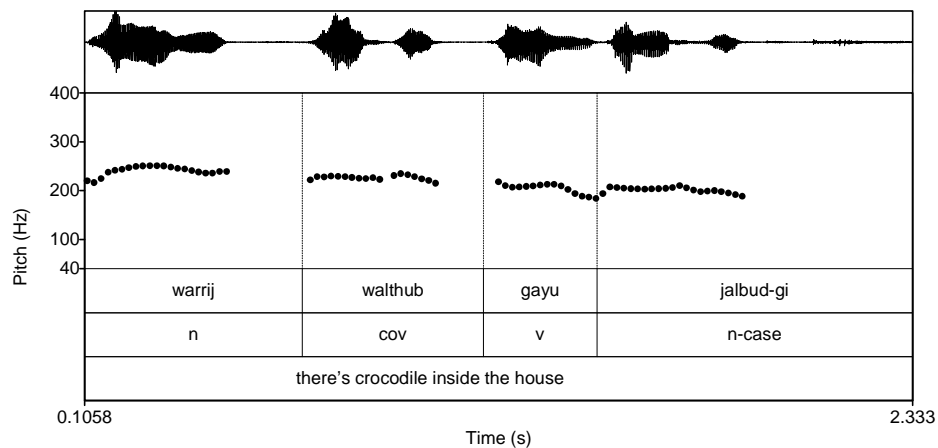


Figure 3-17 An existential clause.

3.2.3.4.3 Finite subordinate clauses

The general finite subordinate clause (with a verbal predicate) often, but not always, functionally corresponds to relative clauses. It can also be adjoined to the main clause and function like a non-specific adverbial clause (Schultze-Berndt 2000:110). It is marked by the clitic *=ma*, *=mang* following the first constituent of the clause.

An example is given in (34) (Figure 3-18) of a prosodic sentence formed of two IUs. The first is a main clause, with a predicate *malajagu gurrija ngagba*. The second is a finite subordinate clause which functions as a relative clause; the first constituent of the clause is marked with the clitic *=mang*.

(34)

marlajagu	gurrija	nga-gba...	thuny=mang	ga-gba
goanna	digging	1sg-be.PST	buried.in.ground=SUBORD	3sg-BE.PST

'I was digging for a goanna, one that was buried in the ground.'

[IP:ES97_03_02]

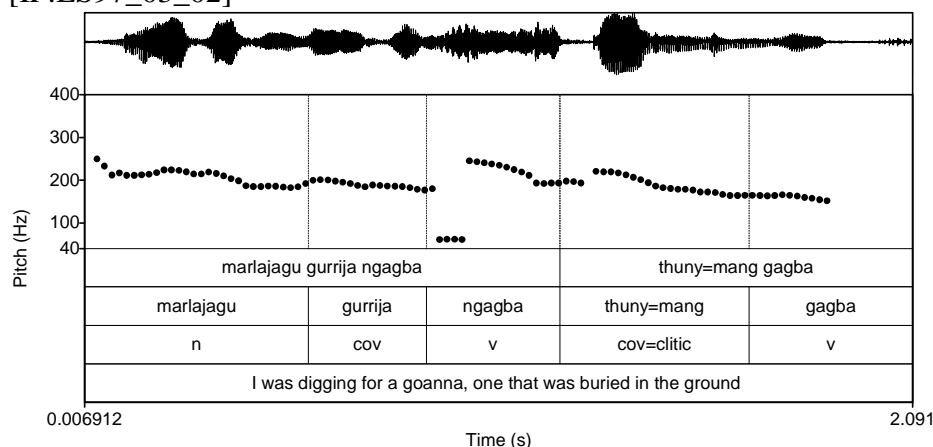


Figure 3-18 A prosodic sentence consisting of 2 IUs: the first is a main clause, with a verbal predicate; the second is a finite subordinate clause which functions as a relative clause.

3.2.3.4.4 Non-finite subordinate clauses

Two subtypes of non-finite subordinate clauses will be distinguished, the 'embedded' type, and the 'edge' type.

The main predicate in an embedded non-finite subordinate clause is always a coverb. Subordinate clauses of this type always take a case marker in 'complementising' function. The 'TIME' suffix *-mindij* is grouped with the case markers partly because it also has the same complementising

function. The non-finite subordinate clause is fully embedded in the main clause, that is, it is not restricted to a marginal position and may function as adverbial or as secondary predicate.

‘The case-marked coverb is often the only constituent of the subordinate clause. If an argument of the coverb is present, it is either unmarked, or (more rarely) has the same case marking as the coverb; both patterns of case marking are also found in noun phrases’ (Schultze-Berndt 2000:111).

Example (35) (Figure 3-19) is an illustration, where the first clause *guruwuny barrbarr burramanyi* is the main clause and consists of a predicate, and the second is a non-finite subordinate clause, formed by the coverb and a dative suffix *thawayawu*.

(35)

guruwuny	barrbarr	burra-ma-nyi	thawayawu
bottle.tree	hit.against-RDP	3pl:3sg-HIT-IMPF	eating-DAT

'they were smashing up boab (nuts) for eating,'
[IP:ES97_03_02]

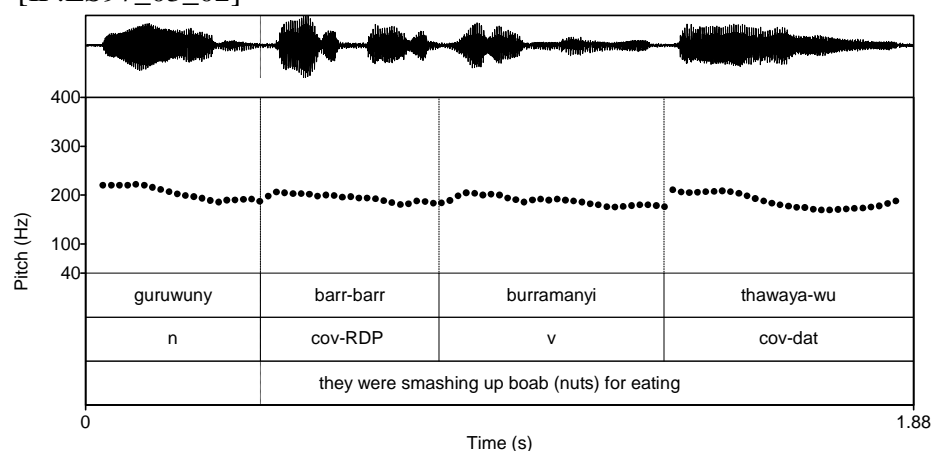


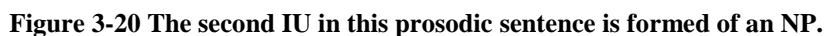
Figure 3-19 A prosodic sentence with 2 clauses: the first is the main clause; the second is a non-finite subordinate clause, formed by a coverb and the dative suffix.

Finally, coverbs can be used as *non-finite predicates* in a distinct, stylistically marked type of clause which will be explored more fully in §5.2.3.4. Coverbs in these constructions usually predicate on one of the arguments of the predicate in the preceding main clause and may receive a resultative or depictive reading. Alternatively, they can also be interpreted in a more general manner as delimiters of the event in the preceding clause.

3.2.3.4.5 Non-clausal IUs

Dislocated elements, usually NPs, may occur at the left or at the right edge of the main clause. Functionally, they may be new or reactivated topics, or afterthoughts. Figure 3-20 shows such a prosodic sentence in which the second IU is a non-clausal construction, consisting of the NP *bujmawu buyud* added after the main clause. It is used here as an afterthought, to give more information on one of the arguments of the clause, i.e. to specify what kind of sand is used (see §5.2.3.4 for an analysis of these constructions).

"They got sand for me with a shovel, the bush kind of sand."
[IP:ES97_03_02]



136

combination of sounds not otherwise found in the language. I will follow Ameka (1992:108) in considering interjections as syntactically independent, in that they can constitute an utterance by themselves, and thus “always constitute an intonation unit by themselves”. There are few interjections in Jaminjung, distinguished from particles and clitics in that they are always stressed. Figure 3-21 shows example (37) with an interjection, *awu*, appearing at the beginning of a sentence (it is also reported speech).

(37)

awu, *nganthan=gun diwu* *nga-yu* *nguntharlg*
no what=CONTR fly1SG- say/do.PST heavy

“No, what is it that I threw away, something heavy”

37 [IP:ES97_03_02]

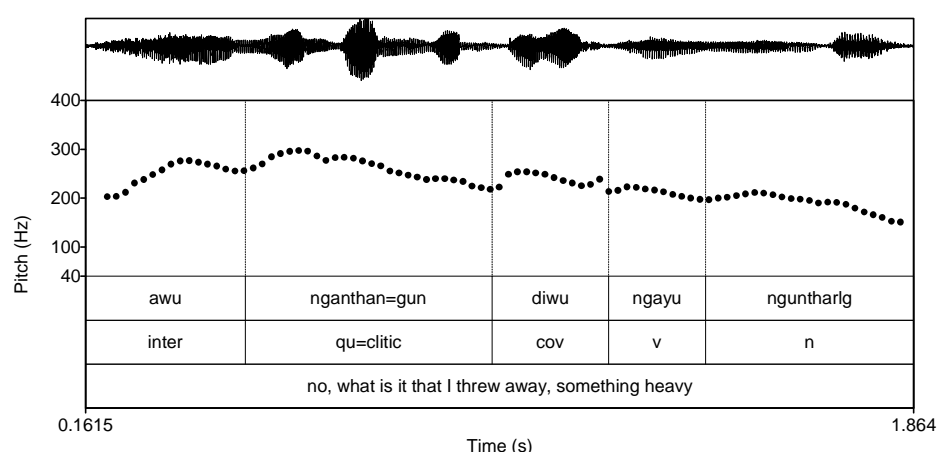


Figure 3-21 The interjection *awu*, at the beginning of a sentence.

Before concluding this section, a few examples of prosodic sentences will illustrate how subsequent IUs can form a semantic whole, even if syntactically they correspond to independent clauses, which has been the original motivation for positing this larger unit.

Figure 3-22 shows example (38), a prosodic sentence which has two IUs. This is the main event in the narrative (the speaker recalls being bitten by a centipede). The two IUs form a semantically cohesive whole as far as the events they describe are in immediate temporal succession.

(38)

<that> jalarriny <bin> wirriny ngarrgu mam
that centipede been turn 1SG.OBL hold.tight

gan-ba=biyang
3sg:1sg-bite.PST=now

'That centipede turned on me, and bit me with a tight grip (i.e. it didn't let go).'

[IP:ES97_03_02]

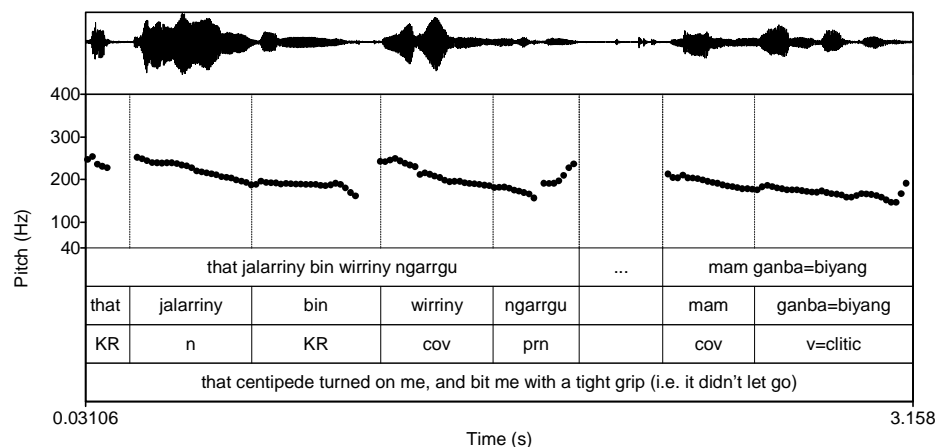


Figure 3-22 A prosodic sentence containing two IUs, describing a succession of events in a narrative.

In example (39) (Figure 3-23), the prosodic sentence consists of two IUs, corresponding to a clause containing a reporting verb *burruyurrgu* followed by a sequence of direct speech as its complement.

(39)

<en> burru-yu=rrgu <maitbi> jalarriny-ni
and 3pl:3sg-say/do.PST=1SG.OBL maybe centipede-ERG/INST

ganiny-ba
3sg:2sg-bite.PST

'And they said to me, 'maybe a centipede bit you'.

[IP:ES97_03_02]

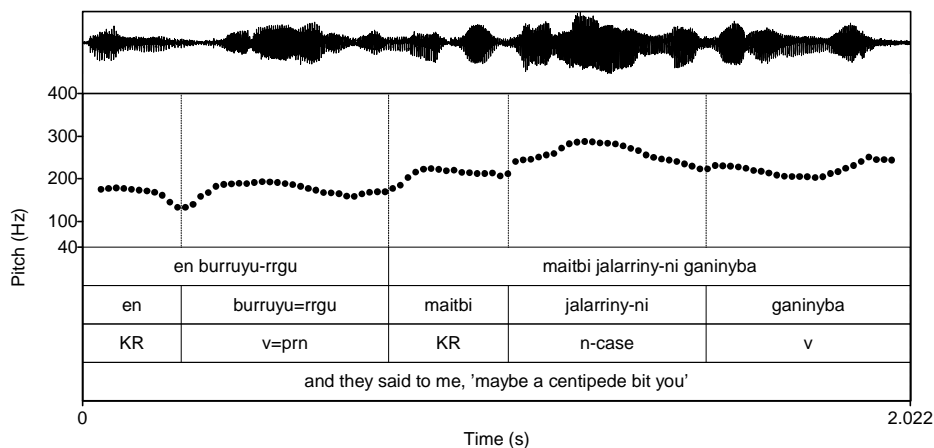


Figure 3-23 A prosodic sentence consisting of two IUs, the first corresponding to a clause containing a reporting verb followed by a second IU, which is a sequence of direct speech as complement to the first clause.

3.2.3.5 Summary

This section has presented an overview of the grammar of Jaminjung, and the potential units investigated in this study. It is important to keep in mind that for the purposes of this study, all of the posited units are defined on grammatical grounds, with the exception of the IUs. For ease of reference they are summarised in Table 3-6, below.

Unit	Morpho-syntactic correspondence
Prosodic words	Content words (nominals, free pronouns, coverbs, verbs, interjections) and associated bound morphemes (case markers, pronominal prefixes, TAM suffixes, derivational markers, and clitics).
Phrasal constituents	NPs, Complex predicates (coverb + verb or inflected verb), non-finite predicates (coverb without an inflected verb), particle groups (particle + NP or particle + CP etc).
IU	≈Clause, also dislocated elements (NPs, coverbs, etc), interjections.
Prosodic sentences	Adjacent clauses: main clause and subordinate clause, dislocated elements, etc. which form a semantic whole, also possibility of independent clauses.

Table 3-6 Morpho-syntactic equivalence of the potential units.

In the next section, the analysis and evidence for the potential units in Jaminjung is presented.

3.3 Analysis: prosodic constituents in Jaminjung

This section reports on the results of the measurements effected in order to ascertain whether the posited units are truly prosodic units, that is, whether the units established on morpho-syntactic grounds (with the aid of some phonological criteria in the case of IUs) are also units at the prosodic level, and what parameters are used in their encoding.

Prosodic features associated with the right- and the left-edge of the units are examined. At the right boundaries, the durational cue is examined to assess the extent of final syllable lengthening in each unit and whether the inter-onset interval (IOI) is a good measure of boundary strength. Amongst the prosodic words, coverbs, because of their specific phonological characteristics (§3.2.2.2.3) are given special attention. Final pitch lowering, also expressed in absolute and relative terms, and the final underlying targets are evaluated. At the left edge, the correlates of pitch reset, underlying target of the first syllables, pitch excursion, and duration are

measured and discussed. In addition, the correlates associated with stress will be discussed.

3.3.1 Methodology

The first step consists in labeling the potential units in the tokens in the dataset. The previous section has detailed the criteria on which the labeling is based: they are grammatical for the prosodic words and phrasal constituents, while IUs also rely on some prosodic parameters.

Figure 3-24 (example (40)) illustrates the segmentation and labeling. In this example, the different prosodic units are indicated in separate tiers for ease of reference. In the actual textgrids used for the analysis, the labeling of the units is concentrated in the first tier and diacritics are used to indicate boundary units.

The first tier shows the syllabic segmentation, the second tier shows the prosodic words, the third tier the prosodic constituents, the fourth shows the IUs, and the fifth the prosodic sentence. The sixth tier contains a free translation.

(40)

wirib-ni=gayi wib gani-ngawu janju buwuny ...
dog--ERG/INSTR=ALSO look.behind 3sg:3sg-see.PST DEM marsupial.rat

jarriny-gi=mang bul gani-ma
hole-LOC=SUBORD emerge 3sg:3sg-hit.PST

'The dog also looked back at that rat, the one that appeared in the hole.'

[CP:ES96_018_02]

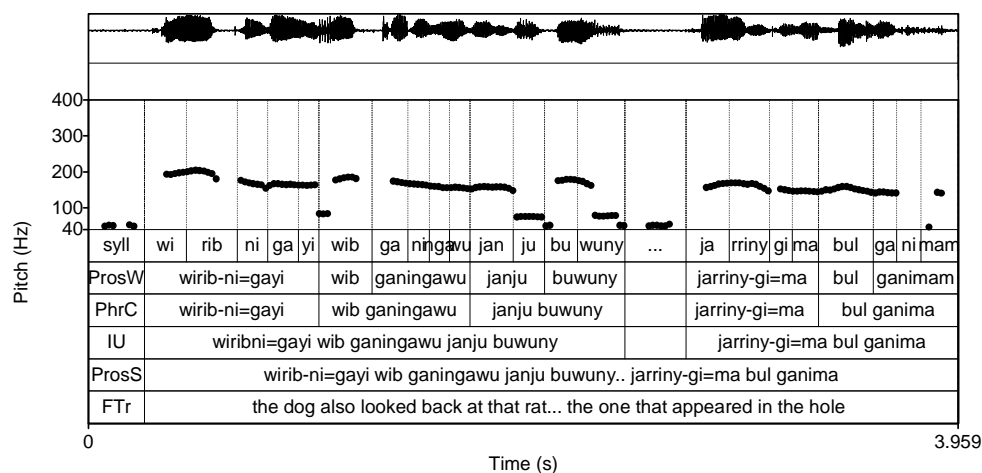


Figure 3-24 The annotation of a textgrid. The first tier is the syllable parsing, the boundaries of the other prosodic units are shown in the subsequent tiers.

Measurements of absolute duration are used to calculate final lengthening. The relative duration of the last syllables is established as a ratio of the length of the final syllables compared to that of the preceding syllables in the same unit. It is worth noting that for speech sounds (segments), which usually range in duration from 30ms to 300ms, studies have shown that differences must be 10 to 40 ms in length before they are judged by listeners to have a ‘just-noticeable difference’ in duration (Lehiste 1970:13). This is taken into account in the interpretation of the results below. In order to establish a threshold for this study of Jaminjung, I will follow Amir et al. (2004) in considering a syllable to be lengthened if it is 10% longer than the preceding syllables (ratio > 1.1).

The occurrences and length of pauses are counted. Boundary strength, or affinity index, is measured as the inter onset interval (IOI), calculated as the duration from the onset of the final syllable to the onset of the following syllable in the next unit. Pauses are thus added to syllable duration when they occur.

Final lowering is a relative value, comparing the mean F0 of the final syllable to that of the previous syllable in the unit.

At the left edge, mean pitch measures are used to calculate the degree of pitch reset, by comparing the pitch of the unit-initial syllable to that of the final syllable in the preceding unit.

Underlying targets are evaluated with final velocity measurements (see §2.3.3 for a detailed explanation of this measure).

3.3.2 Dataset

The dataset consists of unrehearsed narratives:

- (1) Two spontaneous narratives, the *jalarriny* story, in which the speaker tells how she was once bitten by a centipede [IP:ES97_03_02]; and the ‘twitching’ story, where the speaker retells an anecdote from a recent fishing trip [DR:CS06_a021_01].

- (2) One stimuli-based narrative, ‘the Frog story’ [CP:ES96_018_02])³².
- (3) One mythological narrative, the Emu and the Brolga
[DM:MH96_A19_01].

Some of the materials from the texts are excluded from the analysis when they include unclear speech which makes an acoustic analysis impossible, e.g. long hesitations or ambient noise interfering with the analysis.

Table 3-7 shows a count of the total number of potential units in each text.

text		Prosodic word	Phrasal constituent	IU	Prosodic sentence
Jalarriny	N	197	140	77	42
	%	20.7%	19%	17.6%	19.7%
Frog_story	N	384	293	184	81
	%	40.3%	39.7%	42%	38%
Emu_Brolga	N	257	214	122	65
	%	27.0%	29%	27.9%	30.5%
‘Twitching’	N	114	91	55	25
	%	12.0%	12.3%	12.6%	11.7%
Total	N	952	738	438	213
	%	100.0%	100.0%	100.0%	100.0%

Table 3-7 Summary of the potential units in each of the texts used in the dataset.

The length of the potential units is indicated by the number of syllables they contain. The counts are displayed in Table 3-8 below. Prosodic words are more likely to be up to four syllables long, with much fewer occurrences of words up to seven syllables long. Phrases have a wider range, from one syllable up to twelve syllables, but are more likely to have five or six syllables. The length of IUs can be anything from one to twenty-five syllables, with very few tokens of more than fifteen syllables. Finally, the longest prosodic sentence has thirty-eight syllables, but prosodic

³² By way of example, the sound file and textgrid for the ‘Frog Story’ is provided with the accompanying sound files (under the name ES96_018_02_Sample text _grouping).

sentences are more likely to have up to thirteen or fourteen syllables. These counts corroborate the expected accrual in complexity of the potential units.

Number of syllables	Prosodic words		Phrasal constituents		IUs		Prosodic sentences	
	N	%	N	%	N	%	N	%
1	952	33.5%	738	26.0%	438	15.4%	213	6.9%
2	863	30.4%	706	24.8%	422	14.8%	212	6.9%
3	568	20.0%	571	20.1%	382	13.4%	213	6.9%
4	314	11.0%	413	14.5%	333	11.7%	211	6.9%
5	104	3.7%	231	8.1%	268	9.4%	201	6.5%
6	30	1.1%	118	4.2%	224	7.9%	192	6.2%
7	11	.4%	45	1.6%	181	6.4%	184	6.0%
8			10	.4%	139	4.9%	180	5.9%
9			5	.2%	117	4.1%	166	5.4%
10			3	.1%	93	3.3%	150	4.9%
11			2	.1%	67	2.4%	144	4.7%
12			1	.0%	53	1.9%	132	4.3%
13					36	1.3%	121	3.9%
14					25	.9%	109	3.5%
15					17	.6%	90	2.9%
16					13	.5%	74	2.4%
17					9	.3%	67	2.2%
18					8	.3%	63	2.1%
19					5	.2%	55	1.8%
20					2	.1%	45	1.5%
21					2	.1%	41	1.3%
22					2	.1%	37	1.2%
23					3	.1%	31	1.0%
24					2	.1%	26	.8%
25					1	.0%	24	.8%
26							19	.6%
27							15	.5%
28							14	.5%
29							12	.4%
30							8	.3%
31							5	.2%
32							4	.1%
33							3	.1%
34							3	.1%
35							2	.1%
36							2	.1%
37							2	.1%
38							1	.0%

Table 3-8 Number of syllables in each of the potential units. N is the number of occurrences and % is the percentage of the total for this unit.

The table presents all the units labeled. For the analysis, not all units are used, for example, for the phenomena associated with the right boundaries, only the units that do not share their boundary with a larger unit are kept for the investigations, to confirm that any observed effect can truly be associated with the measured unit. This is illustrated with example (41), including an interlinear gloss, and the units shown in the last line. The units that do not share a right boundary with the unit of the next hierarchical level are shown in bold.

(41)

garlagarla buragba Wirlma=nguji=gun
play 3pl-BE.PST proper.name=ETC=CONTR

'They were playing, W. and others,'

[[[garlagarla]ProsW [buragba] ProsW]PhrC [[Wirlma= nguji=gun] ProsW]
PhrC]IU
[IP:ES97_03_02]

A similar, but reversed, selection is effected for the investigation of the phenomena associated with the left boundaries.

The pitch range for each of the speakers in the texts used in this dataset is shown in Table 3-9, where they appear both in Hz and in semitones. The results reveal that there is surprisingly little variation in the female speakers' range (IP, CP, and R), while the range for the male speaker of the Emu-Brolga story is slightly less wide. The ranges vary from 23 to 31 semitones, or a little more than two octaves.

Speaker/text	Max (Hz)	Min (Hz)	Range (Hz)	Range (st)
IP/ Jal	394.41	72.53	321.88	29.31
CP/ frog	365.72	62.73	302.99	30.52
DM/ emu_brolga	170.45	45.49	124.96	22.87
DR/ twitching	354.82	59.77	295.05	30.83

Table 3-9 Pitch range of the speakers in the corpus, expressed in Hz and in semitones.

3.3.3 Right-edge boundaries

This section investigates the phenomena occurring at the right-edge boundaries of the units, including durational events such as final syllable

lengthening, pausing and affinity index; and pitch phenomena including final F0 lowering and phonation events.

3.3.3.1 Duration

The first prosodic cue to be measured is syllable duration. This section will give details of the measurements of both absolute and relative syllable durations.

3.3.3.1.1 Absolute duration

The means of the raw measurements of duration of the final syllables for each identified potential unit are compared here.

The results are shown in Table 3-10, where the mean value is expressed in ms. The number of units and the standard deviation³³ are also shown. The statistical relevance of the correlation between duration and type of unit is supported by the results of an ANOVA test with duration as dependent variable and unit type as independent variable ($F(3, 928) = 16.8$, $p=.000$).

Unit	Mean (ms)	N	Std. Dev
ProsW	185.63	214	68.91
PhrC	161.56	285	66.89
IU	204.05	220	72.85
ProsS	190.87	209	69.30

Table 3-10 Mean duration of the last syllables of the potential units.

The following observations can be made from these measurements. The duration of the final syllables increases with the complexity of the unit, but the gradation is not linear, with phrasal constituents and prosodic sentences breaking the pattern. The final syllables in prosodic words are notably longer than in phrasal constituents, with 185.63ms and 161.56ms, respectively. This is somewhat surprising. It is possible that the longer

³³ The standard deviation measures the spread of the data about the mean value. It is useful in comparing sets of data which may have the same mean but a different range.

durations of prosodic words are due to the overrepresentation of coverbs in this unit, given their specific phonological make-up (§3.2.3.1.4). In effect, they account for 60% of tokens, a rate which is probably more an artifact of the selection process for this dataset than a reflection of their occurrences in natural speech. This issue is addressed in the subsequent section on the affinity index calculation. The final syllables in prosodic sentences are shorter than those of IUs, with 190.87ms to 204.05ms. In fact, the values for the prosodic words and the prosodic sentences are very close. Post hoc tests show that the most significant variation is found in phrasal constituents.

The values of the mean duration of the last syllable in the potential units do not indicate a linear progression in the durations of the last syllables of the units from prosodic words to sentences. The following section will take into account the relative duration of the final syllables to the other syllables in the same unit.

3.3.3.1.2 Relative duration

Final syllable lengthening implies a comparison of the duration of the last syllable of a unit with all (and each) of the preceding syllables in the same unit, a comparison which is usually expressed as a ratio. This measurement is presented here. Following Amir et al. (2004) in their study of Hebrew, it is considered that lengthening occurs whenever the ratio is greater than 1.1 (>1.1), that is to say, the duration of the final syllable is more than 10% longer than the average duration in the other syllables.

The boxplot shows the results of the comparison (Figure 3-25); a horizontal line shows the 1.1 mark. The table on the right shows the values of the ratios for each potential unit, as well as the number of tokens and standard deviations. An ANOVA test with the occurrences of ratios as a dependent variable and the type of unit as independent variable shows that the two are significantly correlated ($F(3, 79) = 4.247, p = .008$).

The measures indicate that only IUs and prosodic sentences have values above 1.1; thus lengthening occurs only in these two units. Interestingly, the lengthening is greater in the final syllables of IUs than in those of the prosodic sentences. This may be in part explained by the fact that IUs may have an overall level contour, which occurs only in non-

sentence final positions and is marked with longer final syllables, as discussed in chapter 4 (§5.2.3.2).

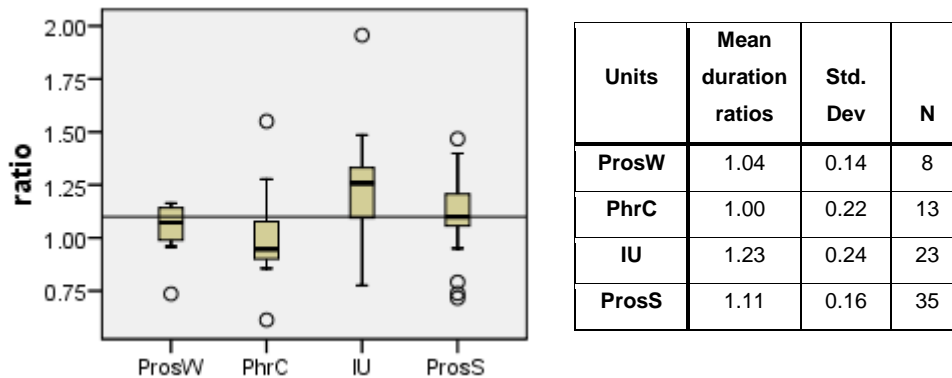


Figure 3-25 The boxplot on the left show the average ratios per type of unit. The table on the right shows the actual values of the ratios.

3.3.3.2 Pauses

In this section, pauses are examined. Firstly, their absolute duration is calculated; these measurements are then used to calculate the affinity index. In order to test pause as a possible boundary cue in Jaminjung, the durations of all breaks between the posited units are measured; any gap in the waveform of 30ms or more is labeled as a pause. The pauses are then associated with the previous prosodic unit. A label *other pause* is created to accommodate events such as coughs, hesitations etc. This type of discontinuity is not included in the following measurements.

Firstly, the total number of pauses associated with each potential unit is counted. Prosodic words in the dataset are never followed by a pause. There are also very few cases of phrasal constituents delimited by pauses, only 11 in total. Pauses occur mostly after IUs and prosodic sentences, in 189 and 200 instances, respectively.

The duration of the pauses after each unit is shown in Figure 3-26. Phrases are followed by shorter pauses, 287.28ms; IUs by longer pauses, 689.98ms; and prosodic sentences by much longer pauses, averaging 1653.40ms. The standard deviation also indicates a wide variety in the length of the pauses for each unit. The results of an ANOVA test with pause duration as a dependent variable and the type of units associated with the

pause as an independent variable show a significant correlation between them ($F(2, 399) = 61.389, p = .000$).

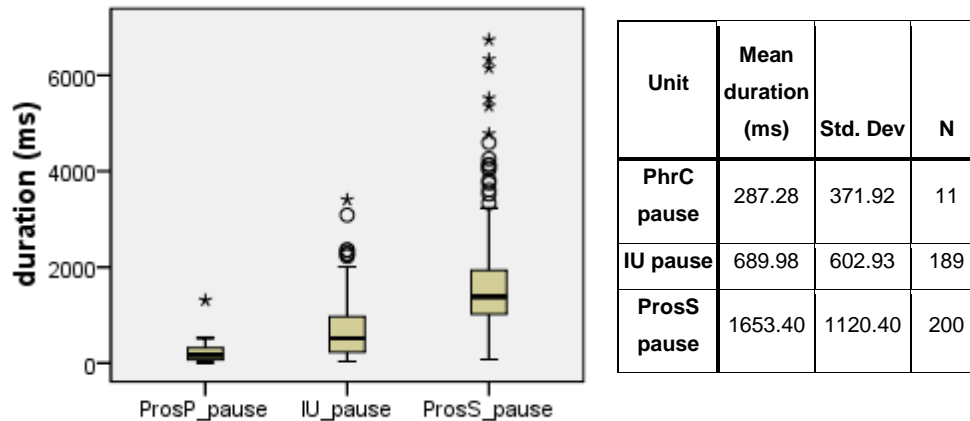


Figure 3-26 Duration of pauses after each potential unit (ms).

Given that pauses do not occur after prosodic words, and very rarely do so after phrasal constituents, it is fit to posit them as encoding parameters of IUs and prosodic sentences only.

3.3.3.3 Affinity index

To account fully for the durational correlate at the right-edge of the potential units, a further measurement is made, that of the inter-onset interval (IOI), as an indication of the *affinity index* (§3.1.4). The IOI values for each potential unit are counted and then compared.

The results are quite evidently in harmony with the measurements of the durations of the final syllables and that of the pauses, displaying a gradation in length at least between the phrasal constituents, IUs and prosodic sentences. The average IOI values of prosodic words and phrasal constituents are very close, 183.83ms and 182.61ms respectively. The affinity index for IUs is 794.08ms and for prosodic sentences 1862.13ms. The ANOVA test shows a strong correlation between the average duration of the IOI and the type of unit ($F(3, 914) = 354.754, p = .000$). Post hoc tests indicate that the difference between the first 2 units is not significant, but those between these units, IUs, and prosodic sentences are significant. Figure 3-27 below summarises the findings.

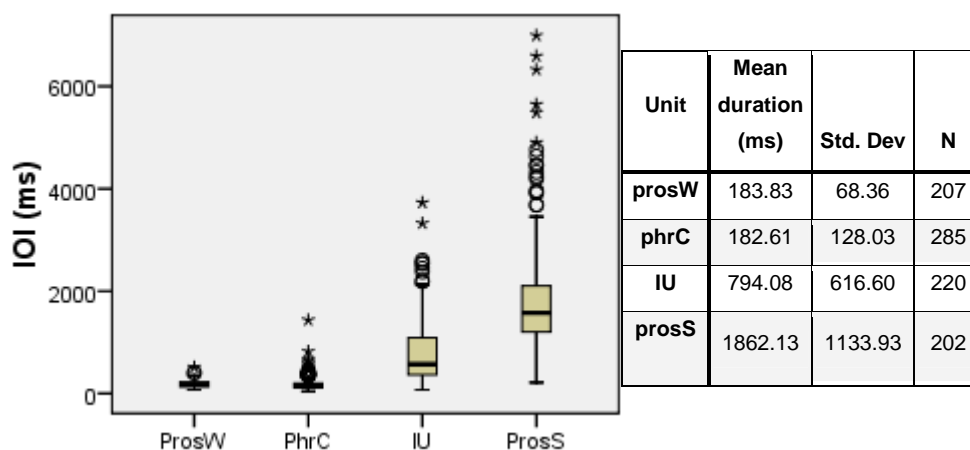


Figure 3-27 Affinity index, a measure of the inter-onset interval for each potential unit.

The measurements of the affinity index for the potential units in Jaminjung distinguish three levels of groupings, corresponding to:

Level 1

prosodic words → content words with affixes and clitics

phrasal constituents → NPs, complex predicates

Level 2

IUs → clauses

Level 3

prosodic sentences → association of clauses

To complete this examination of the durational cue, the possible effect on the measurements of the phonological specificities of the coverbs is examined.

3.3.3.3.1 Coverbs as prosodic words

So far, the results of the durational cues do not yield unambiguous evidence to distinguish between prosodic words and phrases. In accordance with the affinity index, a shorter Inter Onset Interval (IOI) is expected in units that are closely related and a longer interval in units that are less closely related. Hence, prosodic words which consist of content words and their associated bound morphemes should have shorter IOI than phrasal constituents, which consist of NPs or complex predicates, etc (and can

contain two or more prosodic words). However, both units have very similar IOIs, as reported in the previous section. A final test is done to validate these findings. The hypothesis is that the high incidence of coverbs as prosodic words could influence the results, notably because of their specific phonological properties.

A filter is thus applied to remove the coverbs, and the affinity index is recalculated, revealing that the IOI of prosodic words then becomes 165.36ms, which is much shorter than the 183.83ms average first calculated. That of phrasal constituents is slightly larger, at 182.61ms (

Table 3-11). Statistically, the difference between the units is significant $F(3, 780) = 272.953, p = .000$), however, post hoc tests show that the difference in the measurements of the first two units is not significant.

Unit	affinity index (ms)	Std. Dev	N
ProsW	165.36	65.97	73
PhrC	182.61	128.03	285
IU	794.08	616.60	220
ProsS	1862.13	1133.93	202

Table 3-11 Average duration/IOI of the last syllable of the prosodic words, phrasal constituents, IUs and prosodic sentences, with coverbs excluded from measurements in the prosodic words category.

These measurements confirm the validity of the affinity index, by showing that prosodic words have weaker boundary cues (in terms of duration) than prosodic phrases when coverbs are removed from the dataset. It further highlights the specific characteristics of coverbs as a word class.

So far, the measures of the durational cue suggest that the units posited can be ordered in the following manner, according firstly to their absolute values:

Phrasal Constituents < Prosodic words < Prosodic sentences < Intonation Units

Relative duration, expressed as the ratio of the duration of the unit-final syllable to that of the other syllables in the same unit, suggest a slightly different ordering, where the first two units have very close values.

Prosodic words /Phrasal Constituents < Prosodic sentences < Intonation Units

The affinity index reveals a re-ordering of the units, indicating weaker boundaries (as far as duration is concerned) from the prosodic words to the prosodic sentences. The results are strengthened when the coverbs, which are overly represented in prosodic words, are excluded from the measurements. The results further highlight the specific characteristics of coverbs as a word class. The units can be represented as:

Prosodic words < Phrasal Constituents < Intonation Units < Prosodic sentences

Thus, this makes the IOI the most reliable cue for distinguishing all units.

The next section will examine the cues associated with pitch at the right-edge of the posited units.

3.3.3.4 Pitch and phonation events

3.3.3.4.1 Lowering of F0

Final lowering, functioning as a marker of finality, is not expected to occur in units smaller than the prosodic sentences, as they are necessarily non-final. Results showing both absolute values of mean F0 and relative values for lowering are presented. Absolute values of the mean pitch of the final syllables in each unit are compared; then relative lowering is determined through a comparison of the change in pitch within each of the units. The pitch of the final syllables is thus compared to that of their first syllables.

The absolute mean F0 values for the final syllables of each unit are presented in Table 3-12.

The results show a clear decrease in F0 from unit to unit. An ANOVA test ($F(3, 872) = 13.334, p = .000$) reveals a significant correlation

between the type of unit and the mean F0 of its final syllable. The average value for the prosodic words is 176.51Hz, the values for the phrasal constituents and IUs then decrease, with respective averages of 165.41Hz and 164.22 Hz. Finally, the final syllables of prosodic sentences have an average F0 of 147.69Hz.

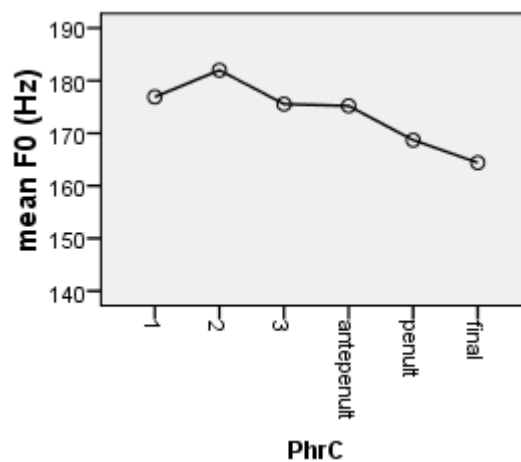
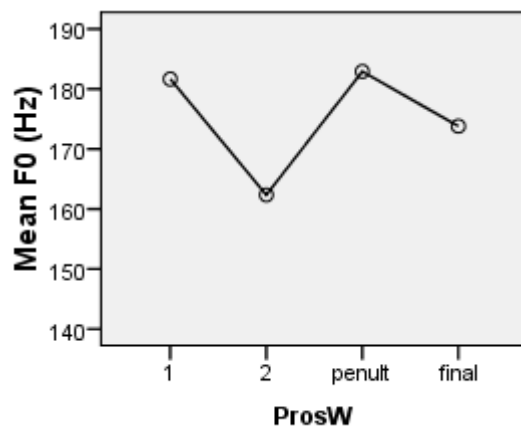
Units	Mean F0 (Hz)	Std. Dev	N
ProsW	176.51	38.29	214
PhrC	165.41	41.42	281
IU	164.22	46.37	207
ProsS	147.69	53.42	170

Table 3-12 Mean F0 (in Hz) measures of the last syllables of each unit.

Next, the relative lowering of pitch is measured for each posited unit. It is assumed that the end lowering is relative to the pitch at the beginning of a unit, so only the syllables at the left and right edges of the units are retained. For the prosodic words, which tend to be shorter, the comparison includes only the first and the last syllables; monosyllabic prosodic words are excluded. For the phrasal constituents, IUs, and prosodic sentences, the comparisons include the first and last three syllables, when units are long enough. With shorter units, of 2, 3, 4 or 5 syllables, only the first and last syllables are used. All monosyllabic units are excluded. Firstly, the rates of occurrence of final lowering are counted for each of the units. Amongst the 171 prosodic words examined, 99, or 58%, have a lower F0 in the final syllables than the first syllable. Out of 271 phrasal constituents, 111 (59%) have lower F0 on their final syllable, accounting for 59%. 142 out of 206 IUs, or 69%, have final lowering; and out of 210 prosodic sentences, 168 have lowering, or 80%.

The graphs (Figure 3-28) below show the averages over the syllables at the left and right boundaries of the units, illustrating the patterns of F0 lowering. All units undergo some lowering from the penultimate to the last syllables, except for the IUs, which again may be explained by the occurrence of an overall rising non-final contour. When the values at the left boundaries are taken in account, the overall pitch in prosodic words does not decline, rather, it appears to rise slightly throughout the unit. In phrasal

constituents, the general pattern is that of a slow decline. In IUs, the syllables at the left boundary have higher mean F0s than those at the right boundary, even if the decrease is not very sharp. In prosodic sentences, the lowering is quite indisputable. The results of the ANOVA tests for the measures of mean F0 by syllable position is significant for all units, except for the prosodic words (ProsW $F(3, 412) = 1.732$, $p = .160$; PhrC $F(5, 1074) = 4.127$, $p = .001$; IUs $F(5, 1630) = 3.886$, $p = .002$; ProsS $F(5, 1120) = 15.389$, $p = .000$).



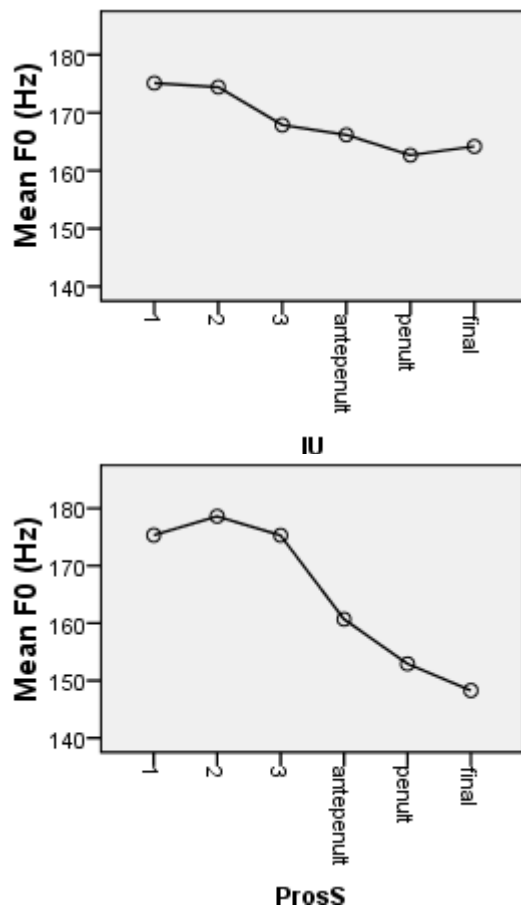


Figure 3-28 The lowering of F0 shown for each type of unit.

In summary, the measurements indicate that lowering occurs in larger units, possibly in phrasal constituents, but most convincingly in IUs and prosodic sentences. The measures presented in this section reflect the surface F0 movements only. In the next section, the final velocity measurement is used to assess the pitch targets associated with unit boundaries.

3.3.3.4.2 Final Velocity

In PENTA, the final velocity measure is used as an indicator of the underlying pitch target. In this section, the values of final velocity of the syllables at the right boundaries of each of the posited units are compared. They are expected to reflect structural organization to some degree, but the targets at the right edge of the potential units in this dataset could also express the notion of completeness, or finality, particularly as all units smaller than prosodic sentences are non-final.

The results are shown in Table 3-13. The differences in final velocity in the last syllables of the units are significant, as revealed in the ANOVA test ($F(4, 1294), 14.918, p=0.000$). The values suggest two major subgroupings: prosodic words and phrasal constituents have velocities around -10st/s, while IUs and prosodic sentences have values around 2st/s. This suggests different underlying targets: the low negatives in the smaller units suggesting [mid] targets, and the low positive values of the larger units [low] underlying targets (see §2.3.3 for the interpretation of targets values).

Units	N	Mean final velocity (st/s)	Std. Dev
ProsW	204	-9.18	20.01
PhrC	281	-8.42	18.48
IU	216	1.94	22.73
ProsS	206	3.08	21.27

Table 3-13 The mean velocity measurement in the final syllable in each of the potential units.

The suggested [low] targets in larger units and [mid] targets in smaller units are congruent with their positions in the utterances as smaller units are indeed expected not to bear any marking of finality.

The investigation of final lowering leads us to consider another cue, that of phonation events, which occurs regularly in this dataset.

3.3.3.4.3 Phonation events

In Jaminjung, speech is sometimes still perceptible after the last calculated point of the pitch track. The aperiodicity in the signal is associated with either breathy or creaky phonation and is believed to be a potential boundary cue.

In our dataset, the presence or absence of breathy and creaky phonation is identified firstly through perception, then using the waveform, pitch tracker, and spectrogram. Figure 3-29 below shows a spectrogram of breathy and creaky phonation in Jaminjung. The breathy part occurs at the end of the first IU, on the final syllable, *na*, which is also considerably lengthened. The pitch track ends about midway through the syllable. The

creaky phonation is found sentence-finally, in the last syllable, *ra*. In this case, the pitch track ends just after the onset of the syllable.

(42)

<bottle>=marlang janju buny-ngawu= <na:>... gara
bottle=GIVEN DEM 3du:3sg-see.PST=NOW nothing

'The two looked at that bottle, nothing!'
[CP:ES96_018_02]

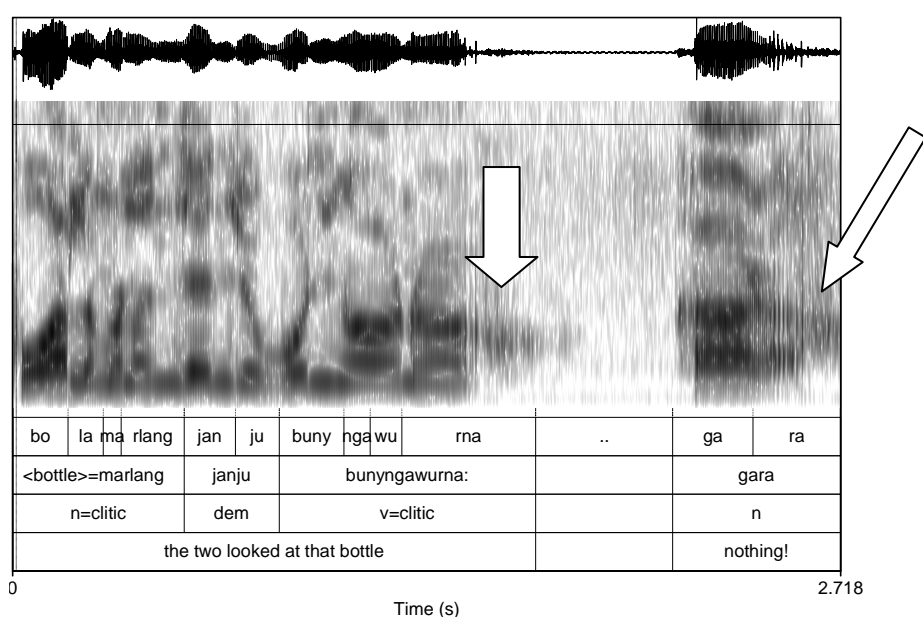


Figure 3-29 Spectrogram showing breathiness and creakiness.

Phonation events do not occur at the end of prosodic words, and very rarely in phrasal constituents, as shown in Table 3-14 below. Breathiness occurs predominantly in IUs (71% of occurrences), while creakiness is associated with the boundary of prosodic sentences (89%). The correlation between creakiness and type of unit is verified with a Chi-square test, which shows a strong correlation ($p = .000$).

	Phonation	PhrC	IU	ProsS	Total
creakiness	Count	0	12	96	108
	% within phonation	.0%	11.1%	88.9%	100.0%
	% total	.0%	5.1%	40.9%	.0%
breathiness	Count	1	15	5	21
	% within phonation	4.8%	71.4%	23.8%	100.0%
	% total	100.0%	55.6%	5.0%	16.3%

Table 3-14 Total number units that end with creakiness and breathiness in each text over the total number of units (1st line), and percentages (2nd line).

Creakiness being associated with low pitch, the average pitch of final syllables in prosodic sentences with and without creakiness is compared, resulting in averages of 136.23Hz and 152.74Hz respectively, which almost reaches statistical significance ($F(1, 162) = 3.738, p = .055$). Unsurprisingly, syllables where creakiness occurs have a lower pitch than those without.

The rates of occurrences of phonation events make them a good cue for boundary marking: breathiness is more likely to occur at the end of IUs, and creakiness at the end of prosodic sentences.

3.3.3.5 Summary: cues at right-edge boundaries

In this section, the cues marking the right-edge boundaries were tested, including firstly, durational cues (final syllable lengthening and affinity index) and secondly, pitch and phonation events such as F0 lowering, pitch target, creakiness and breathiness. The findings can be summarized as follows.

There is a non-linear increase in the absolute duration values of final syllables from the smaller to the larger units which can be expressed as:

Phrasal Constituents < Prosodic words < Prosodic sentences < Intonation Units

The measure of the relative duration of the final syllables indicates that lengthening occurs in IUs and prosodic sentences, but not in smaller units. Moreover, the ratios for the IUs are slightly higher than those of prosodic sentences. Pauses only occur after IUs and prosodic sentences and not after the smaller units. The measurements of the affinity index show a marked difference between the first two units (prosodic words, phrasal constituents), IUs, and prosodic sentences. When coverbs are removed from the prosodic words subset, because they are the only word class that can be monosyllabic and accept some specific consonant clusters which can affect the duration of the syllables, the IOI (inter onset interval) points to the expected distinction between:

Prosodic words / Phrasal Constituents < Intonation Units < Prosodic sentences

As to the pitch related correlates, F0 lowering occurs in larger units, possibly in phrasal constituents, but most convincingly in IUs and prosodic sentences. Final velocity values suggest that larger units receive [low] targets, and smaller units are marked with [mid] targets, suggesting that the [low] target could be interpreted as signaling finality. The lowering leads to the occurrence of phonation events, at least in the case of creakiness which is found at the end of prosodic sentences. However, the occurrence of breathiness at the end of IUs suggests that both phonation events may be boundary markers at the disposal of speakers, and not solely a phonetic consequence of pitch lowering.

3.3.4 Left-edge boundaries

The prosodic cues associated with left-edge boundaries are usually tonal in nature, that is, they are associated with pitch variations. The left boundary is particularly interesting in Jaminjung as most observers perceive the first syllables of words as stressed (Schultze-Berndt 2000:42) and it is assumed that *stress* serves a demarcative function (§3.2.2.4). The following section aims at delimiting its domain and endeavours to find the prosodic parameters that encode it. Pitch related phenomena are measured, notably pitch reset and final velocity (as the best indicator of the underlying pitch targets).

3.3.4.1 Pitch reset

In this study, pitch reset is defined as the difference in pitch range between two adjacent units. Pitch reset is calculated by subtracting the mean F0 of the last syllable of a unit from that of the following syllable, as formulated below:

$$\text{F0 Reset} = \text{MeanF0 LastSyll} - \text{MeanF0 Next Syll}$$

It is well-known that for the study of intonation, pitch distances are more relevant than absolute pitch: the same melody can be recognized in different pitch ranges, for example those of a male and a female speaker.

This is why the semitones scale, one possible log scale derived from the Hertz scale, is often used in prosodic research. As the measurements of the syllables' mean pitch is calculated in Hz in this study, the calculations of the pitch resets are also expressed in Hz. It is understood that an absolute measure may not be the best representation of pitch differences, particularly in a dataset that comprises many speakers. They are presented nevertheless as they point to interesting differences between the units.

The results are shown in Table 3-15. The value of the reset for the prosodic words is negative at -1.65Hz, and that of the phrasal constituents is a modest 2.00Hz. On the other hand, IUs and prosodic sentences both trigger salient resets, of 11.21Hz and 22.47Hz, respectively. These results are statistically significant ($F(3, 920) = 21.808, p = .000$). While prosodic words continue the line of declination, the very small positive value for the phrasal constituents indicates that the line is broken and there is a reset, albeit a much reduced one. Post hoc tests indeed show that the difference in the values of prosodic words and phrasal constituents is not significant.

Unit	Mean pitch reset (Hz)	Std. Dev	N
ProsW	-1.65	18.41	217
PhrC	2.00	20.76	282
IU	11.21	37.21	224
ProsS	22.47	52.32	201

Table 3-15 Value of pitch reset for each of the potential unit.

The resets are illustrated in the example sentence (43) below (Figure 3-30), showing two IUs, separated by a short pause. The reset at the beginning of the second IU is fairly important. Resets at the left edge of phrasal constituents and prosodic words are much less evident, the line of declination being only slightly interrupted. The exception is with the last constituent, *malajagugun* 'goanna', which bears the main focal accent in the IU (§4.1.3).

(43)

[[[ngayug] prosW]phrC [[ngiya] prosW]phrC
1SG PROX

[[ganba=ga]prosW]phrC]]IU
3sg:1sg-BITE.PST=YOU.KNOW

‘Me, I was bitten here’

[[[[gurrija]prosW [ngagba]prosW]phrC
digging 1SG-be.PST

[[malajagu=gun]prosW]phrC]]IU]prosS
goanna=CONTR

‘I was digging for a goanna’

[IP:ES97_03_02]

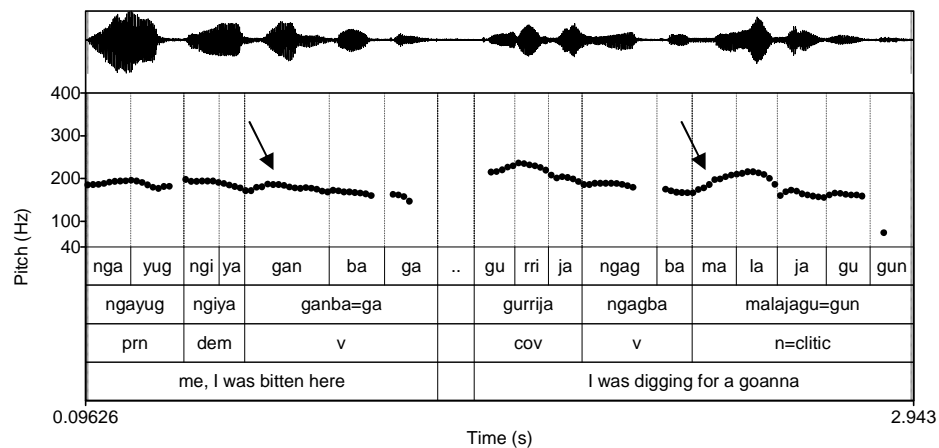


Figure 3-30 The onset of prosodic words, shown by the arrows, where the pitch reset is minimal, but still breaks the line of declination.

In brief, prosodic words and phrases just break the line of declination of the utterance, while IUs and prosodic sentences both trigger a reset of increasing value.

3.3.4.2 Initial high tone

The perception of *stress* on the initial syllable and the observation of the F0 tracks in the Jaminjung dataset lead us to posit that (at least one of) the units start with an underlying [high] pitch target.

This is illustrated in example (44) and Figure 3-31 below, where the pitch peaks which are hypothesized [high] targets are marked by an arrow.

The peak does not always occur on the first syllable, a phenomenon which is associated with the realisation of focus (§ 4.2.2.3.3). This section seeks to establish whether such a target really occurs at the left boundary, and what unit constitutes its domain.

(44)

yinju=biyang wurrgba gan-u-ga
PROX now chuck 3sg:3sg-take.PST

wirib=marlang=biyang jamurrurun yugung
given=GIVEN=now below-LOC.ABL run

gani-yu=nu
3sg:3sg-say/do.PST=3SG.OBL

'Here now, he took and chucked him , the dog ran at the bottom.'
[CP:ES96_018_02]

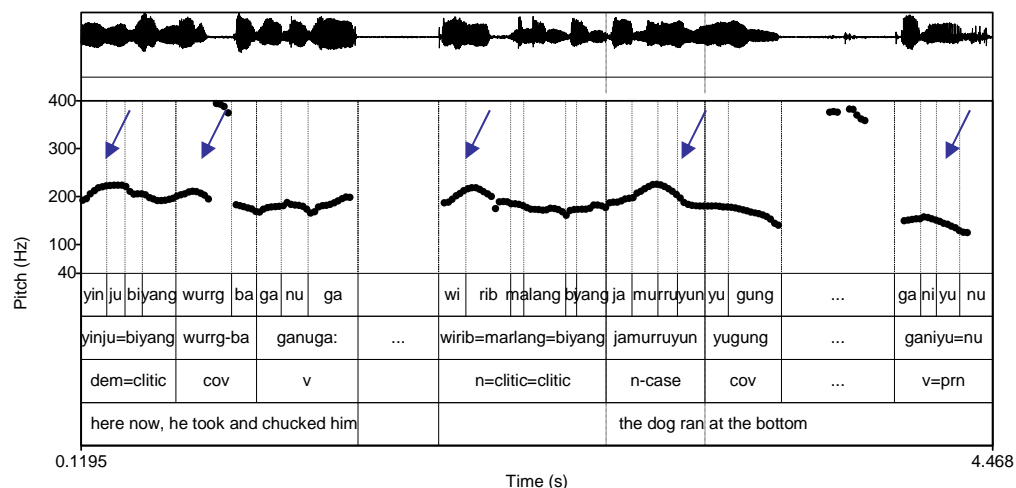


Figure 3-31 The hypothesised high tones in an utterance in Jaminjung. The pitch events under investigation are indicated by arrows.

The first and second syllables of each unit are measured, this time with a filter isolating units that do not share a leftward boundary with a larger unit. Final velocity values are used as an indication of the underlying targets. The patterns are confirmed by a measure of the mean F0 of the syllables at the left boundary.

The results are shown in Figure 3-32 below. The values of the final velocity of the first syllable are very close in prosodic words, phrasal constituents and IUs, hovering around -3st/s, while those of prosodic sentences are slightly higher, at 5.44st/s. In the second syllables, the smaller units have values around -6st/s, but those of IUs fall sharply to -12.01st/s,

close to the -9.95st/s of the prosodic sentences. Over the two syllables, the patterns are similar in prosodic words and phrasal constituents, and in IUs and prosodic sentences. The pitch targets of the prosodic words are probably [mid] targets, while those of the larger units are probably a [high] and a [fall]. These targets, however, are not understood to mark the unit of grouping but rather focus, as discussed further below. These results are statistically significant ($F(3, 1755) = 3.726, p=.011$). Post hoc tests confirm that the only significant difference is between the prosodic sentences and the other three units.

Unit	Mean final Velocity (st/s)		St. dev		N	
	1 st syl	2 nd syl	1 st syl	2 nd syl	1 st syl	2 nd syl
ProsW	-4.13	-7.72	21.35	22.97	214	195
ProsP	-3.66	-5.24	19.93	21.48	277	237
IU	-2.23	-12.01	22.38	20.61	234	205
ProsS	5.44	-9.95	25.83	21.91	209	184

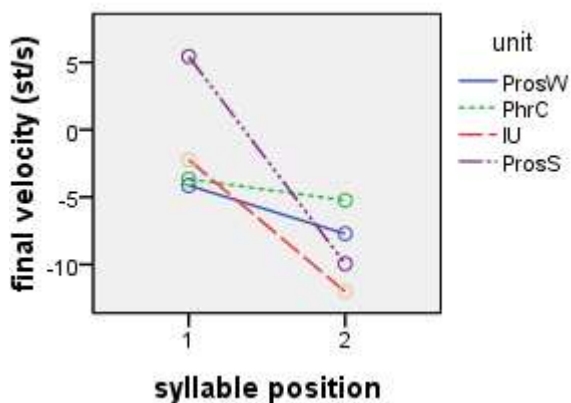


Figure 3-32 Final velocity in the first syllables of each unit.

To confirm the findings of the velocity measures, the mean F0 of the syllables is investigated. Similar patterns are found in prosodic words and phrasal constituents which have a very slightly higher F0 in the first syllable than in the second, while in IUs and prosodic sentences there is a very slight rise. The differences are significant ($F(3, 1748) = 5.492, p=.001$), particularly that of the phrasal constituents to the other units. The results are shown in Figure 3-33.

Unit	Mean F0 (Hz)		St. dev		N	
	1st syl	2nd syl	1st syl	2nd syl	1st syl	2nd syl
ProsW	172.83	170.64	38.84	41.79	211	192
ProsP	166.40	164.03	38.82	40.81	276	237
IU	174.74	175.12	41.83	47.74	233	208
ProsS	174.79	175.30	45.72	48.10	206	185

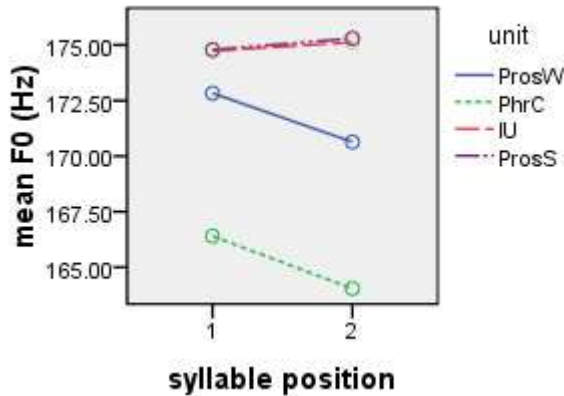


Figure 3-33 Mean F0 in the first 2 syllables of the units.

These findings do suggest that the pitch peaks which are hypothesized to demarcate prosodic units in Jaminjung with a [high] target are indeed found in IUs and prosodic sentences, but not in the smaller units. However, because the initial syllables at the left edge of the larger units are found to be the locus for the encoding of focus and topic (see ch. 4), it is not possible to associate specific underlying targets to the encoding of the grouping function.

To disentangle the correlates used to encode stress in Jaminjung, it is necessary to search for patterns that would distinguish the first syllables of words, where stress is expected to occur, based on perception, to other syllables in the same words, as well as to those of other units. The measurements of excursion size, duration, and intensity³⁴ of the first and second syllables of each unit are presented here. The results are shown graphically in Figure 3-34.

³⁴ This measure is used as an indication only, for the methodological reasons mentioned earlier (§2.4).

The excursion sizes, that is to say the difference between the minimum and maximum in pitch within the syllables, suggest a gradation on values from smaller to larger units, with the very notable exception of IUs which have more salient excursions on their first syllables than even in those of prosodic sentences. ($F(3, 1767) = 4.511, p = .004$). Prosodic words have excursion sizes that are smaller in the first than in the second syllable, indicating that this correlate is probably not a contender for the encoding of stress.

The duration of the first syllables is also measured. The only statistically significant difference is between IUs and prosodic words ($F(3, 1768) = 3.256, p = .021$). The differences in durations between the first and second syllables are more marked in prosodic words and phrasal constituents than in IUs and prosodic sentences, thus duration could be interpreted as serving to encode stress.

Both excursion sizes and duration are understood as a further encoding of focus, discussed at length in chapter 4.

The values for the correlate intensity are also presented. The results of the test suggest that, notwithstanding the unreliability of the recording conditions, this correlate could be an important cue for stress. There is a gradation in the measurements from the smaller to the larger units and the strongest statistical significance is found to occur, differentiating all four types of unit ($F(3, 1768) = 29.357, p = .000$).

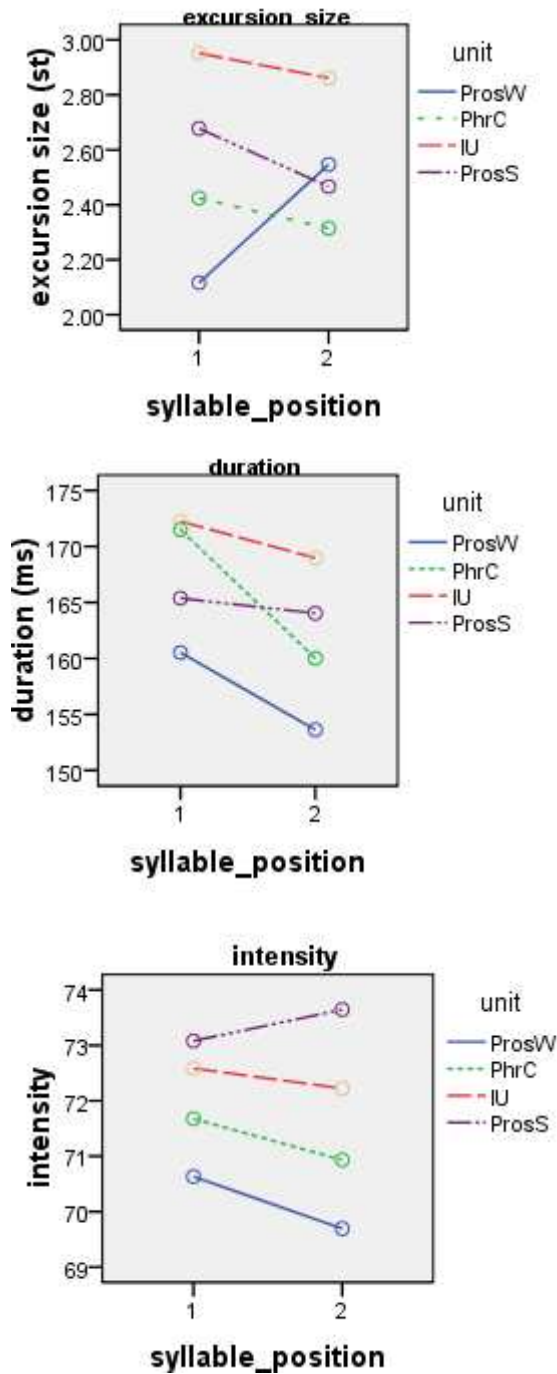


Figure 3-34 Excursion size, duration and intensity measurements of the first syllables of the units.

The findings of the correlates at the left boundary can be summarized as follow. Pitch resets occur in at least the IUs and prosodic sentences, and less convincingly in phrasal constituents. Prosodic words do not trigger pitch reset.

The values of final velocity suggest [mid] targets in smaller units. Those in IUs and prosodic sentences suggest a possible [high] underlying target, but this target probably does not serve as an encoding of the demarcative function, but is rather associated with the marking of topic focus (the former marked by a [high] target and the latter by a [fall] or [high] and [fall] sequence of targets, see ch. 4).

The observed pitch movement on the first syllable is thus more a reflection of the pitch reset and not the application of a specific target. As to stress, the results suggest that its domain is the word, rather than the phrasal constituents, and that its most likely cues are a longer duration and higher intensity. Further research would be needed to establish its exact correlates.

3.4 Conclusion: grouping

The first part of this chapter was devoted to presenting the phonological and morpho-syntactic characteristics of Jaminjung, establishing the basis for the set of prosodic units tested in the second part. Four units were defined:

- (1) Prosodic words which consist of content words (nominals, free pronouns, coverbs, verbs, interjections) and their associated bound morphemes (case markers, pronominal prefixes, TAM suffixes, derivational markers, and clitics).
- (2) Phrasal constituents which correspond to NPs, complex predicates (coverb and inflected verb or inflected verb), non-finite predicates (coverb without an inflected verb), particle groups (particle + NP or particle + CP).
- (3) IUs which correspond for the most part to clauses, but can also be dislocated elements (NPs, coverbs, etc), or interjections.
- (4) Prosodic sentences which consist of semantically connected independent clauses or adjacent clauses, usually a main clause and a subordinate clause or a dislocated element, etc.

Cues for the units were sought at both the right and left edge of the units. At the right edge they were: final lengthening, final lowering and phonation events; at the left edge, pitch reset and a specific underlying

target. The findings from the measurements are summarized in the following descriptions of the units.

Prosodic words are not set off by a pitch reset. The mean pitch of their first syllables is slightly lower than that of the final syllable of the unit preceding it, thus continuing the declination line. Their first syllables have a [mid] target and are demarcated by stress, probably cues by longer duration and higher intensity. As to their right boundaries, the final syllables are not lengthened, do not undergo final lowering, and have velocities that suggest [mid] tones. They are not followed by a pause, which lead them to have very low IOI values.

Phrasal constituents do trigger a small pitch reset, with the consequence that the declination line appears to be halted. Their first syllables usually receive a [mid] target. Further evidence for the phrasal status is provided by the observed prominences on the first syllables of phrases in sentence focus constructions (see §4.3). Such constructions are not partitioned into topics and comments, and still display a pitch movement on the first syllable of the phrasal constituents. As these movements cannot serve to mark information structure categories, they are thought to be solely associated with the grouping function.

At the right edge of the phrasal constituents, final syllables are not lengthened, but may, in a few cases, be followed by a pause, which is reflected in the slightly higher values of their affinity index. They also display some final lowering, but are not associated with phonation events. Their final underlying target is more likely to be a [mid] target than a [low] target.

Prosodic words and phrasal constituents have similar prosodic encodings, and apart from the distinction based on phonological grounds, i.e. prosodic words being the domain of phonological processes such as lenition (§3.2.2.3.2), evidence for their constituting different units comes from their difference in pitch reset mentioned above and from the affinity index measurements. Prosodic words have lower IOI values than phrasal constituents.

IUs and prosodic sentences are both set off by substantial pitch resets, that of IUs being smaller. The final velocity values in IUs and

prosodic sentences do suggest a possible [high] underlying target, but this target is taken not to have a demarcative function. I will argue later that this [high] target often found on the first syllable of IUs and prosodic sentences is associated with the marking of focus which is more likely to occur there, or that of the topics (ch. 4).

At the right boundary, both larger units are lengthened. Interestingly, the values of IUs' ratios are slightly greater than those of prosodic sentences, results which may be explained by the presence of IUs with an overall level contour in the dataset. These contours are usually not found in sentence-final position and are marked with longer durations in their final syllables.

IUs are usually followed by a pause, which is shorter than the pauses following sentences, thus accounting for the lower values of the affinity index of IUs. In other words, the final syllable in IUs is more likely to be longer than that of prosodic sentences, but the pauses that follow are usually shorter. IUs may end in breathy phonation, and prosodic sentences in creaky phonation. Final lowering occurs in both IUs and prosodic sentences, and both units have final syllables that receive a [low] target.

In short, the units are distinguished at their left boundaries by pitch resets which increase from unit to unit, a finding similar to that of Schuetze-Coburn et al. (1991) for American English. Units are not encoded with a specific pitch target; the observed movement in pitch is the effect of the pitch reset.

At their right boundaries, prosodic words are not lengthened, phrasal constituents only slightly, and neither have notably F0 lowering. This shows that they do not receive the marking usually associated with finality, rightly so, as the units tested were not in final positions. However, as the measures often display a gradience, it could be argued that prosody does not mark finality in a categorical manner (at least not in a binary final/non-final manner).

The strongest cue for the identification of the units posited in this chapter is the affinity index, which clearly delimits the four units. The values of the IOI of the units indicate that strong dependencies between immediately adjacent units weaken the boundary between them.

Finally, measurements on the tokens in this dataset confirm that intensity, as unreliable as the measurements may be in the context of this research, is probably the strongest marker of the perceived stress in first syllables of prosodic words. This would need be tested with more controlled data, if possible, in further research.

Chapter 4 Marking information structure

In this chapter, a quantitative study investigating the effects of information structural context on the prosodic realization of declarative sentences in Jaminjung is presented. As pointed out by Beaver et al. (2004), it is not uncommon in prosodic research to rely on semantic criteria to provide evidence for posited phonological distinctions. “If two productions are consistently associated with distinct meanings, then the phonetic difference between those productions must be phonologically significant.” The notions of information structure provide such an anchoring point.

The study of information structure (IS hereafter) interacts with many subfields of linguistic inquiry. As information structure can be marked solely by prosodic features, semanticists and syntacticians are forced to call upon phonologists and phoneticians to provide a clear account of the features used. Prosodists, for their part, agree that marking focus and other IS categories is an important function of intonation.

The central argument here is that in Jaminjung topics and focus domains are marked by prosodic means. Evidence of their specific encoding will be provided through measurements of the parameters of pitch and duration over the syllables.

This chapter is organized as follows: in section 4.1, a review of the literature on information structure and prosody is presented, including definitions of topics and foci used in this chapter. The terminology and the definitions for the categories employed in this analysis are based on a synthesis of existing proposals and can be regarded as fairly widely accepted. Care was given in the preparation of the Jaminjung dataset to select examples that correspond as closely as possible to these definitions. In the next sections, the Jaminjung data is analysed - firstly, the realization of focus over three different focus domains, the predicate/comment domain (§4.2.2), to which the sentence/thetic domain, and the argument domain are compared (§4.3); different types of topic are examined in §4.5, given and new topics are compared. Section 4.6 is concerned with the encoding of contrast, both in focused argument and in topics; and section 4.8 concludes

with an overview of the prosodic encoding parameters of the information structure categories in Jaminjung.

4.1 Information structure and prosodic analysis

Information structure (IS) is now widely used to refer to grammatical or phonological phenomena which reflect the relation of an utterance with the context in which it takes place (see Halliday 1967, Chafe 1976, Lambrecht 1994, inter alia). Lambrecht (1994:6) defines IS as “that component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts”.

The pragmatic structure of a proposition reflects the speaker’s assumptions about an addressee’s state of knowledge at the time of an utterance, and also about the representations of discourse referents in the addressee’s mind. IS is based on the intuition that utterances are ‘about’ something that both the speaker and the hearer have some shared knowledge about (the topic), and that utterances contain information on this topic which the speaker can assume is new to the hearer (the comment/focus). Various dichotomies are used to define this relation: theme-rheme; topic-comment; topic-focus; background-focus; given-new. They are not equivalent, but all aim at explaining this basic intuition.

Languages of the world make use of various strategies to realize information structure, they can be:

- Intonational: type and placement of accents.
- Syntactic structure: e.g., cleft structures or fronting (‘topicalization’), left/right dislocation.
- Morphological marking: specific grammatical morphemes marking topical or focal status of a constituent e.g., particle ‘*wa*’ to indicate topics in Japanese.

Moreover, languages often use a combination of means, rarely only one, depending on their typological characteristics.

While there is no general agreement on the status of information structure in grammatical theory, many authors, including Lambrecht (1994:6-13) and Féry, Fanselow and Krifka (2007) assign information structure a place in the core grammar, of no lesser status and importance than e.g. the signaling of grammatical relations. This is the view I will follow here.

As to prosodic research, highlighting part of an utterance in a particular way in order to convey distinctions in informational value is one of the most widely recognized functions of intonation, indeed it may be one of its most robustly encoded non-lexical functions. Some researchers orient their studies towards understanding the prosodic realization of the topic and comment categories while others are interested in defining the correlates of the focused part(s) of the utterance which is mostly assumed to be minimally marked by a ‘pitch accent’. Others point to differences in the realization of new and given information (see Bauman et al. 2006, for an overview of the research addressing the theme-rheme, given-new, and background-focus dimensions).

The next section provides an overview of the terminology used in the analysis of the information categories of Jaminjung.

4.1.1 Topic/Comment

I will make use of the notions of topic and comment which pertain to the organization of utterances. According to Krifka (2006:43), the topic/comment structure is ‘a packaging phenomenon’ which may have its basis in the cognitive organization of the human brain. A well-established view in current linguistics relates the topic-comment division to the predication structure of the utterance, i.e. to the distinction of a predication base, or topic, and a comment on this topic.

4.1.1.1 Definitions

Topic is not used here in the sense of *discourse topic*, what part of a discourse is about, but rather in the sense of *sentence topic*, the entity which

is predicated about in a sentence (Lambrecht 1994: 117; Krifka 2006).

Gundel (1985:86) defines topics in this way:

“an entity, E, is the topic of a sentence, S, if in using S the speaker intends to increase the addressee's knowledge about, request information about, or otherwise get the addressee to act with respect to E”.

Topics may have different functions and different degrees of integration into the sentence, as shown in example (45), ranging from i) full integration, where they have a grammatical function in the main clause; to ii) partial integration, where they are realised outside the clause, but co-indexed with one of its elements; to iii) no integration, where they are neither within nor co-indexed to one of its elements (see in particular Maslova and Bernini 2006).

- (45) i) [My dog_{TOP}] was playing with a ball.
ii) [That other dog_{TOP}], it came from the other side and stole it.
iii) [As to this walk_{TOP}], there is no point in continuing.

A sentence can have more than one topic or, as is very relevant in the case of Jaminjung, sentences can lack an overt topic in which case they are either comments on an implicit (discourse-given) topic or sentences that are in focus in their entirety (orthetic sentences). Such cases have brought some authors to distinguish between ‘topic expression’ and ‘topic referent’, where the former is the linguistic expression of a topic and the latter refers to what it stands for, its denotatum (see Lambrecht 1994:127-131, Krifka 2006:5-6). Apart from the general ‘aboutness’ topic discussed so far, the following subtypes of topics will be useful in the following analysis: ‘new’ topics which highlight a change of topic; and ‘contrastive’ topics, when choosing among several possible referents. Contrastive topics are regarded as topics which at the same time have focal status, following Krifka (2006), since a contrastive topic typically implies that there are alternatives in the discourse.

In this chapter, only NP topics will be considered; whether other types of constituents can also be topics in Jaminjung will be a matter for investigation in further research.

The criteria used to identify sentence topics in the dataset are inspired by the guidelines set by the Interdisciplinary Studies on

Information Structure (ISIS) project in Potsdam (Dipper et al. 2007). Thus, only the following expressions can function as *aboutness* topic: referential NPs (i.e. definite descriptions and proper names), indefinite NPs with specific and generic interpretations, and indefinites in adverbially quantified sentences that show Quantificational Variability Effects³⁵, bare plurals with generic interpretations, and bare plurals in adverbially quantified sentences that show Quantificational Variability Effects.

Dipper et al. (2007:163-164) further suggest the following test to identify topics:

An NP *X* is the ‘aboutness topic’ of a sentence *S* containing *X* if

- *S* would be a natural continuation to the announcement *Let me tell you something about X*.
- *S* would be a good answer to the question *What about X?*
- *S* could be naturally transformed into the sentence *Concerning X, S'* or into the sentence *Concerning X, S'*, where *S'* differs from *S* only insofar as *X* has been replaced by a suitable pronoun.

In a bipartite view of sentences as the expression of information about a topic under discussion, the *comment* is the remainder of the sentence – corresponding to the predication that is made about the topic. I will employ here the definition of comment by Gundel (1988:210): “A predication, *P*(*), is the comment of a sentence *S*, if, in using *S* the speaker intends *P* to be assessed relative to the topic of *S*”.

4.1.1.2 Specific tunes associated with the topic and comment categories

For English, it has been claimed that topics end with a rising accent, and that the comment of a sentence typically ends in a falling accent

³⁵ “Quantificational Variability Effects can be defined as follows: An adverbially quantified sentence that contains an indefinite NP or a bare plural is roughly equivalent to a sentence where the combination Qadverb + indefinite NP/bare plural has been replaced by a quantificational NP with corresponding quantificational force” (Dipper et al. 2007:164).

(Bolinger 1986, Jackendoff 1972, Gundel 1978, Steedman 1991, Pierrehumbert and Hirschberg 1990, inter alia). The following example, from Pierrehumbert (1980:40) is given as illustration:

(46) What about Anna? Who did she come with?

Anna came with Manny.

H* LH% H* LL%

Anna is associated with the topic and receives an H* LH%³⁶ tune while *Manny* is associated with the comment and receives an H* LL% tune. Gundel et al. (1997) experimentally tested the often-quoted relationship between topic and comment structure and specific tunes, especially the assumption that a specific pitch accent or tune is associated with the focus (or comment) of an utterance, which is distinct from the pitch accent or tune associated with the topic. Their study used transcripts of radio news stories read by a professional newscaster. They found no compelling evidence for the assignment of specific tunes to the topic and comment categories. Most of their topics receive a rising tune (H* LH%), but their comments may receive both rising and falling tunes (H* LH% or H*LL%), depending of where they occur in the sentence. Further, the overwhelming majority of sentence final tones are low tones. If they are high they are more likely to be topics, but then again, most sentence final topics end in a low tone. This leads them to suggest that the rising and falling tunes (H* LH% and H*LL%) rather indicate continuation and non-continuation. This is important in the context of this research, highlighting the difficulty of mapping a function to a specific encoding, and suggesting a methodology that will ensure that factors such as ‘position in the sentence’ is taken into account.

³⁶ The H*-L and L-H* notation is adapted from Pierrehumbert (1980). The starred tone, H(igh) in this case, is associated with the strongest syllable of the pertinent domain. The non-starred tone, L(ow), is then associated with the syllable following or, respectively, preceding the strongest syllable (§2.2.3.5)

4.1.2 Focus

A note of terminological caution is warranted at this stage. The category of *focus* is amply referred to in the literature. It may be studied as a component of phonetics, phonology, syntax, semantics, pragmatics or even psychoacoustics (see Beaver et al. 2004:1-14, for a review) which can lead to confusion. Von Stechow (1999:84) summarizes various conceptions of focus thus:

Focus is understood as a semantic-pragmatic concept, indicating a certain relation of the focused expression to the discourse context. Jackendoff (1972) introduced the syntactic focus-feature F at the level of syntactic structure in order to describe the syntactic correlate to the semantic notion of focus. Thus, the focus concept is understood as a hypothetical construct that has a phonological correlate (the sentence accent), a semantic-pragmatic correlate (the new information), and a syntactic correlate (a syntactic feature [+F]). In the short history of the development of a comprehensive theory of focus, each correlate has been chosen from different schools of thought as a conceptual starting point.

I will use the term in its information structure sense, which is defined further below. To account for the focus systems in the languages of the world, and that of Jaminjung, two more parameters are needed, the *scope* of the focus function, that is, the part of the underlying clause structure which is in focus; and the *communicative point*, what pragmatic reasons underlie the assignment of focus to the relevant part of the underlying clause structure (Dik 1997:330). These will also be presented.

Focus can be informally defined as referring to the discriminating part(s), the most salient or highlighted or most foregrounded part(s) in the utterance. An expression is in focus if the speaker is choosing it from amongst a (possibly infinitely large) set of alternatives, thereby contrasting it with them.

A more technical definition of focus is given by Lambrecht (1994:213) as “[t]he semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition”.

Lambrecht (1994:213) uses the terms *presupposition* and *assertion* to refer to “the set of propositions lexicogrammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered” and “the proposition expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered”, respectively.

Some authors also distinguish focus from the notion of *background* which refers to the non-discriminating remaining part(s) of the utterance (see Krifka 2006). It is peripheral, and serves the purpose of specifying the way in which either the topic or the focus is to be understood. Relative clauses, subordinate clauses and adjuncts (adverbials and PPs) are often part of the background.

4.1.2.1 Prosodic correlates of the focused part(s) of the utterance.

The focus of an utterance, semantically its most salient part, is usually predicted to be the most prosodically prominent. In the languages of the world, a variety of prosodic resources are employed³⁷ to mark material as focused:

- the placement of the pitch accent (in AM terminology), and deaccenting elements in post-focal position, as in English (Ladd 1996);
- the placement of phrase boundaries which may either precede or follow the focused element. Indeed, pausing before an accented element is a well-known and wide-spread device in languages for drawing a listener’s attention to what is coming and thus highlighting its importance. Again, post-focal elements may be ‘dephrased’, that is, material following the focus is incorporated into the same phrase, as in Japanese;

³⁷ See also Büring (2009) for a review of the various strategies used to realize focus, including prosodic means.

- the expansion of pitch range on the focused elements and the narrowing of the post-focal pitch range, as in Mandarin Chinese (Xu 2005), or English (Xu and Xu 2005);
- and finally, duration (Flemming 2008, see also Dik1997:327). It was found that the syllable which is the focus exponent is lengthened in many languages, e.g. in English (Turk and Shattuck-Hufnagel 2000), in Mandarin (Xu and Xu 2005), in Dutch (Sluijter and van Heuven 1996), and in German (Kügler 2008).

As noted earlier for English, it is generally agreed that topic is marked with a rise in pitch (or L-H* in AM terminology), and focus by a fall (H-L*). More specifically, focus is marked by a ‘nuclear pitch accent’ on the primary stressed syllable which is usually the last accent in a phonological phrase. When the nuclear accent occurs early in the phrase, the remainder of the phrase is said to be ‘deaccented’ (see Büring 1999, and Ladd 2001, ch. 6 for a review). Interestingly, Hedberg and Sosa (2007) remark on the difficulty of identifying such specific patterns in natural speech in their study of American English: “In particular we deny that there is any prosodic category as distinctive as a ‘topic accent’ as opposed to a ‘focus accent’”. Gussenhoven (1999:45) proposes that semantic constituents defined as arguments, predicates, and modifiers, are marked for focus, rather than syntactic constituents. In his model of intonation, English has three pitch accents (fall (HL), fall-rise (HLH), rise (LH)), and nine additional pitch accents resulting from modifications in the basic pitch accents. He further proposes a Sentence Accent Assignment Rule (SAAR) which applies to all pitch accent types. According to the SAAR, every focused argument (subject and internal verbal complements), predicate (verbs, predicate nominals, predicate adjectives, etc.), and modifier (adverbials and other adjuncts) must be accented - with the one exception: a predicate adjacent to an argument can be unaccented. Consequently, a focus domain is defined as “any constituent whose focus can be marked with a single pitch accent”. He further proposes that the size of the excursion and the accent height is correlated with the degree of prominence.

Another important factor, referred to as alignment, describes the precise position of the accent in time within the accented syllable. More recently, Steedman (2003) describes the intonational patterns for themes and rhemes in English, in AM terms, as L+H* and H*, respectively. For Steedman, theme and rheme can both be partitioned into background and focus. He suggests that focus is marked by prominence in pitch compared to the background. The background is usually unaccented and can even be omitted entirely from conversations. In Steedman's analysis, the theme is to be thought of as that part of an utterance which connects it to the rest of the discourse and the rheme is that part of an utterance that advances the discussion by contributing novel information. In short, in Steedman's view, the overall intonation pattern is determined by the theme/rheme partition, and the actual placement of the pitch accents is determined by the focus/background partitioning in these categories (see Louwerse et al. in press, for an overview). Louwerse et al. confirmed this partitioning in a series of tests based on conversations elicited with Map tasks (Anderson et al. 1991). They found that the average pitch of the rheme (comment) in a turn is significantly higher than the average pitch of the phrasal theme (topic) of that turn.

Research conducted under the PENTA model umbrella is starting to yield interesting results. Wang and Xu (2006:1) have found specific correlates associated with the focused part of an utterance in English, notably an extended pitch range on the focused syllable and a compression of the same range in the post-focus part of the utterance while leaving the pitch range of pre-focus component neutral.

Post focus compression of pitch and intensity is thus found to be an encoding parameter of focus in English, but also in Mandarin, Cantonese (Wu and Xu 2010), and Hindi (Patil et al. 2008), however current research indicates that it may not occur in all languages. Parallel studies in Taiwanese and Taiwan Mandarin have shown that both lack post-focus compression (Chen et al. 2009).

There are also reports of languages where focus is not manifested through prosodic means such as Navajo, a language of the Athabaskan family, where each syllable of the word has a tonal target, and statements,

questions, and focus constructions (contrastive topics in our description) share almost identical pitch tracks (McDonough 2002); or Chitumbuka, an underdescribed Bantu language of Malawi, in which focus is not consistently signaled by prosody (Downing 2006).

For Australian languages, Heath (1984) describes “pregrammatical discourse categories” listed as Focus, Topic, Definiteness, Given and New, which may be marked intonationally (see Mushin and Baker 2008 for a review of the literature on discourse in Australian languages). Blake (1987: 155) proposes two common principles of discourse organization: (i) topic precedes comment and (ii) focus comes first. He further states that it is not unusual for phrases acting as topics to be placed outside the predication proper and be separated by an intonation break, either before or after, which could be represented schematically as: topic [focus + rest of comment]. He goes on to note that the initial position in an utterance is preferred for highlighting functions, such as in Warlpiri, where prominence-marking or focusing is accomplished by placing a word in phrase-initial position (Laughren et al. 1996; Mushin and Simpson 2005). Hale (1992) pointed to the pre-verbal position as the choice position for focus marking in the same language. Austin (2001) states that Jiwari also makes use of the initial position to introduce new topics, significant new information, or to make a contrast.

Bishop (2003:211) suggests that accent distribution in Bininj Gunwok (BGW), a term used to cover a large group of related dialects spoken in Western Arnhem Land, serves mainly a word delimitative function. She even raises the issue of whether there is any interaction between post-lexical accent assignment and information structure in BGW. She goes on to describe relative *pitch scaling* parameters, the global, or at least phrase-sized choice of pitch range, as the main means of encoding focus and other aspects of information structure in BGW. She also finds that there is no systematic use of deaccenting in BGW.

Indeed, the lack of evidence of a culminative function for accents and the use of pitch range extension on the focused element, or at least in part of the focused element, are reported for other Australian languages. For example, Chapman (2007) has found that there is no evidence in Warlpiri

for a culminative accent on words with either narrow or broad focus. He also suggests that a final rising intonation associated with pauses is used to mark off left dislocated topics, adjuncts, and words and phrases used in apposition. Singer (2006) notes that in Mawng, a non-Pama-Nyungan language of Northern Australia, all verbs and nouns usually receive some type of pitch accent. Thus, the fact that a word receives a pitch accent does not in itself tell about its information status.

Fletcher et al. (2002) in a qualitative study of a corpus of Kayardild, also a non-Pama-Nyungan language, note that expanding the range of the phrase-internal pitch target is a relatively common strategy used to mark focus. For Dalabon, another Australian language of the non-Pama-Nyungan family, Fletcher (2007:10) reports that the initial syllable of the word in focus clearly bears the main pitch prominence. Furthermore, the word in focus does not “necessarily have to be intonation phrase initial, although Dalabon does not have as much freedom as other Australian languages (e.g. Warlpiri) to move around grammatical words due to the complex form of the verbal word”. In another study of BGW and Dalabon, Fletcher and Evans (2002) found that syllables associated with ‘pitch accents’ are phonetically lengthened.

Informal observations of the Jaminjung data suggest that its speakers, just as those of the neighbouring languages in Northern Australia reported above, use prosodic means to emphasize elements they wish to highlight.

4.1.2.2. *Scope of focus*

Speakers may choose to emphasize constituents of various sizes. In the analysis of the encoding of focus in Jaminjung, I will test whether the encoding varies when the size of the focal constituents changes. The domains tested are those proposed in Lambrecht (1994). They are described here, as well as some prosodic research associated with each of the domains. In other words, focus as a special pragmatic status can be defined in terms of its scope, whether it concerns the truth value of the predicate, the whole sentence or only a constituent in the utterance. Lambrecht (1994:223) thus distinguishes three main types of constructions. He illustrates these types

with example sentences reproduced below from English, Italian, spoken French, and Japanese in which the small caps indicate the prosodic prominence:

- PREDICATE FOCUS (PF) is used when an entire verb phrase³⁸ is in focus. This type is claimed to be a universally unmarked type of focus correlating with the topic-comment structure as the unmarked pragmatic articulation (Lambrecht 1994:296). Roughly, the subject of the sentence corresponds to the topic and the remainder, is a comment on that topic.

(47) [context: What happened to your car?]

- a. My car/It broke DOWN.
- b. (La mia macchina) si è ROTTA.
- c. (Ma voiture) elle est en PANNE.
- d. (Kuruma wa) KOSHOO-shi-ta.

- ARGUMENT FOCUS (AF) is used when the focal constituent is a single constituent (subject, object or oblique), and the remainder of the sentence is a presupposed proposition. Argument-focus can be expressed through a number of strategies: by changing the stress, as in the English example (a); by inverting the order of the constituents, as in the first Italian examples (b); by using cleft constructions as in the second Italian example and the French example in (c); and finally, by using morpho-syntactic markers as in the Japanese example in (d). Argument-focus constructions are those most often treated in prosodic analyses of ‘focus’.

(48) [context: I heard your motorcycle broke down?]

- a. My CAR broke down.
- b. Si è rotta la mia MACCHINA.
E la mia MACCHINA che si è rotta.
- c. C'est ma VOITURE qui est en panne.
- d. KURUMA ga koshoo-shi-ta.

³⁸ With the understanding (mine) that Verb + Argument do not form a Verb Phrase in all languages.

The commonly used terms *broad focus* and *narrow focus* are avoided here because of their vagueness and because the term *narrow focus* is, in practice, often equated with argument focus (whether contrastive or not) discussed hereafter.

- SENTENCE FOCUS (SF) is used when an entire sentence is being focused. Such expressions are also termed *thetic* sentences (Sasse 1987). In SF, both the subject and the predicate are part of the focus domain, thus distinguishing it from both PF, in which the subject is the topic and hence not part of the focus domain, and from AF, where the predicate is part of a presupposed proposition and, as such, not part of the focus domain. Lambrecht suggests that SF needs to be minimally distinct from PF, namely in that the subject in a SF construction takes on properties normally associated with objects in a PF sentence (Lambrecht 2001:627).

- (49) [context: What happened?]
- a. My CAR broke down.
 - b. Mi si è rotta (ROTTA) la MACCHINA.
 - c. J'ai ma VOITURE qui est en PANNE.
 - d. KURUMA ga KOSHOO-shi-ta.

Importantly, Lambrecht (2001:336) states that these three types of focus structure correspond to three basic communicative functions: “that of predicating a property of a given topic (predicate focus: topic-comment); that of identifying an argument for a given proposition (argument focus: identificational function); and that of introducing a new discourse referent or reporting an event (sentence focus: presentational or event-reporting functions)”. Sentence focus or thetic statements appear “at any point in a text where information is not given about someone or something, but about an entire state of affairs” (Sasse 1987:535). In short, a defining criterion for thetic sentences is that they present information about situations rather than about entities.

Languages vary in the strategies they use to distinguish topic-comment and sentence focus. Some attention has been given to the prosodic encoding of thetic statements, particularly since Lambrecht. He proposes

that SF needs to be minimally distinct from PF, namely by detopicalizing the subject that would normally be interpreted as the topic in a PF sentence (Lambrecht 1994:235). This is conceptualised in the *Principle of subject-object neutralization*: in a SF construction, the subject tends to be grammatically coded with some or all of the prosodic and/or morphosyntactic features associated with the focal object in the corresponding PF construction. Languages that have a more rigid syntax and a more flexible focus structure move the accent, as in English. Conversely, languages with the reverse organization i.e. more flexible syntax (word order) and less flexible focus structure, as in Italian, move constituents instead of shifting the accent.

4.1.3 Contrast

Another notion I will rely on is that of *contrast*. The basic requirement for contrastiveness is the existence of highlighting. Constituents that are contrastive are understood to belong to a contextually given set out of which they are selected to the exclusion of at least some other members of the set. Both topics and foci can be interpreted contrastively (Kiss 1998). Thus topics can be ‘about something’, and can be contrastive; focus can be ‘information’ and contrastive as shown graphically in Table 4-1 (from Neeleman et al. 2009:1). Example (50a) illustrates contrastive topics (from Kiss 1998) and (50b) a contrastive argument in focus.

	Topic	Focus
Non-contrastive	aboutness topic [topic]	new information focus [focus]
Contrastive	contrastive topic [topic, contrast]	contrastive focus [focus, contrast]

Table 4-1A typology of topic and focus constructions (from Neeleman et al. 2009:01)

(50) a. Where do Anna, Kati and Mikkor live?

ANNA_{Contrastive topic}, she lives here.

b. What did Peter buy, a book or a magazine?

Peter bought A BOOK_{Contrastive focus}.

To understand the contrastive and non-contrastive focus distinction, it is useful to refer to Dik (1989), who proposed a typology of focus according to the communicative point expressed. He first defines new (or completive), and contrastive categories, which can be further subdivided into replacing, expanding, restricting, and selecting. For the purpose of this research, I will use *information* focus interchangeably for Dik's new focus.

Information focus contributes information that is new or 'context-incrementing' (Drubig and Shaffar 2001:1079) e.g. in an answer to a constituent question. *Contrastive focus* 'identifies a subset within a set of contextually given alternatives' (Drubig and Shaffar 2001:1079; Kiss 1998; Götze et al. 2007); it is often syntactically and morphologically marked with respect to information focus. Examples are given in (51) and (52); both involve a contrast between the constituent in focus (capitalized) and an alternative piece of information which is not explicitly presented (or presupposed) in (51), but is contained in the question in (52). This type of focus construction, termed replacing or corrective focus, is fairly common in the Jaminjung data, where speakers often seek the best way to express some concept or idea.

(51) YOU_{Contrastive Focus} are the murderer. (Sherlock Holmes)

(52) John bought apples?
No, he bought BANANAS_{Replacing focus}

Topics can also be contrastive. As pointed out by Umbach for English (2001:176):

A contrastive focus may occur in any position in a sentence. If, however, it occurs in the topic part, it represents a contrastive topic and will typically be marked with a rising accent. Being in the topic position, a contrastive topic refers to an entity the speaker wants to talk about. Moreover, due to its contrastiveness, it indicates that there exist alternatives the speaker wants to talk about, too.

Example (53) below, also from Umbach, illustrates this (capitalized words indicate a focus accent).

- (53) Adam: What did the children do today?
Ben: [The small children]_{contrastive topic} stayed at HOME and/but [the bigger ones]_{contrastive topic} went to the ZOO.

It is usually held that both contrastive topics and foci are marked by a special intonation (at least since Bolinger 1961:87), typically described as a fall-rise accent in English (Büring 1997:69). However, whether topics and foci receive the same prosodic encoding cross-linguistically is still undecided. Braun and Ladd (2003) have shown that in German, sentence-initial themes in contrastive contexts are prosodically distinguished from those in non-contrastive contexts by peak height and alignment, range and duration of the rise, and duration of the stressed vowel. They also suggest that contrastiveness may not be categorical but gradient. This is supported by Zimmermann (2007), who points out that the intonation of contrastive focus differs only gradually from information focus in intonation languages, while in languages such as Hungarian or Finnish, in which contrastive elements are realized in a particular syntactic position, the opposite appears to be true.

In Australian languages, Bishop and Fletcher (2005) suggest that a rising accent (L+H*) (for certain varieties) tends to signify contrastive stress in Bini Gun-wok, without clarifying whether this applies to topics or focused arguments specifically.

For Mawng, Singer (2006) notes that the two functional categories of contrastiveness and information focus are signaled by being the most prominent elements in the utterance. The most prominent accents in contrastive NPs are realised as a very steep rise and fall. More specifically, Hellmuth et al. (2007) describe a special focus accent distinguished from a neutral counterpart in scaling (by means of steep pitch excursions in contexts of contrastive focus), but not in alignment, which they interpret as a hyperarticulated version of the neutral accent.

4.1.4 Given / New

The analysis of the prosodic encoding of information structure in Jaminjung will also make use of the notions of *given and new*, a division first proposed by the linguists of the Prague School. The notions are often used interchangeably with topic and focus; strictly speaking, the notions refer to cognitive *activation states* of referents in utterance and need to be differentiated from these categories. Chafe (1976) actually distinguishes three levels of activation: given, new and accessible information. *Given* refers to the part of an utterance that the speaker expects the addressee already knows or at least should already know; while *new* refers to the part of the utterance that is new, or at least new relative to the discourse in question, and relative to the given. *Accessible* information is semiactive and becomes activated at the time of speaking. Lambrecht (1994:109) further divides accessible information into textually, inferentially and situationally accessible information, taking the nature of the context into account. While topic referents are often discourse-active (given) or at least accessible to the hearer, and the referents of focal constituents are often unidentifiable (discourse-new), the association does not hold in all situations. However, for the purpose of this study, the given/new notions will be used in conjunction with the category of topic.

In intonation research, it is often assumed that the different states of activation of information can be associated with specific prosodic realizations, following Chafe (1994:75), who proposes that “both new and accessible information are expressed with accented full noun phrases, whereas given information is expressed in a more attenuated way”. Ladd (1996) argues along the same line, proposing that, at least for West Germanic languages like English and German, new information is marked by a pitch accent, while given information is deaccented. He further considers the possibility that the distinction may not be categorical but gradient, following Chafe. More specifically, in American English, Pierrehumbert and Hirschberg (1990) claim that given information is assigned an L* accent (if accented at all), new information is marked by an H*, and accessible information is marked by either an H* pitch accent

followed by a !H* (downstepped H*, inferable item) or an H+!H* (predication already mutually believed by speaker and listener). For German, Kohler proposes that new information is marked by a medial peak on an accented syllable; and activated information is marked by an early peak contour, a pitch movement with a peak immediately preceding a considerably lower accented syllable. These contours do resemble those proposed for American English.

4.2 The categories analysed in Jaminjung: main hypotheses

The categories discussed in §4.1 are conceptually fairly straightforward, but can prove difficult to apply when natural language is analysed. Further, as the aim of this analysis is to provide a general description of the prosodic encoding of information structure in Jaminjung, I will make use of the elements described above to define the categories in the analysis.

I will first seek to establish the prosodic means used to encode focus in Jaminjung. I will test whether the encoding varies in the three focus domains described by Lambrecht: the predicate focus, the sentence focus and the argument focus. I prefer the term ‘comment’ focus domain to ‘predicate’ focus domain, as the former refer to an information structure category and the latter to a syntactic category. The comment focus domain will be analysed in detail and used as a point of comparison for the two other domains. Following Baumann et al. (2006), I predict that prominence-lending cues increase as the focus domain narrows, so that the argument focus domain should be the most marked and the sentence focus domain the least marked. As reported in the previous section, prosodic research has addressed various aspects of the question, i.e. in distinguishing between ‘broad’ and ‘narrow’ focus, the former more or less corresponding to the topic/comment structure with a focus on the comment; and the latter with argument focus, where only one element in the comment is emphasized. I

propose a systematic comparison of the three domains to verify whether and how the encoding differs.

After looking at the prosodic realization of focus in Jaminjung, I will examine topics, using as a benchmark the given topics in the utterances with a topic/comment structure. Jaminjung, as many other Australian languages (see Mushin and Simpson 2005, and Mushin 2005), is said to have ‘free’ word order. Given topics may either precede or follow the comment and I will test whether the position of the given topic triggers variation in their encoding. I will also check if the use of the clitic =*marlang*, used to mark the givenness of referents, changes their encoding.

I will make use of the degree of activation *new* to describe the encoding of topics. They are usually found at the left periphery of the utterance, either integrated in the IU or as a separate element (left detached). Whether the latter should be interpreted as topics rather than as (presentational) sentence focus constructions (Lambrecht 1994:184) will be tested through a comparison with the dataset analysed for the sentence focus domain. The encoding of given and new topics will also be compared based on the hypothesis that both subtypes are minimally distinguished by prosodic means.

Finally, I will also explore the feature of contrast. I will consider whether differences in the ‘communicative point’ of the focus are reflected in its prosodic encoding within the argument focus domain. I will distinguish between arguments that serve as information focus and those that serve as contrastive focus.

I will describe the encoding of contrastive topics, and compare it to that of given and new topics. I will also compare it to that of contrastive arguments, in order to find whether contrastiveness receives a specific encoding in Jaminjung.

I will also test whether the use of a morphological marker, the clitic =*gun*, marking contrast, used with both focused arguments and topics³⁹, affects the prosodic encoding.

³⁹ The clitic is also found with verbs (§3.2.3.1.6)

4.2.1 Analysis: Focus domains

In this section, I will assume that Information Structure is manifested overtly through prosody in Jaminjung. I will assume that speakers make some words and phrases more prominent than others, that these are perceptually more salient to the listener and are presumably employed at least in part to draw the listener's attention to informationally salient words. It is assumed that the function of marking part of an utterance as more important, as focused, is achieved through making one of the syllables more prominent than others (in Jaminjung). I shall assume that the functional differences associated with focus realization are reflected in variations in the acoustic correlates of the perceived prominent syllables.

I will examine the three domains of application of focus with the aim of showing that each of these domains makes use of different prosodic parameters.

4.2.1.1 Methodology

The IU is the domain of investigation. The methodology follows the guidelines described previously (§2.4), firstly by selecting a set of utterances that conform to the criteria specified for each category. Importantly, the labeling of perceived prominent syllables in the dataset used in this analysis is done before any measurements or analysis. In some cases a primary and a secondary prominence is assigned.

As native speakers could not be called upon to participate in the labelling of prominences for practical reasons such as their being too far away at the time of the analysis and lack of literacy skills, they were assigned by two labelers, the author and a non-linguist English speaker. An explicit selection procedure was applied: only IUs where there was consensus on the prominences between the labelers were retained.

To check on our judgments, a test was conducted with a linguist fluent speaker of a neighbouring Australian language. Thirty IUs were selected randomly from the dataset, the subject was asked to indicate the most prominent syllable; a secondary prominence could be assigned if that was deemed necessary. The results were marked as follows: 3 marks were given when all listeners agreed on the prominence, 2 marks when the same

syllables were selected but given a different degree of prominence (primary and secondary). In 23 of the 30 IUs (77%), there was perfect agreement on the prominence, but the percentage rises to 87% in those cases where only the degree of assigned prominence diverged. There is only one case where the listeners heard a prominence on entirely different syllables.

The labeling of prominences and segmentation of the utterances into syllables is followed by coding for specific functions or categories. Syllable positions are labeled in relation to the most prominent syllable⁴⁰.

For all focus domains, the locus of prominences and their morpho-syntactic correspondence are examined. The analysis of the prominence in the comment focus domain includes tests to assess whether it receives a specific pitch target through an examination of the correlates of final velocity and F0. As mentioned in §2.3.3, the evaluation of pitch targets is based on the final velocity measurements, but not uniquely. The examination of the pitch track, as well as auditory analysis is necessary, particularly for the [low] and [high] targets which may both display velocities between -5st/s and 5st/s.

The measures are compared according to word length and syntactic constituency. Finally, a test is done to compare the final velocity measurements in the prominent syllable in comments and the syllable that occupies the same position in topics. The correlate of mean pitch is then examined. The measurements are checked with a series of ANOVA tests with the correlates as dependent variables and syllable prominence as independent variable. Finally, excursion size and duration are similarly assessed. For the latter, syntactic constituency and word length are taken into consideration.

After this detailed evaluation of the parameters used to mark the prominent syllables in the comment focus domain, a comparison is made with the other domains: the sentence and arguments focus domains. As more than one prominence is heard in the sentence focus domains, the correlates of the mains and secondary prominences for this domain are

⁴⁰ The effects of conditioning factors e.g. segmental make-up, speech rate, and the like, can not be isolated in a study based on spontaneous speech such as this one.

presented prior to the comparison of the three domains. For the argument focus, the structure of the comments in which the prominence occurs is described, as well as the position of the word containing the prominence within the comments. Concerning the argument focus domain, the methodology developed for this analysis is based on the work of Xu (2005 *inter alia*) who examined the realisation of *broad* and *narrow* focus, the latter corresponding to the argument focus domain in question in this section. However, it is not possible, with the natural data used in this dataset, to reproduce the experiments that lead Xu (2005:234) to propose that pitch range of post-focus words is compressed extensively while the pitch range remains the same in pre-focus words (although tokens of pre-focus pitch lowering have also been observed in English). In the experiments in question, the speakers were required to read out prepared sentences in order to produce various focus domains. The encoding schemes of each of the domains were then compared.

4.2.2 Comment (predicate focus)

In this section, I aim to establish the encoding parameters of focus when its domain coincides with the entire comment. I will first present the hypothesis and dataset on which the analysis is based. I will examine the locus of the prominent syllable and the syntactic category of the constituents in which they occur. I will then present the results of the measurements of the correlate of final velocity, mean F0, pitch excursion, and finally, duration.

4.2.2.1 Hypothesis

Based on the observations of other languages reported in §4.1.2.1, and on a first informal assessment of the Jaminjung data, it is hypothesised that focus is realized by making the first syllable of the comment prominent by assigning it a specific target, and by pronouncing it with a wider pitch excursion and a longer duration.

4.2.2.2 Dataset

In Jaminjung, as in many languages, it is not unusual to find sequences of clauses in which a continuing topic is coreferential to one of the arguments of the clause, nearly always the subject/agent. Further, given topics are often elided so that clauses lack an overt subject, or rather, the subject in such clauses is expressed only by the inflection on the verb, here bound pronominals.

The tokens in this dataset have been selected because they all include an overt given topic – referring to an entity in the realm of the shared knowledge of both hearer and speaker – and a comment that contains new information but no contrastive element.

In Jaminjung, clear cases of IUs with topic/comment structure with a predicate focus include (see also §3.2.3.4):

- Verbless clauses: [NP]_{top} + [NP]_{PredF}
- Intransitive clauses: [NP]_{top} + [(coverb) V (OBL)]_{PredF}
- Transitive clauses: [NP]_{top} + [NP (coverb) V]_{PredF} or
[NP]_{top} + [(coverb) V NP]_{PredF}

Verbless clauses have a predicative NP which in Jaminjung does not require a copula. They occur very frequently in the dataset, as they were collected in the course of documenting the biological knowledge of the speakers. Speakers were shown pictures of various plants and animals and asked to either identify them or comment on their use etc. In Figure 4-1, the demonstrative *nginju* ‘this one’ introduces a nominal predicate *yag* ‘fish’.

(54)
 nginju yag
 PROX fish

Here (this one) (is) a fish.
 [JM:CS07_068_03]

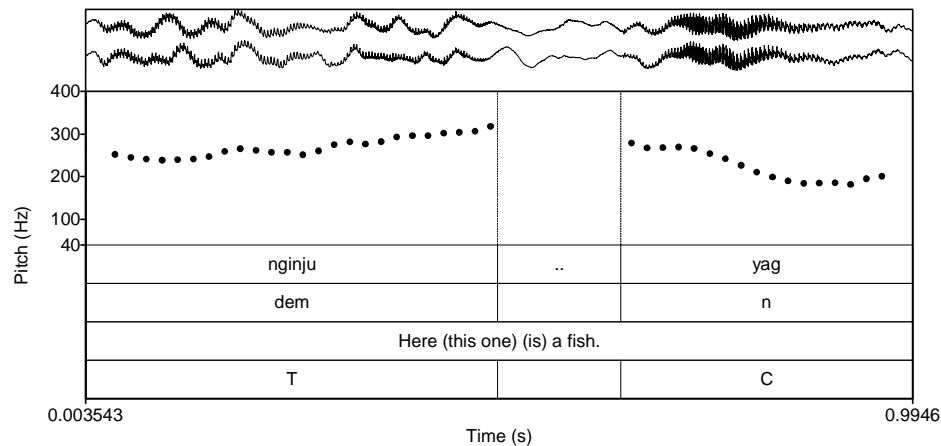


Figure 4-1 Example of IU consisting of a topic NP and a comment NP.

A peculiarity of this dataset is that many nominal predicates are made up of a single noun, so that their final syllables are also comment final and thus bear the encoding associated with boundary marking, a condition that is fully compatible with the idea that syllables receive parallel and multiple encodings in the PENTA model, but which may make identifying the specific parameters associated with each function less easy. Verbal predicates are usually longer, so that the words in initial and final position are less likely to bear the encodings of multiple functions. This will need to be taken into account in the interpretation of the results.

Semantically, intransitive clauses express different meanings, such as an action, a state change or a state, essentially following Chafe's description (1970:98). In example (55) Figure 4-2, the topic *janyung*, 'the other', is followed by the comment consisting of a complex verb (a coverb and an inflected verb), *buru gajgany* 'has gone back'.

(55) (also in (59) and (67))
 janyung buru ga-jga-ny
 other return 3sg-go.PST

'The other has gone back.'
 [IP:ES96_08_03]

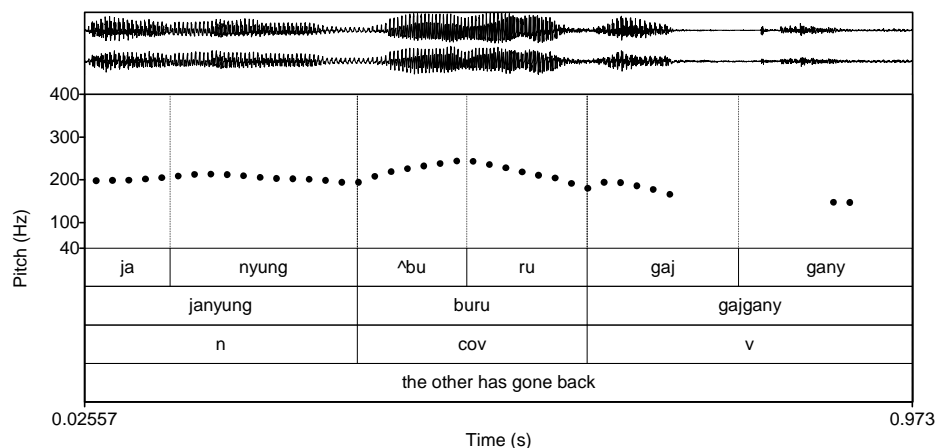


Figure 4-2 Example of IU consisting of an intransitive clause.

In transitive clauses, the action or activity is transferred from the subject/agent to the object (patient). The object is said to be at least affected by the activity. Example (56) is extracted from a retelling of the Frog Story, the subject/topic *wirib-ni* 'the dog' is followed by a comment with a complex verb *jag ganadgiyany* 'has thrown down', and an NP argument, the NP *thanthiya munurr* 'those bees'.

(56)
 wirib-ni jag gan-adgiyany thanthiya munurr
 dog-ERG_INSTR go.down 3sg:3sg-throw-PST DEM paperwasp

'The dog has thrown down those bees.'
 [DP:ES96_07_01]

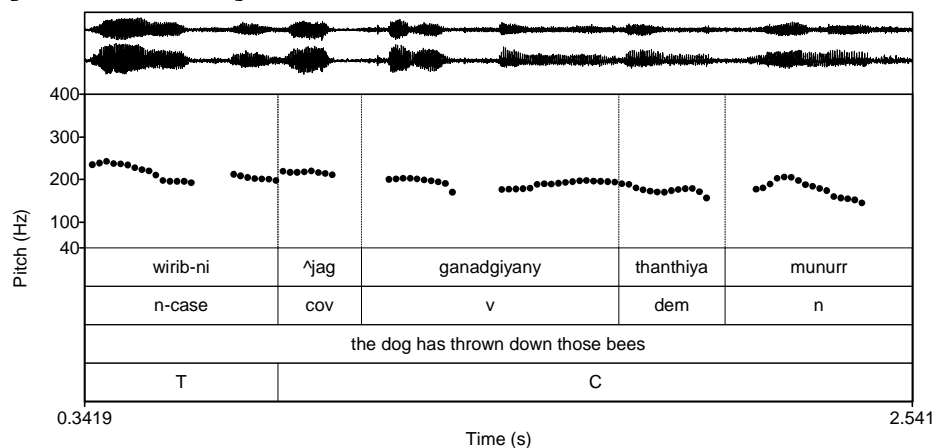


Figure 4-3 An IU consisting of a transitive clause: NP(top) + coverb V NP(comment).

The set of data for the comment focus domain comprises 73 IUs. There are 43 IUs of types 2 and 3 (intransitive and transitive clauses), where the subject/agent functions as the topic, and the predicate as the comment about the subject-topic; and 30 IUs of type 1 (nominal predicates) where the first NP is a topic and the second NP a comment on this topic. The distribution of verbal to non-verbal clauses in the dataset (43 verbal clauses to 30 non-verbal clauses) does not reflect the real frequency of these structures in natural language. It occurs in our data for practical reasons related to fieldwork contingencies: as mentioned earlier, the non-verbal clauses occurred frequently in documenting plant use in the area, and in labeling pictures for dictionary use.

Example (56) (above) can be used as an illustration of the labeling of the topic-comment structure of the IUs in this dataset. The topics are marked with (T) and the comments/predicates with (C). Arguments of transitive verbs or other oblique arguments are coded as part of the comment when they are within the same IU. Arguments in a separate IU were given a separate code (C+). Other parts of a sentence that were not part of the topic-comment structure, such as framesetting topics (fairly common in speech data elicited in picture description tasks) were given a different code.

4.2.2.3 First syllable prominence

The scope of the focus in a topic-comment/predicate structure is the whole comment/ predicate; however, the accent usually falls on one syllable within the focus domain. The need to explain why the accent does not fall on all of the lexical items in a broader focus domain has given rise to the notions of *focus projection* or *accent percolation* (see Krifka, 2006 and Beaver et al. 2004, for an overview), the former meaning that focus is assigned to a simple constituent and can be projected to a larger constituent, the latter that focus can be assigned to a complex constituent which is marked by an accent somewhere within it. These notions represent views on how syntactic/semantic components of grammar interface with the phonological aspects of focus realization and remain highly controversial in their applications to natural language.

Table 4-2 shows that in 66 of the 73 IUs in the dataset (90.41%), the main prominence is heard on the first syllable of the comment in focus; in four tokens, the prominence is heard within the predicate but not on the first syllable. Out of the three tokens where the main prominence is heard on the first syllable of the topic, two tokens receive a secondary prominence on the first syllable of the predicate.

A secondary prominence is assigned in 13 instances, twice in topics, once in another argument and in the remaining 10 tokens, in the predicate, indicating that some predicates receive both a primary and a secondary prominence.

Category	Primary prominence		Secondary prominence
	First syllable	Other syllable	all
topic	3		2
comment	66	4	10
other argument			1
Total	73		13

Table 4-2 Rates of occurrences of first and second prominences

4.2.2.3.1 Syntactic category of the first constituent in the comment

As the perceived prominence mostly occurs in the first syllable of the first constituent of the comment, it is interesting to determine what this constituent consists of, particularly in a language such as Jaminjung, in which word order is not constrained by syntax. The first constituent of the comment may be a nominal, a coverb, an inflected verb, or a negation particle. The counts for each are shown in Table 4-3.

In verbal clauses, it corresponds to the first syllable of a coverb (marked as ‘cv’ in Table 4.3) in 31 out of a total of 43 tokens (72.09%), to an inflected verb in 4 tokens, and to an argument of the verb, placed before it, in 3 tokens. In nominal clauses, the main prominence is heard on the first syllable of the predicate NP in 27 of the 30 tokens in the dataset (90%).

As to the tokens where the prominence is heard elsewhere, in 5 tokens the main prominence is heard on a syllable that is not in initial position in the comment/predicate. In one of them, the prominence is heard

on the first syllable of a coverb which is preceded by the negative particle *gurrany* (which do not seem to attract prominences); one on the first syllable of the inflected verb in a complex verb construction, similarly not in initial position. In nominal clauses, the prominence is heard in 3 tokens on syllables that are not in initial position in the predicate. Finally, there are 3 tokens where the prominence is heard on the first syllable of the NP topic. This generalisation about the locus of the prominence on the first syllable of the comment, according to their syntactic category, is statistically significant, chi-square (9, N= 85) = 41.394, p=0.000.

syntactic constituent clause type	Primary prominence	
	verbal	nominal
First syllable of Comment		
arg	3	
cv	31	
v	4	
nominal (pred)		27
Total	38	27
Other syllable within Comment		
arg		
cv	1	
v	1	
nominal (pred)		3
arg (2nd in pred)		
Total	2	3
Elsewhere (not in comment)		
top-np	3	
other		
Total	3	
TOTAL	73	

Table 4-3 The occurrences and positions of primary and secondary prominences and their correspondence with syntactic constituents, shown in verbal and nominal predicates.

Figure 4-4 is an illustration of a verbal predicate, in which the main prominence is most often heard on the first syllable of the first word, usually a coverb. The topic *thanthubiya ngayin* ‘those animals’ has a fairly flat contour and the verbal predicate *bunburr burrangga* ‘all take off’ is marked by a pitch excursion of the first syllable of the coverb *bunburr*.

(57) (also in (63))

thanthu=biyang	ngayin	bunburr	burr-angga
DEM=now	meat/animal	many.take.off	3pl-go.PRS

'Those animals all take off.'

[IP:ES97_03_01]

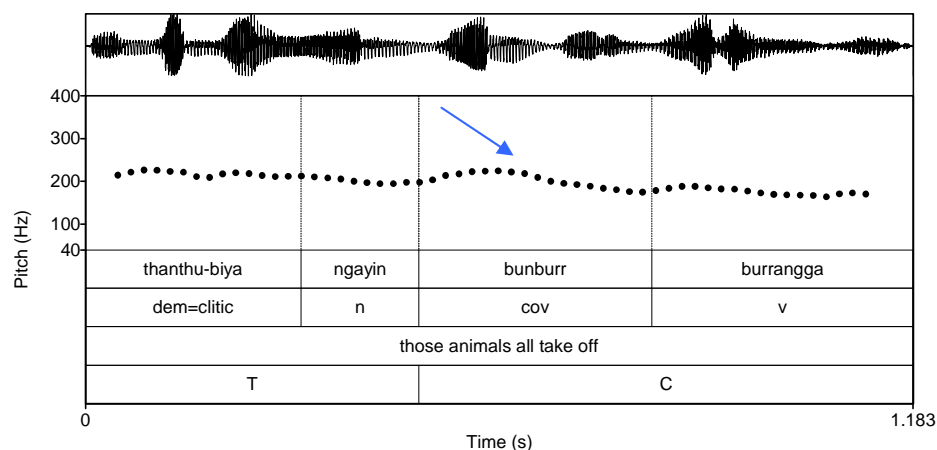


Figure 4-4 Illustration of topic-predicate construction with a prominence on the first syllable of the predicate. This syllable corresponds to a coverb (the arrow points to it).

In summary, in most of the tokens in the dataset, the main prominence is heard on the first syllable of the first prosodic word of the comment. In verbal clauses, the prominence most often coincides with a coverb, or with an inflected verb, even when the clause contains an NP. The perception of the prominence on these syllables will now be corroborated by measurements of their acoustic parameters - duration, pitch, and velocity - to verify the prosodic cues used to mark them as such. In the following sections, only IUs with a prominence on the first syllable of the comment are retained in order to control for any effect on the measurements deriving from the position of the syllables within the word and the IU.

4.2.2.3.2 Prosodic correlates of the prominent syllables: specific pitch target

In this section, I will report on the measurements made to test the hypothesis that the prominent syllable is marked with a specific pitch target. The patterns under investigation are illustrated in examples (58-61) illustrated in Figure 4-5 below. In the first three, the comments are verbal clauses with coverbs of various lengths (1, 2 and 3 syllables) in initial position, *wib* 'look back', *buru* 'return' *yarrajgu* 'afraid'. In the last

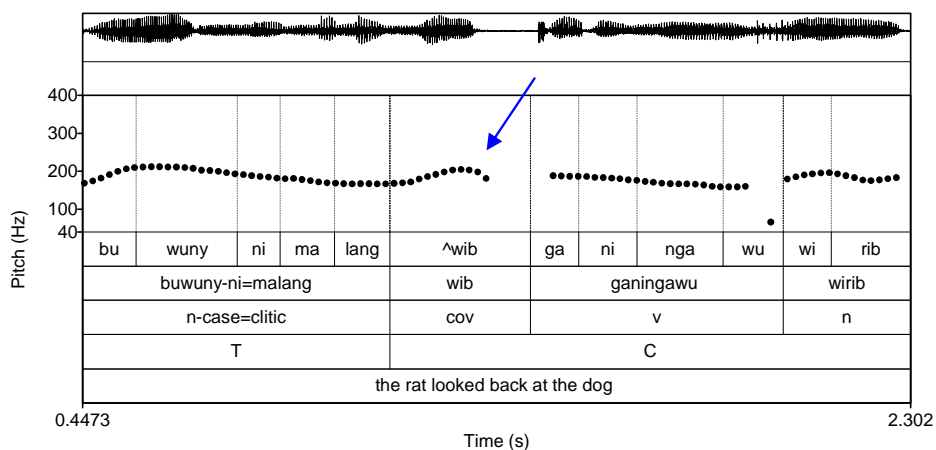
example, the nominal predicate comprises two words, the prominence falling on the first syllable of the initial word *jarlag* ‘good’. An arrow points at the syllables in question.

(58)

buwuny-ni=marlang wib gani-ngawu wirib
marsupial.rat-ERG/INSTR=GIVEN look.behind 3sg:3sg-see.PST dog

'The rat looked back at the dog.'

[CP:ES96_18_02]

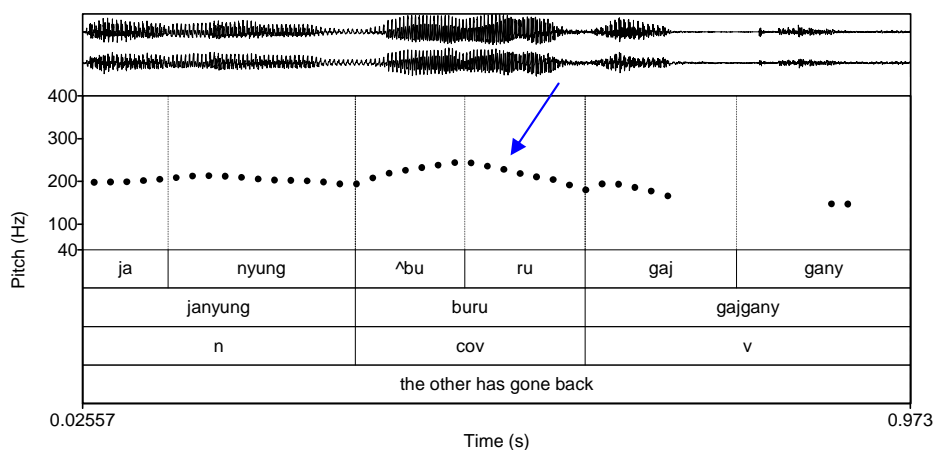


(59) (also in (55) and (67))

janyung buru ga-jga-ny
other return 3sg-go.PST

'The other has gone back.'

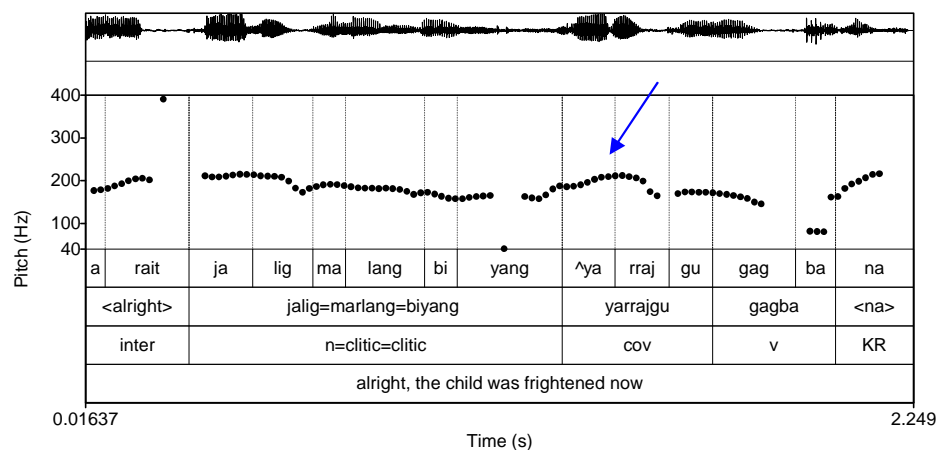
[IP:ES96_08_03]



(60)

<alright>	jarlig=marlang=biyang	yarraigu	ga-gba =<na>
all.right	child=GIVEN=now	afraid	3sg-be.PST=now

'All right, the child was frightened now.'
[CP:ES96_18_02]



(61)

nginju	jarlag	wirib
PROX	good	dog

'This one is a good dog.'
61 [BH:CS07_72_01]

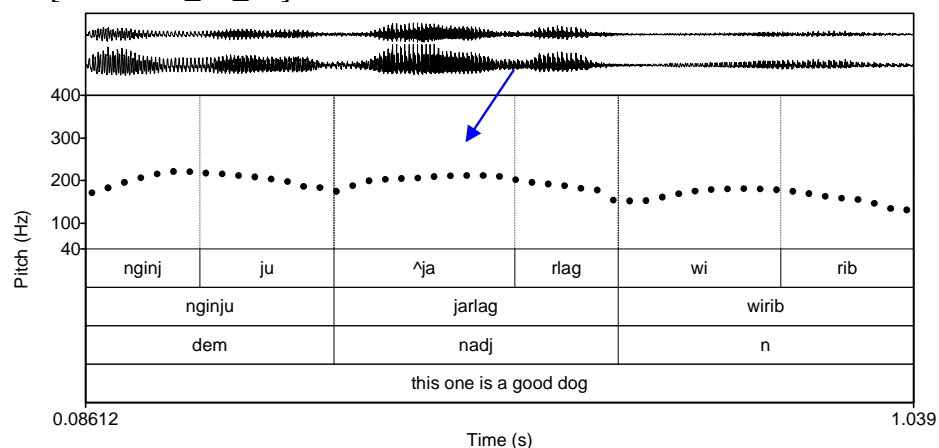


Figure 4-5 Examples with words of 1-, 2- and 3- syllables in comment-initial position which have a prominent first syllable, shown by an arrow.

4.2.2.3.3 Final velocity in the syllables of the first word of the comment

The first correlate examined is final velocity which is considered the best indication of the underlying pitch targets. A first overall statistical test shows that the differences in the values according to syllable position in the

IU are statistically significant ($F(15, 247) = 1.911, p = .023$). I will present a detailed analysis of the measurements of the initial word in the comment (which contains the main prominence), taking into account its syntactic category and number of syllables.

Initial words may be coverbs, inflected verbs, or nouns. The number of syllables and their relative position within the words are also considered. The calculations only include categories that have a sufficient number of tokens for comparison: monosyllabic coverbs, bi- and tri-syllabic coverbs, inflected verbs, and nouns. The results are presented in the summary Table 4-4 and graph Figure 4-11 at the end of the section.

A series of ANOVA tests show that both syntactic category and number of syllables have a significant effect on the final velocity ($F(19, 154) = 1.899, p = .018$), and ($F(29, 145) = 1.586, p = .041$).

Monosyllabic coverbs in verbal predicates are quite common (12 tokens), their final velocity, -23.62st/s , suggests a [fall] target. In this dataset, monosyllabic coverbs are always followed by an inflected verb with an average velocity of 0.04st/s in its first syllables, reflecting a [mid] target. The suggested overall pattern in monosyllabic coverbs is:

Monosyllabic coverb	Inflected verb
	1st syll
[fall]	[mid]

Example (62) in Figure 4-6 illustrates these findings. The contour in the monosyllabic coverb *bul* ‘emerge’ reaches a peak early in the syllable and then falls sharply (its velocity is -16.06st/s). The first and second syllables of the inflected verb *ganima* ‘hit’ display a much less steep contour (-6.88st/s and 4st/s , respectively).

(62)

jabarlng=marlang bul gani-ma burrinyi
 frog=GIVEN emerge 3sg:3sg-hit.PST 3DU

'The frog however went out on them.'

[CP:ES96_18_02]

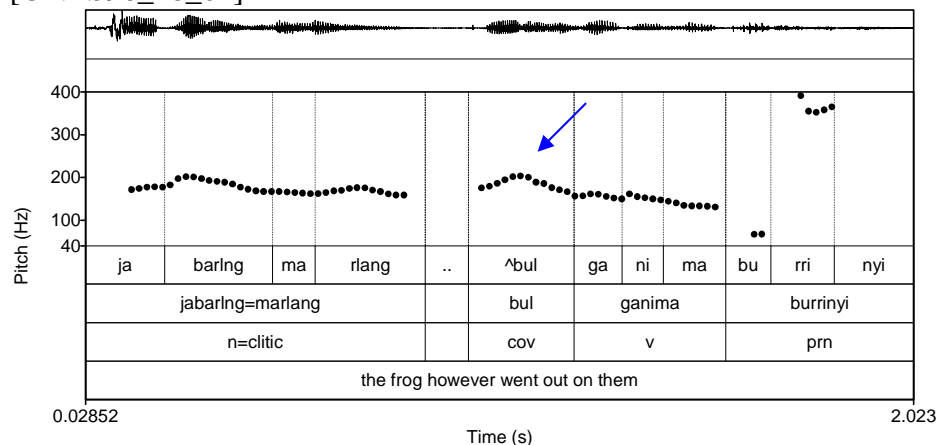


Figure 4-6 The sentence contains a monosyllabic coverb, with a [fall] pitch target, followed by [mid] targets on the inflected verb.

Bisyllabic comment-initial words make up the largest subset in this dataset. They correspond to coverbs in 14 tokens, to inflected verbs in 4 tokens, and to nouns in 14 tokens. The average final velocity values of the first and the second syllables in bisyllabic coverbs are -12.91st/s and -5.63st/s; in nouns -0.47st/s and 4.67st/s; and in verbs -2.98st/s and -10.36st/s. The value of the final velocity for the first syllable of bisyllabic coverbs is nearer zero than in its monosyllabic counterpart (-23.62st/s), but it is nevertheless markedly lower than in inflected verbs or nouns. The final velocity in the first syllable of the nouns and the inflected verbs indicates a fall, although a very slight one. For the nouns in this subset, the second syllable is also comment-final, which may be reflected in its velocity. The values for the verbs, with a lower average negative value in the second syllable, could suggest that the [fall] target is realized in this syllable rather than the first. However, there is great variability in the values, probably due to the very small number of tokens in the analysis, which indicates that the results, at this stage, can only be interpreted as a possibility which would need further investigation. The targets suggested are:

Bisyllabic coverbs, nouns and (maybe) verbs

1st syll	2nd syll
[fall]	[mid]

The examples below illustrate these findings. In example (63), a bisyllabic coverb, *bunburr* ‘take off (of many animates)’, has a prominence on its first syllable (-34.49st/s), and receives a [fall] target. Its second syllable (-8.15st/s) continues with this falling movement, and the first syllable of the following word, the inflected verb *burranga* ‘go’, has the gently falling contour typical of a succession of [mid] targets (-8t/s).

(63) (also in (57))

thanthu=biyang	ngayin	bunburr	burr-angga
DEM=now	meat/animal	many.take.off	3pl-go.PRS

'Those animals all take off.'

[IP:ES97_03_01]

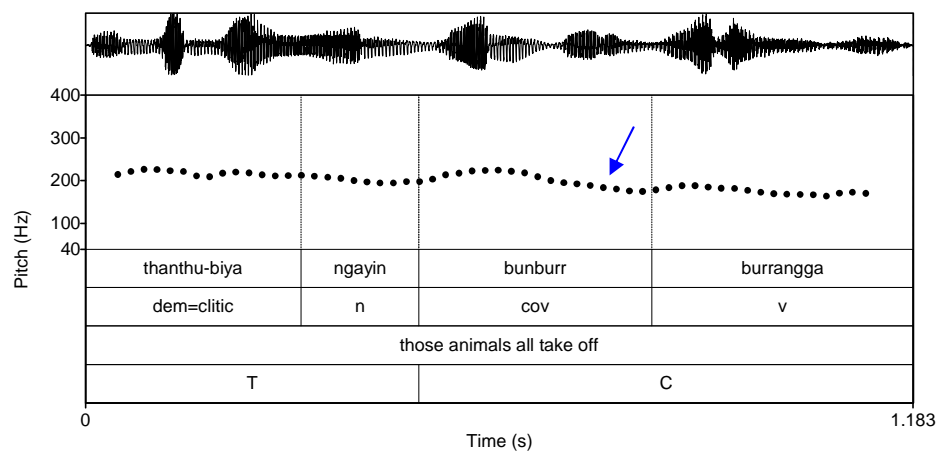


Figure 4-7 A bisyllabic coverb *bunburr* with a pattern of a [fall] and [mid] targets on the first and second syllables.

Example (64) (Figure 4-8) is of a bisyllabic nominal predicate *juwud* ‘eye’ displaying a less sharp falling pitch track, more typical of the [high] target (1st syll.= 5.18st/s, 2nd syll.= -14.36st/s).

(64)
 nginju juwud
 PROX eye

'Here is the eye.'
 [JM_CS06_14_01]

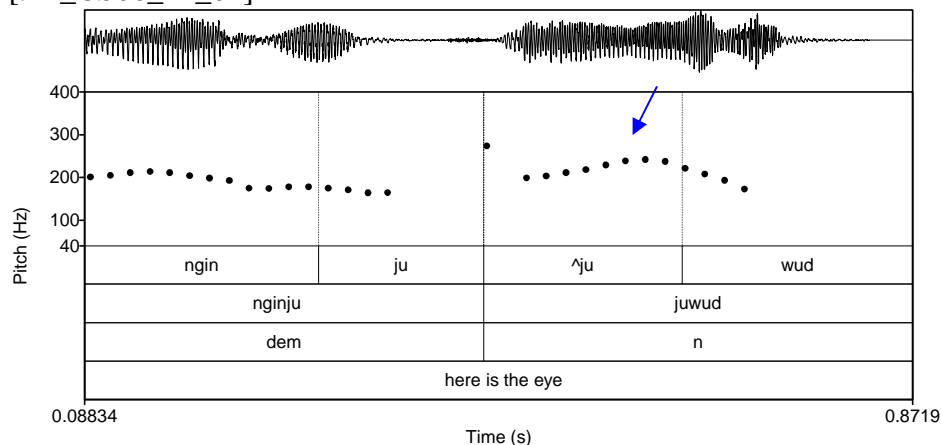


Figure 4-8 A bisyllabic nominal *juwud* with a [high] on the first syllable, followed by a [fall].

In longer trisyllabic comment-initial words, the average final velocity value of prominent syllables in coverbs is -1.89st/s, and -7.15st/s in nouns⁴¹. The average value for the second syllables in coverbs is -16.47st/s, and in nouns -14.75st/s; that of the third syllable is -12st/s and 1.42st/s, respectively. This suggests the following pattern of targets:

Trisyllabic coverbs and nouns

1st syll	2nd syll	3rd syll
[high]	[fall]	[mid]

Example (65) below shows a trisyllabic coverb *yarrajju* 'frightened' which displays this pattern. The first syllable rises slightly (11.53st/s), the second syllable is the locus of the [fall] (-50st/s), which expands into the third syllable (-36.97st/s). The first syllable of the following word, the inflected verb *gangga* 'went', has a gently falling contour, similar to those observed in the verbs following mono- and bi-syllabic coverbs displaying a [mid] target.

⁴¹ There are no instances of comments consisting of trisyllabic verbs in the dataset, although they do occur in natural speech.

(65)

wirib=marlang=biyang
dog=GIVEN=now

yarraju
afraid

g-angga
3sg-go.PST

'The dog is frightened.'

[CP:ES96_18_02]

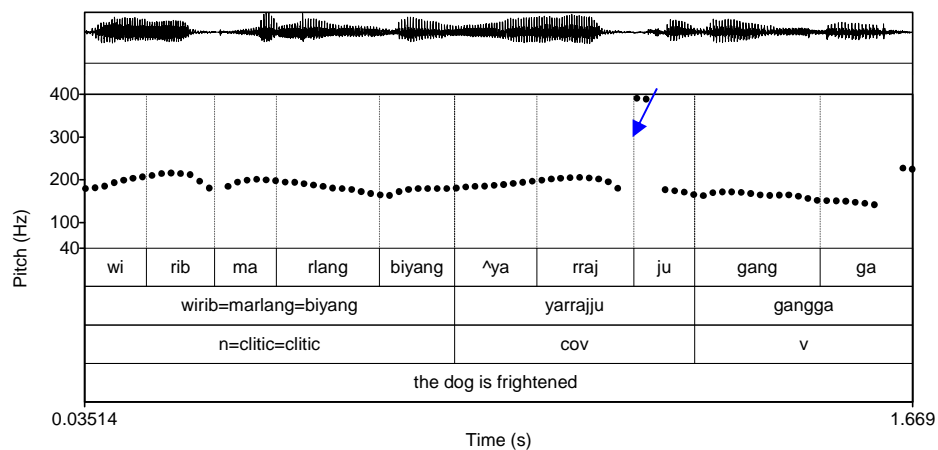


Figure 4-9 A trisyllabic coverb *yarraju* displaying the pattern of pitch target [high] [fall] [mid].

The trisyllabic nominal predicate in example (66), *mulurru*

'woman', has a similar pattern (1st syll. = 5.81st/s, 2nd syll. = -29.41st/s, 3rd syll. = 1.34st/s).

(66)

nginju
PROX

mulurru
old.woman

'Here is an old woman.'

[JM:CS06_34_03]

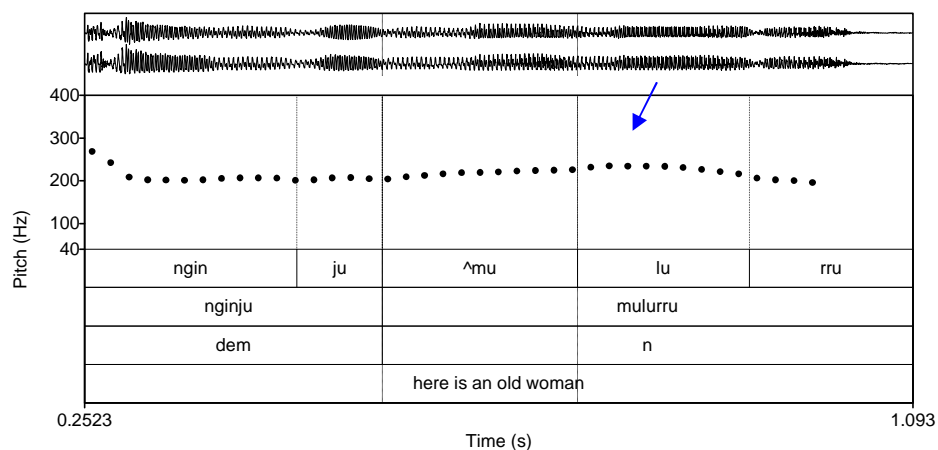


Figure 4-10 A trisyllabic noun, *mulurru*, also displaying a pattern of [high] [fall] [mid].

The overall results of the final velocity calculations for monosyllabic coverbs, bi- and tri-syllabic coverbs, inflected verbs, and nouns are presented in Table 4-4 and graph Figure 4-11 below.

Syntactic category/ Syllable position			Final velocity (st/s)		
			PROM_1st syll	2nd syll	3rd syll
cv	cv-1syll	Final vel	-23.62		
		St.dev	25.20		
		n	12		
	cv-2syll	Final vel	-12.91	-5.63	
		St.dev	16.52	22.23	
		n	14	14	
	cv-3syll	Final vel	-1.89	-16.47	-12.00
		St.dev	42.77	28.03	18.02
		n	5	4	4
v	v-2syll	Final vel	-2.98	10.36	
		St.dev	24.46	40.32	
		n	4	3	
n	n-1syll	Final vel	-35.72		
		St.dev			
		n	1		
	n-2syll	Final vel	-0.47	4.67	
		St.dev	17.35	16.64	
		n	14	14	
	n-3syll	Final vel	-7.15	-14.75	1.42
		St.dev	7.34	27.28	10.86
		n	9	7	6

Table 4-4 Final velocities by syntactic categories and by number of syllables.

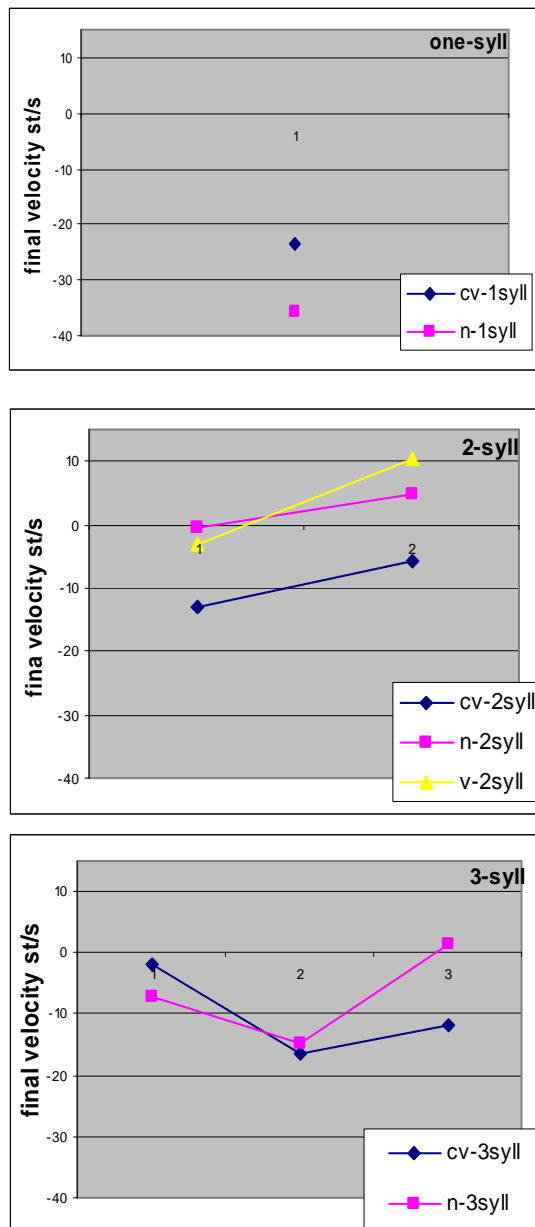


Figure 4-11 Final velocities by word length.

Summarising, the values obtained for coverbs, inflected verbs and nouns suggest that coverbs differ in their patterning from the other two categories. The final velocities of their first and prominent syllables have lower values, most particularly in monosyllabic coverbs.

For the three categories, it is suggested that the overall pattern is of a [fall] target on the prominent syllables, followed by a [mid] target on the adjacent syllable. This pattern is somewhat spread in longer words, at least in the trisyllabic words in this dataset, with the [fall] target occurring on the

second syllable. The prominence in this case is nonetheless perceived on the first syllable. The implications of these measures will be discussed further in §4.2.3 below.

4.2.2.3.4 Comparing the first syllable of the topic and the first prominent syllable in the comment

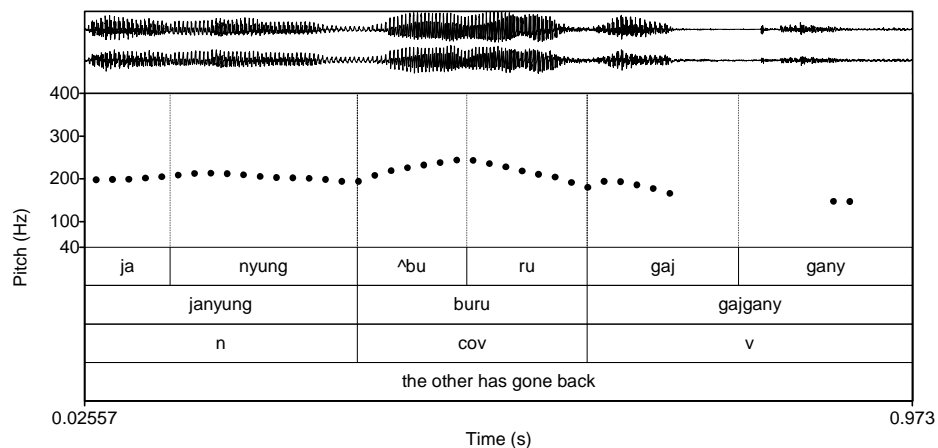
At the level of the IU, it is interesting to compare the first syllable of the comment and that of the topic, to assess whether the patterns reported above for comments really have a focus encoding function and are not simply marking the first syllable of a major grouping unit. In the latter case, one would expect the first syllable of the topic, which coincides with the first syllable in an IU and a prosodic phrase, to have the same encoding. The average final velocity of the prominent syllable in initial position in the comment is -9.65st/s, and that of the first syllable of the topic is 0.34st/s, suggesting different underlying targets, a [fall] target for the comment and a [high] target for the topic (see §4.5 for an analysis of the topics). The differences are statistically significant ($F(1, 133) = 7.017, p = .009$). Figure 4-12 shows an example where the difference in the first syllable of the topic, *janyung* ‘the other one’ and that of the comment, *buru* ‘return’ is easily observable in the pitch track. The velocity curves for the two syllables are also shown.

(67) (also in (55) and (59))

janyung	buru	ga-jga-ny
other	return	3sg-go.PST

'The other has gone back.'

[IP:ES96_08_03] (repeated ex.46)



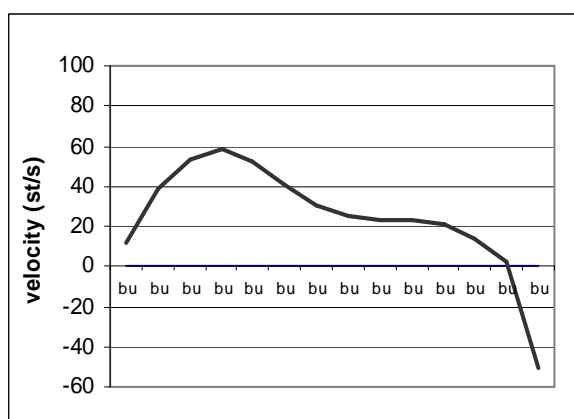
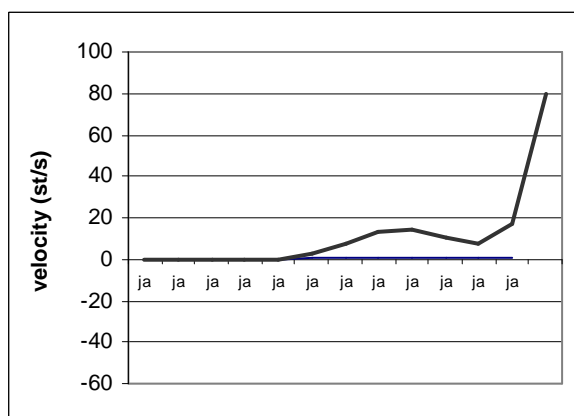


Figure 4-12 The topic in the example is the noun *janyung*, and the comment is *buru gajgany*. The main prominence is heard on the first syllable of the comment, here the coverb *buru*. The velocity measurements for the initial syllables of *ja-* and *bu-* are also shown.

The starting hypothesis of a specific tone associated with the encoding of focus for the prominent syllable in the predicate is borne out by the measurements. The pitch target most often associated with this syllable is found to be a [fall], results which are consistent with those obtained by Liu and Xu for English and Mandarin (2005, 2007), and El-Zarka for Egyptian Arabic (to appear), although the slope of the fall is not very steep compared to the values found in Liu and Xu (2007). The final velocity values are lower in coverbs than in the other word categories, particularly in monosyllabic coverbs, pointing to an important difference between word categories in the timing of the fall. The difference could be attributed to syllable structure or to whether the word in question is preceded by a pause (however minimal). This question would warrant further investigation. F0 measures are effected to confirm these findings.

4.2.2.3.5 Mean F0 in the first word of the comment

To continue characterizing the prominence perceived on the first syllables and its special status within the IU, the correlate of mean F0 is considered.

The mean pitch is measured within the word containing the prominence. Once again, the results are shown by number of syllables and by syntactic constituents, thus distinguishing between coverbs, inflected verbs and nouns of 1-2 and 3 syllables. The values are shown for non-IU final words. The statistical tests show the results not to be significant, they are presented here as they help to develop an overall picture of the patterns, and to confirm the pitch targets associated with the prominent syllables.

Monosyllabic coverbs have a high mean pitch, of 203.73Hz, compared to 185.03Hz in the first syllables of the following inflected verbs. In bisyllabic words, nouns have the highest mean F0 overall, followed by coverbs and inflected verbs, a pattern that is maintained in trisyllabic words, as shown in Figure 4-13 below. The first syllables have higher mean pitch, except in the case of the tri-syllabic coverbs, which have higher averages in the second syllables.

These values can be interpreted in light of the suggested pitch targets described in the previous section. In monosyllabic coverbs, the interpretation of a [fall] pitch target is supported, as it should translate in surface F0 as a pitch peak before a rather sharp fall. The same can be said of the pitch patterns observed in bisyllabic words, where the mean pitch is higher in the first syllable. In tri-syllabic words, the higher values on the second syllables of coverbs suggests the pitch peak associated with a [fall] target does indeed occur during the second syllable, further reinforcing the interpretation of a displacement of the [fall] target away from the first syllable in longer words.

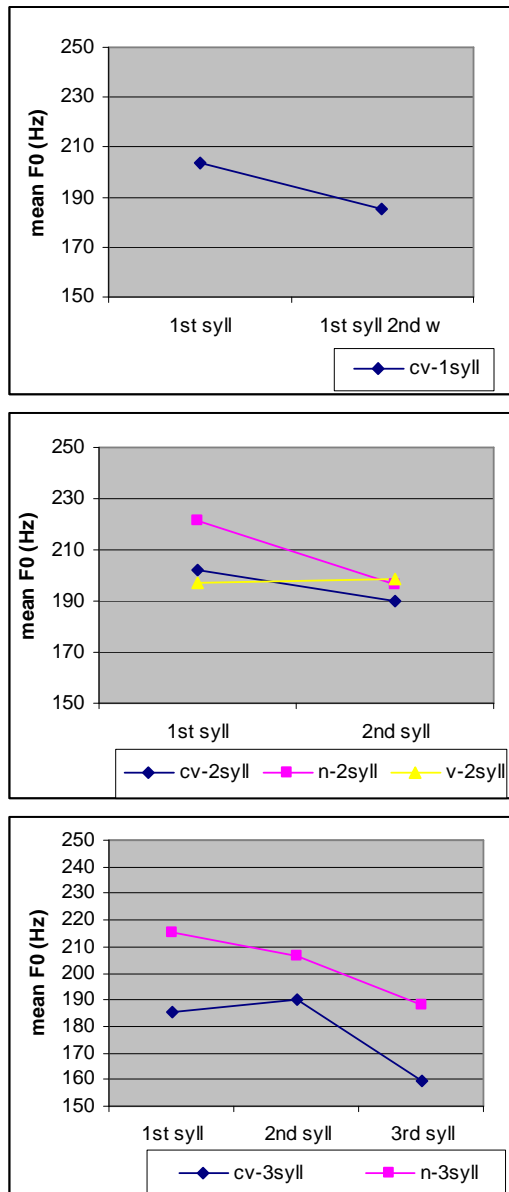


Figure 4-13 Mean F0 values in the monosyllabic coverbs, also showing the first syllables of the following word (top graph), in bisyllabic coverbs, nouns and inflected verbs (middle graph), and in tri-syllabic coverbs and nouns (bottom graph).

4.2.2.3.6 Comparing the mean F0 of the prominent syllable to others in the IU

The F0 value of the prominent syllable is compared to that of the other syllables in the IU. The differences are found to be statistically significant ($F(22, 591) = 10.350, p = .000$). The mean values for each syllable are shown in the boxplot graph in Figure 4-14 below.

The expected pattern of declining mean F0 values throughout the IU with a slight reset at major phrasal boundaries (§3.3.4.1) is found to occur.

Although prominent syllables (marked as MOST-PROM in the graph) do not have the highest mean F0 in the IU with an average value of 208.29Hz, they are very close to IU-initial syllables, corresponding to the first syllable of the topic (1st prePROM), with 211.49Hz. The high value of the prominent syllable is made more salient when compared with that of the first syllable in the second word of the comment, which has a much lower average of 174.85Hz.

Mean pitch is not considered a specific encoding parameter of focus, but rather an effect of the syllable receiving a [fall] pitch target.

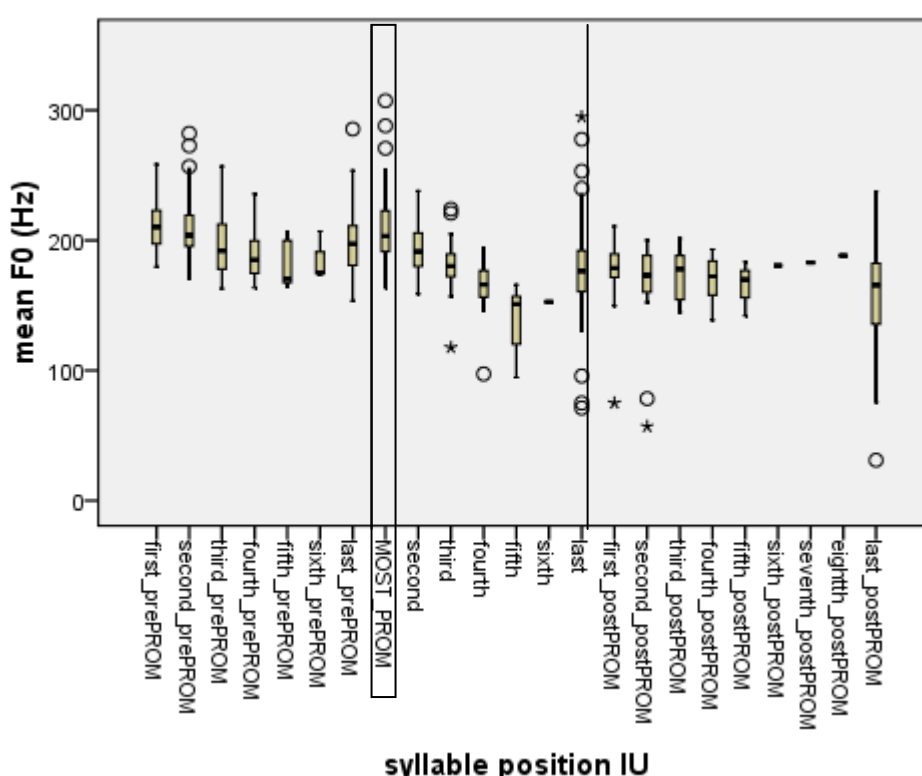


Figure 4-14 Mean F0 values for the syllables before and after the prominent syllables (MOST_PROM, boxed). Constituents are separated by a vertical line.

4.2.2.4 Prosodic correlates of the prominent syllables: pitch range extension

In this section, evidence that the prominent syllables in the comment in focus have a wider pitch excursion than the other syllables is presented. Excursion size is considered as a parameter independent from pitch targets as the same pitch target may be uttered with a wider or narrower excursion. An overall calculation, including all the tokens in the dataset, shows that the

differences in excursion size according to syllable position are statistically significant ($F(22, 576) = .709, p=.000$). The average values for each syllable position are given in the boxplot graph in Figure 4-15.

The average size of the excursion in the most prominent syllables is 3.79st; higher than that of the first syllable in the IU, which has an average of 3.28st. The excursion size of the prominent syllable is even wider than that of the topics (last pre_RPOM) which has a value of 2.4st. This wider excursion size in the topic-final syllable will be discussed further in the examination of the encoding of topics (§4.5). As to the syllables that follow the prominence, those immediately following the prominent syllable have an average of 2.3st, and so on in a decreasing fashion over the words of the IU. The last syllable of the word with the main prominence has a fairly wide pitch excursion, 3st, but still not quite as large as that of the prominent syllable. The only syllable that has a wider excursion is the IU-final, with a value of 4.3st. The values for the IU-final syllables are high, indicating not a peak but the final lowering associated with the major boundary marking.

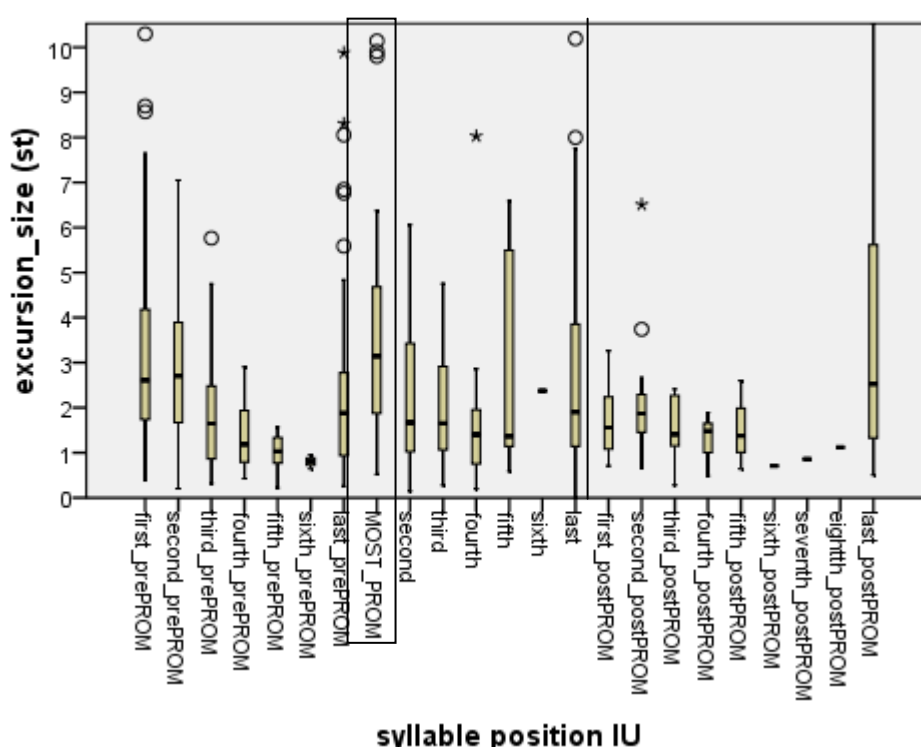


Figure 4-15 Excursion size in the syllables in the words that precede the main prominence (left); and in the syllables that succeed the prominences, in the same word (labeled 2nd, 3rd, 4th, etc) and in the 2nd word, when they occur (1st_postPROM, etc). The most prominent syllable is boxed.

Support for the hypothesis that the pitch range of the prominent syllable is extended is found in a fairly straightforward fashion in the pitch measurement reported here. These measures show that, notwithstanding the underlying pitch target associated with the syllable, the actual pitch excursion is wider in the prominent syllable.

4.2.2.5 Prosodic correlates of the prominent syllables: duration

After pitch, duration is the parameter most often associated with prominence, and thus with the marking of focus. Overall measurements are shown first, and then a more detailed examination of the values by syntactic category and number of syllables. The measurements are shown in Table 4-5.

Firstly, the average duration of the first (most prominent) syllables of the comment, 213.674ms, is compared with that of the other syllables in the IU. The first syllable of the topic (first_prePROM) has an average of 192.33ms; that for comment-final words (1stpost_PROM) is 158.64ms. Syllables in word final positions which are subjected to final lengthening (§3.3.3.1) also have shorter durations than the prominent syllables, with values of 171.310ms for the last syllable of the topic (last_prePROM), 185.47ms for the syllables at the end of the word containing the prominence; and 199.67ms for the IU-final syllables. These differences are statistically significant ($F(22, 593) = 5.608, p = .000$). However, the interpretation of these results must take into account the high rate of occurrences of monosyllabic coverbs in comment-initial position, which are more likely to have coda consonant clusters (§3.2.2.2.3). A further test will be conducted to assess the variations in duration associated with the number of syllables.

Syllable position	Mean duration (ms)	Std. Dev	N
first_prePROM	192.33	72.09	75
second_prePROM	167.88	63.14	40
third_prePROM	113.00	38.77	29
fourth_prePROM	126.04	37.18	19
fifth_prePROM	94.57	29.68	7
sixth_prePROM	154.88	42.01	3
last_prePROM	171.31	74.12	70
MOST_PROM	213.67	69.34	72
second	166.29	51.54	53
third	147.83	51.03	38
fourth	130.25	51.71	24
fifth	149.77	52.75	7
sixth	226.00	.	1
last	185.47	63.84	72
first_postPROM	158.64	51.52	22
second_postPROM	148.51	57.43	16
third_postPROM	131.98	42.35	8
fourth_postPROM	139.77	42.23	7
fifth_postPROM	193.55	69.67	3
sixth_postPROM	132.17	.	1
seventh_postPROM	184.62	.	1
eightth_postPROM	123.78	.	1
last_postPROM	199.67	80.26	24

Table 4-5 Mean duration of the syllables in the IUs.

4.2.2.5.1 Syntactic constituents and number of syllables

This test aims at checking how grammatical constituency and number of syllables in the constituent affect the duration of the prominent syllables. Obviously, syllable duration can be correlated to syllable structure which does not display extensive variation in Jaminjung, making syllable durations comparable, except for the monosyllabic coverbs which can have consonant clusters in coda position (§3.2.2.2.3). This test will compare only words that are non-comment final, to avoid interference with boundary marking.

An ANOVA test with duration of the prominent syllable as dependent variable and constituent type/number of syllables as independent variable shows a significant relation between them ($F(11, 119) = 2.740$, $p = .004$). Table 4-6 Duration of the 1st prominent syllables shown according to the number of syllables in the word, and its syntactic category.

Table 4-6 shows the durations of the prominent syllables, broken down by grammatical category and number of syllables; the same values are illustrated with graphs in Figure 4.16.

Firstly, I will compare the duration of the prominent syllables in words of various length. Monosyllabic coverbs have an average duration of 266.50ms, whereas the prominent first syllables in bisyllabic coverbs have durations of 190.07ms, and 171.83ms in trisyllabic coverbs; in bisyllabic nominals, it is 240.98ms, and 193.36ms in trisyllabic nominals. The prominent syllables are shorter when they occur in longer words; the number of syllables in the words seems to have an inverse effect on the duration of the prominent syllable, an effect that is arguably less marked in nouns than in coverbs.

The duration of the prominent syllable can also be compared relative to that of the other syllables in the word, by syntactic category. Apart from the much longer durations of monosyllabic coverbs already mentioned, this comparison confirms that prominent syllables are longer than second syllables in coverbs, verbs and nouns, even if only slightly so, particularly in bisyllabic coverbs.

Syntactic category/ Syllable position			duration (ms)		
			PROM_1st syll	2nd syll	3rd syll
cv	cv-1syll	duration	266.50		
		St.dev	52.53		
		n	12		
	cv-2syll	duration	190.07	191.06	
		St.dev	40.24	40.94	
		n	14	14	
	cv-3syll	duration	171.83	161.36	132.22
		St.dev	41.12	42.38	
		n	5	5	1
v	v-2syll	duration	194.40	121.61	
		St.dev	27.30		
		n	4	1	
n	n-2syll	duration	240.98	193.35	
		St.dev	77.92		
		n	15	1	
	n-3syll	duration	193.36	189.92	
		St.dev	52.31		
		n	9	1	

Table 4-6 Duration of the 1st prominent syllables shown according to the number of syllables in the word, and its syntactic category.

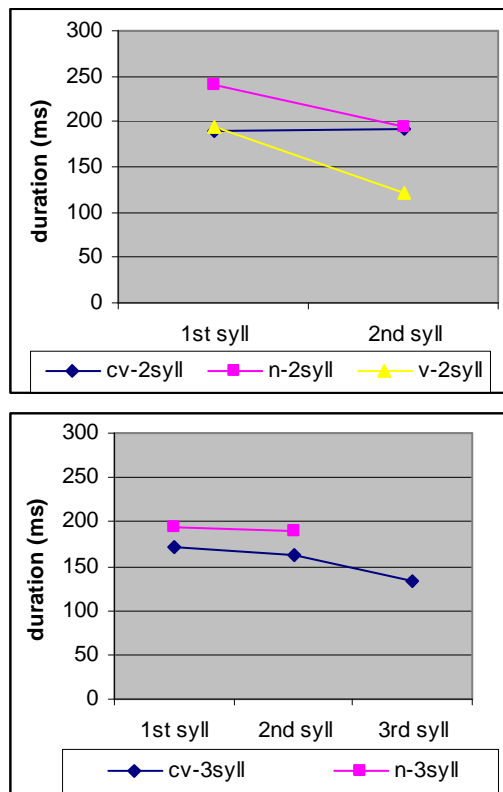


Figure 4-16 Duration of the syllables in the first word of the comment, distinguishing syntactic categories and number of syllables. The prominent syllable is the first. Only non-IU final words are considered.

The present data clearly demonstrate the importance of the durational cue in prominent syllables. The prominent syllable is longer than both the initial syllable in the topic which occurs IU-initially in this dataset, and the final syllables of the comments, which displays lengthening as a marker of inter-constituency boundaries. However, these results must take account of the high incidence of monosyllabic coverbs in comment-initial position. The final test showed that their duration differs significantly from the first syllables of longer coverbs, and from either inflected verbs or NPs that receive the prominence. The duration of the prominent syllable appears to be inversely correlated with the number of syllables in the word in which it occurs; the longer the word, the shorter the prominent syllable is.

4.2.3 Conclusion and Discussion: comment

This section aimed at verifying the encoding parameters associated with the prominent syllables in a focus domain delimited as the comment in

IUs consisting of a given topic and a comment. This structure is considered the least marked structure in terms of information structure. The analysis is based on the assumption that focus is realized in speech through prominences: a property of speech that marks one or many syllables as salient in a chain of speech. That prominence is indeed used in Australian languages to mark focus has been noted, amongst others, by McGregor (2004:95) who notes for Gooniyandi that “the prominent syllable falls within the word that the speaker considers the most important or ‘newsworthy’, the Information Focus”.

Prominence judgments were made based on perception for each IU in the dataset. A quantitative analysis of the prosodic correlates in the prominent and non-prominent syllables was performed in order to find their characteristics.

The prominent syllable is found to coincide predominantly with the first syllable of the predicate/comment in this dataset. For Jaminjung, in which word order is more flexible than English, the prominence most often coincides with a coverb, but it can also occur on an inflected verb or a (preverbal) noun phrase. Obviously, in nominal predicates the prominence falls on the first syllable of the predicative NP, usually a bare nominal. In a few tokens, a secondary prominence is also perceived.

Bolinger (1972) introduced the notion of ‘semantic weight’, proposing that the relative informativeness of a word can be used to predict its pitch prominence: the more information is conveyed by a word, the more likely it is to be prominent. In this interpretation, prominence is assigned to the constituent bearing the highest semantic weight. In Jaminjung, this is most often the coverb⁴² (but could also be an NP). As coverbs typically occur before an inflected verb, a link can be established with the noted significance of the preverbal position in Australian languages: the preverbal position attracts the constituent with the most semantic weight and hence the prominence. This is congruent with Austin and Bresnan’s (1996) and Simpson’s (2004) analyses of information structure in Warlpiri built upon

⁴² See Schultze-Berndt (2000 ch. 6) for an overview of coverb semantics in Jaminjung.

Hale's original observation⁴³ that pre-verbal (or pre-predicator) constituents appear to convey new information.

Once it was confirmed that prominence most often coincides with the first syllable of the comment, measurements were made to find the prosodic correlates of the prominent syllable.

Firstly, measurements were made to define whether a specific pitch target can be associated with the prominent syllable. An analysis of the final velocity suggests a [fall] target. This target is most clearly evidenced in monosyllabic coverbs and bisyllabic coverbs, verbs, and nouns. The syllables that immediately follow the prominent syllable tend to take a [mid] target both after a monosyllabic coverb and as the second syllable in bisyllabic words. The pattern is somewhat spread in trisyllabic words, where the [fall] target occurs on the second syllable, whilst the prominent syllable receives a [high] target. This pattern is not unlike the patterns described by Singer (2006) for Mawng of a L+H* pitch accent which she interprets as 'a special focus pitch accent best analysed as a hyperarticulated instance of the main L+H*', or by Bishop and Fletcher (2005) of a special L+H* tone used in the Gundjeihmi dialect of Bininj Gun-wok. The similarities need to be viewed critically, as the pitch targets described for Jaminjung have the syllable as their domain, while the pitch targets in the AM model may extend over a much larger span. The problem of alignment would need further examination in order to establish the exact correspondences in the patterns across these languages.

The coverbs in Jaminjung demarcate themselves from the inflected verbs and nominals in their having more extreme values. This would warrant further investigation with more controlled data and a wider sample, if feasible. Further, the variation in the patterning of the targets in mono-, bi- and trisyllabic words in Jaminjung suggests that the [fall] target that marks the first word in the comment is in fact aligned with the left of the word, and that the variable linking of the [fall] to a syllable other than the

⁴³ In papers published between 1973 and 1983.

first could depend on the number of syllables in the word or possibly the presence of a preceding pause. The F0 measures confirm these findings.

Secondly, pitch excursions were examined and it was found that they tend to be wider in the prominent syllables than in the other syllables in the IU. This finding is in accordance with many other studies that have confirmed the fundamental role of pitch range in marking focus in typologically diverse languages including Mandarin, Japanese and English. In the context of Australian languages, it has been observed by Bishop (2003:17) for BGW, by Hellmuth et al. (2007) for Mawng, and by Birch (2003) for Iwaidja. As our data is taken from natural speech with virtually no control over the phonological content, it is not possible to test whether there is pitch range compression after the prominent syllable (focus) as noted for English and Mandarin (Xu 2005, 2009). Functionally, these pitch range manipulations would serve to increase the relative prominence of focal compared to non-focal material. This explanation can be linked with Gussenhoven's 'effort code' (2004 ch. 5). Indeed, Gussenhoven proposes that the interpretation of prosody is governed by both biological and properly linguistic codes. The biological codes include a frequency code, an effort code and a production code. The frequency code implies that a raised F0 is a marker of submissiveness or non-assertiveness and is associated with question intonation. The effort code implies that articulation effort is increased to highlight important focal information producing a higher F0 and pitch range. The production code associates high pitch with phrase beginnings (new topics) and low pitch with phrase endings. The results presented in this section suggest that the effort code could predict the encoding of focus in Jaminjung.

Lastly, the duration correlate was assessed. The question of whether to treat duration as a functionally relevant prosodic cue on its own or as a consequence of using higher F0 peaks is still not settled amongst prosodists (see Kügler 2008, Xu 2009, Mo 2008 *inter alia*). Most seem to agree, however, that duration alone is not a sufficient indicator of prominence, which is rather marked by multiple acoustic cues interacting with one another. From the articulatory-functional perspective, timing can be actively controlled for the sake of conveying information, in other words, duration is

one of the available parameters a speaker can manipulate to encode a communicative function. Duration is known to participate in marking focal contrast at least in English and in Mandarin where focus has been consistently found to lengthen the lexical item being focused. After experiments with speech generation of focus using the TA model, Pronom et al. (2009) found that larger surface pitch range expansion occurred when focus was generated with an expanded pitch range and an increased duration than when it was generated with pitch range expansion only. Thus, the duration increase under focus could be interpreted as a means of “allocating sufficient time for the focally expanded pitch range to be articulatorily realized” (Xu 2009:16). This interpretation could hold true for the Jaminjung data. The durational cue obviously singles out monosyllabic coverbs, which display longer duration than their plurisyllabic counterparts, be they coverbs, inflected verbs or nouns. This is explained by the phonotactic peculiarities of coverbs, which can be formed of closed syllables with consonant clusters in codas not found in other word classes. Nonetheless, overall, prominent syllables have longer durations than other word-initial syllables, and even than word-final syllables which make it a strong cue for prominence.

4.3 Comparing the three focus domains

So far, the correlates of the prominent syllables in the comments have been measured, and compared to those of non-prominent syllables, while the effects of factors such as numbers of syllables in the word, or syntactic constituency have been considered. In this section, a comparison is undertaken between comments and the two other domains of focus, the sentence and the argument, to establish whether they make use of the same encoding parameters.

Each domain will be described and illustrative examples given. The comparison will concern the pitch target associated with the prominent syllable through an evaluation of final velocity; mean F0, pitch excursion and duration will also be considered. Note that for the purpose of this

comparison, only IUs with a prominence in the first syllable in the initial constituent of the focus domain are included.

4.3.1 Hypotheses

The hypotheses that underlie the analysis ofthetic sentences in Jaminjung are two-fold. The first hypothesis concerns one of the attested formal correlates of the thetic-categorical distinction and is based on Lambrecht's (1994:235) proposition that sentence focus needs to be minimally distinct from predicate/comment focus, namely by detopicalizing the subject⁴⁴ that would normally be interpreted as the topic in a predicate focus sentence. To do so, languages may make use of various strategies, amongst which are *prosodic inversion*, in which subjects are prosodically marked by an accent instead of the predicate in a topic-comment construction, as in English; or *syntactic inversion*, in which SV order is changed, as in Italian or in Russian; or by using cleft structures (Lambrecht 1994:242) (§4.1.2.2). In Jaminjung, subjects/agents are expected to be marked with a prominence, as changing position in the IU is not a likely strategy in a language with relatively free word order such as Jaminjung; at least it is expected that the part of the sentence that would correspond to the comment should not be singled out by prosodic means.

With regard to the argument focus, it is expected that the argument in focus is shifted to IU-initial position, or at least that the position of the argument in the IU affects the encoding of its prominent syllable. It is also expected that its correlates will be more salient than that of other focus domains.

Before proceeding to the comparison, the datasets used for the analysis of the sentence and argument focus domains will be introduced in detail, as well as the locus of the perceived prominences in each of them.

⁴⁴ Of course, in topic-comment structure, the topic does not necessarily correspond to the subject.

4.3.2 Sentence focus (thetic)

In *sentence focus*, the scope of focus is the entire sentence; in the literature the term *thetic* is also used. The former is preferred in this analysis, as it applies specifically to the scope of the focus, but both terms may be used interchangeably hereafter. IUs in this dataset are defined as *comments on a situation*, rather than as *information about* some entity that is already under consideration (a topic). In sentence focus structures, according to Lambrecht (1994:233), the assertion and focus coincide, “no pragmatic presupposition is formally evoked”. In most cases, both the participants and the event are new, that is, not identifiable by the hearer. Sasse (1987, 2006) describes the main functions of sentence focus/thetic utterances as *annuntiative*, which include out-of-the-blue utterances, exclamations, headlines; *introductive*, which introduce new discourse topics or new referents; *interruptive*, when an event interrupts the normal topic chain; *descriptive*, for scene-setting, usually employing of existential or related verbs; and finally, *explanative*, for situations where an answer is given to the question ‘what happened’, or when an explanation of a preceding discourse context or non-linguistic situation is provided. These characteristics provide the identificational criteria on which the selection of the IUs in this dataset is based.

4.3.2.1 Dataset

Sentence focus/thetic sentences have structures that can vary significantly in Jaminjung. Some of the possible candidates are listed here (where A is Agent):

Intransitive clauses: NP (subj) + (coverb) V + NP or
NP (subj) + V + NP

Transitive clauses: NP (subj) + NP + (coverb) + V or
NP (subj) + (coverb) V + NP (arg (other than subj))

The dataset consists of 20 IUs, 14 of which are intransitive and 6 are transitive. 14 tokens have structures that have an NP subject in initial position, 2 have similar structures but have a linking word in initial position;

and 4 have the complex verb in initial position. Table 4-7 shows the counts and percentages for these structures.

Structure	vtr	vintr	Total	Percent
*NP(subj)+V	2	6	8	40%
NP(subj) + *V	0	1	1	5%
*NP(subj) +CV V	1	2	3	15%
*NP(subj) +NP(arg) V	0	1	1	5%
splitNP ⁴⁵ (subj)+*NP(arg)-V-NP	0	1	1	5%
other+*NP(subj)+V	0	2	2	10%
*CV V+ NP(subj)	1	1	2	10%
*V + NP(other)	1	0	1	5%
*V	1	0	1	5%
Total	6	14	20	100%

Table 4-7 The structure of the thetic sentences in the dataset.

The following examples provide an illustration of the word orders found in the dataset: the first with the more common subject-verb structure and the second with the reverse order, verb-subject. In (68), the speaker is describing a picture showing two children fishing on a riverbank; one of them has caught a crocodile. A third child is seen running towards them in the background. The speaker has talked about the child catching the crocodile and is now talking about the child in the background. The speaker does not expect the hearer to have noticed that other child, thus every constituent in the utterance is new. The focus domain is the whole utterance. Note the prominence of the first syllable of the IU compared with the first syllable of the complex predicate (most frequent locus of prominence is topic-comment structures (see further discussion in §4.3.3 below).

(68)
 jarlig yugung ga-ram
 child run 3sg-come.PRS

'A kid comes running.'
 [JoJ:CS07_65_01]

⁴⁵ It has been proposed that split NPs may be a strategy used to mark thetic sentences in Jaminjung (Schultze-Berndt, p.c.)

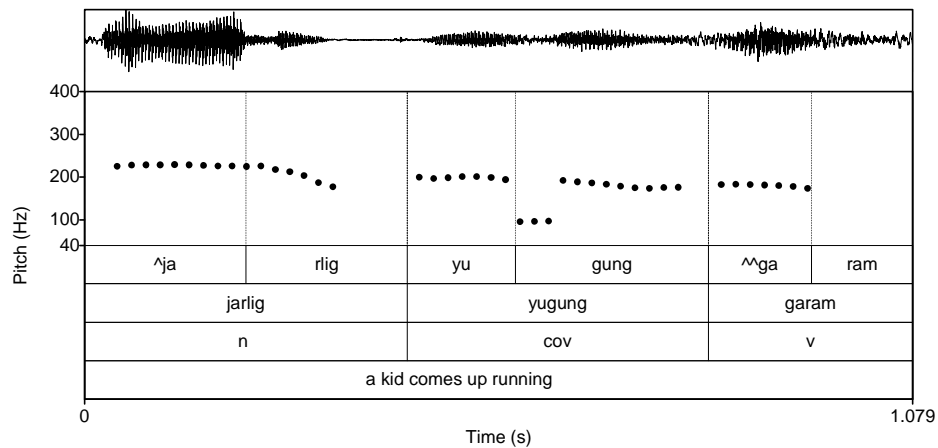


Figure 4-17 Sentence focus construction in annuntiative function; all the constituents are new.

There are other word orders in our dataset. Subjects/agents in some utterances are expressed by NPs placed to the right of the verb. In example (69), the IU is interpreted as athetic sentence with the subject *mali bidiyo* ‘the video machine’ placed after the complex verb *girrb=biya ganiyu* ‘it shut up now’. Here the event interrupts the normal topic chain, that is, the speaker interrupts her elicitation task to comment on the recording equipment.

(69)
 girrb=biya gani-yu mali <bidiyo>
 quiet=now 3sg:3sg-say/do.PST thing video

‘It shut up now, the video thing.’

[IP:ES96_08_02]

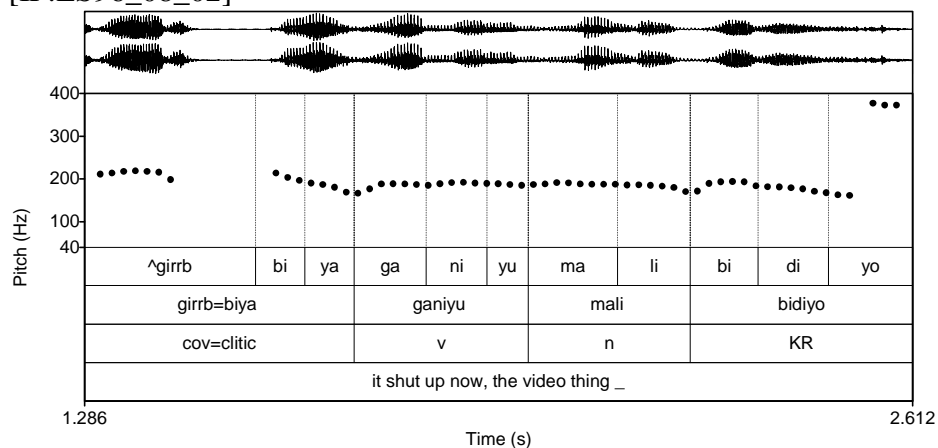


Figure 4-18 A thetic sentence where the speaker interrupts the telling of a story to comment on the recording equipment, thus used in its *interruptive* function.

4.3.2.2 Prominences

Prominences are assigned primarily on auditory perception, prior to analysis and measurements. The labelers (author and one collaborator) agreed that they often heard more than one prominence, in fact, for most of the IUs, two prominences were assigned, and thus the salience of the most prominent syllable appeared to be less marked than in the comments.

The syllable perceived as most prominent is labeled (^) in the textgrids and that with a lesser (or secondary) prominence with (^ ^). Syllable position, and grammatical constituency are also labeled in addition to the order in which they occurred in the IU.

Table 4-8 shows that the main prominence is heard on the first syllable of the first constituent in 16 of the 20 tokens. It is occasionally heard on the first syllable of the second (2 tokens) or of the last constituents (1 token); only once is it perceived on a syllable that is not word-initial, that is, on the second syllable of the first constituent.

syll_position_const	Most prominent	Secondary prominence
first/1st const	16	1
mid/1st const	1	0
last/1st const	0	0
first/2nd const	2	1
mid/2nd const	0	0
last/2nd const	0	0
first/3rd const	0	0
mid/3rd const	0	0
last/3rd const	0	0
first/final const	1	10
mid/final const	0	3
last/final const	0	1
Total	20	16

Table 4-8 Position in the constituent of primary and secondary prominences in sentence focus.

Table 4-9 below links the main prominence to the corresponding syntactic constituent (indicated by a * before the constituent). The prominence is heard on the first syllable of subject NPs in IU-initial position in 12 tokens (light grey in table), and on the first syllable of the complex verb in initial position (either on the coverb or the inflected verb) in 4 tokens (darker grey). There are 2 tokens where the main prominence is

heard on the first syllable of an NP subject that comes immediately after a linking word, such as the Kriol particle *an* ‘and’ used to introduce weather-related events as in ‘and the rain came’ (they appear in the table under first syllable of the second constituent). Excluding these tokens, there are only two tokens where the primary prominence is heard on a syllable that is not in IU-initial position.

Structure	first/1st const	first/2nd const	first/final const	Total
*NP(subj)+V	8	0	0	8
*NP(subj) + cv V	3	0	0	3
*NP(subj) +NP(arg) V	1	0	0	1
other+*NP(subj)+V	0	2	0	2
*V+NP(other)	1	0	0	1
*cv V+ NP(subj)	2	0	0	2
*V	1	0	0	1
NP(subj) + * V	0	0	1	1
splitNP(subj)+*NP(arg)-V-NP	0	1	0	1
Total	16	3	1	20

Table 4-9 Rates of occurrences of primary and secondary prominences according to the syntactic structure of the sentence in focus.

Before proceeding with the comparison between the focus domains, the primary and secondary prominences in sentence focus constructions are examined more closely in order to assess whether prominences really encode focus or simply serve to mark grouping (Figure 4-19). A test comparing the correlates in the words with primary and secondary prominences shows that overall the primary prominences tend to have higher mean F0 throughout, which is expected as they occur IU-initially. The pattern of higher mean pitch in the second syllable of the words with primary prominences (discussed further below) does not occur in those with secondary prominences, although there is a wide variation in the values for these syllables, suggesting the that first prominences encode focus, while the second do not.

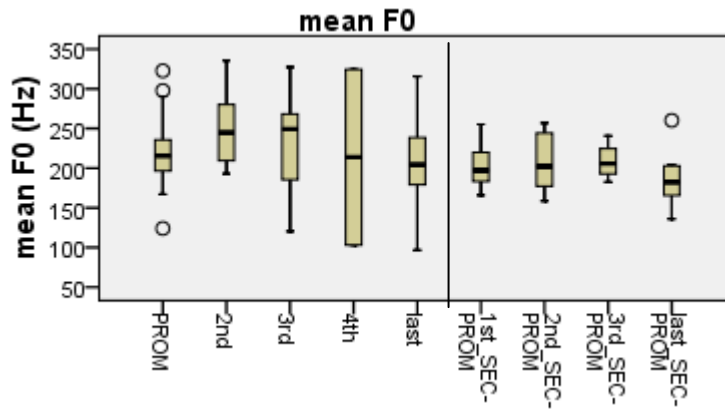


Figure 4-19 A comparison of the words with perceived primary and secondary prominences.

The final velocity values in primary prominences and the following syllables are on average 1.13st/s and -19.24st/s (Figure 4-20). Those of secondary prominences are -7.18st/s and -11.94st/s. The values of mean F0 and final velocity suggest different patterns in underlying targets for primary and secondary prominences.

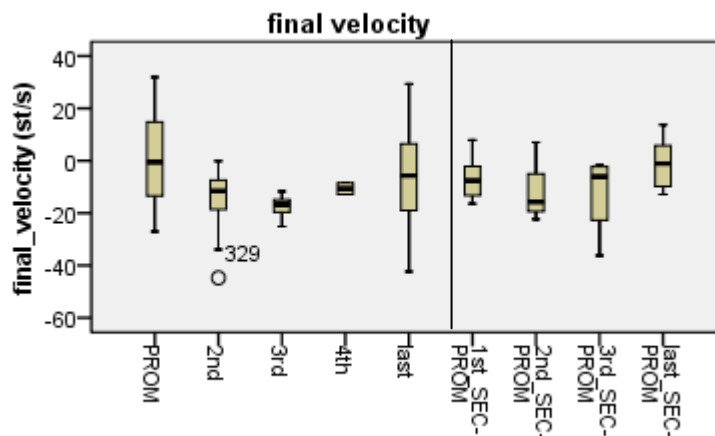


Figure 4-20 A comparison of the final velocity values in the syllables of the constituents containing primary and secondary prominences.

Regarding excursion size, the averages are 2.99st and 2.08st for the primary and secondary prominences, as shown in the boxplot in Figure 4-21.

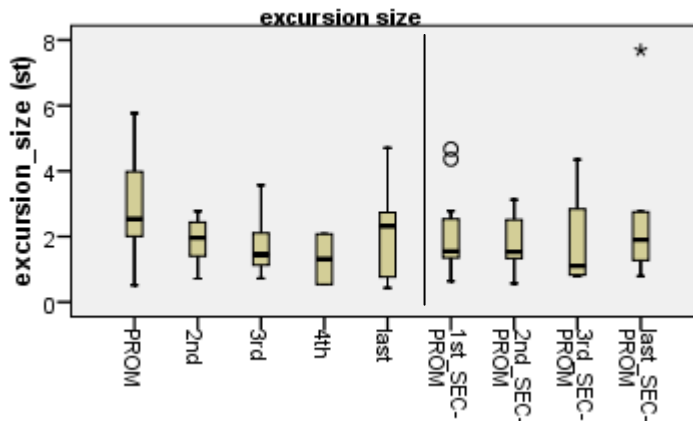


Figure 4-21 A comparison of the excursion sizes in the syllables of the constituents containing primary and secondary prominences

The average duration of the primary prominence is 161.63ms compared with 159ms in secondary prominences. These results suggest that duration is not used to encode sentence focus (Figure 4-22).

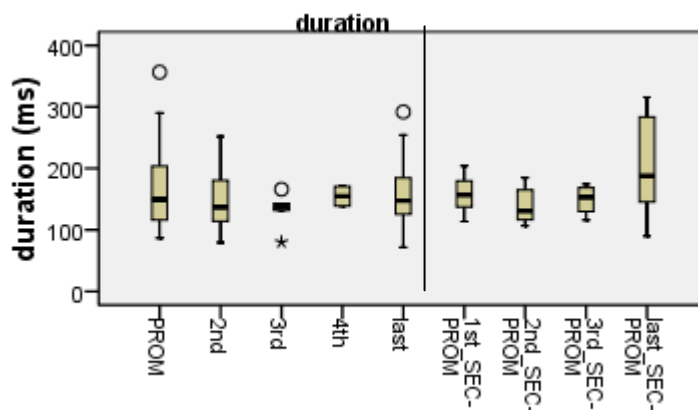


Figure 4-22 A comparison of the duration values in the syllables of the constituents containing primary and secondary prominences.

This comparison between the correlates reveals different patterns for primary and secondary prominences. It is therefore reasonable to assume that they reflect differences in communicative functions, hence that primary prominences are used to encode focus, while secondary prominences are not.

To sum up, in most IUs in the dataset, a primary and a secondary prominence are perceived. Secondary prominences are often assigned to the first syllables of any major syntactic constituents that do not receive the main prominence in the IU.

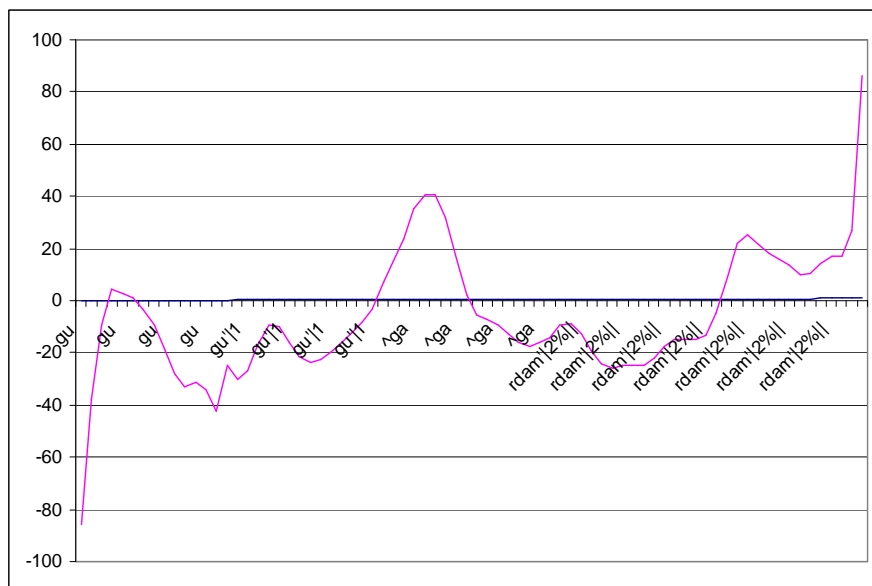
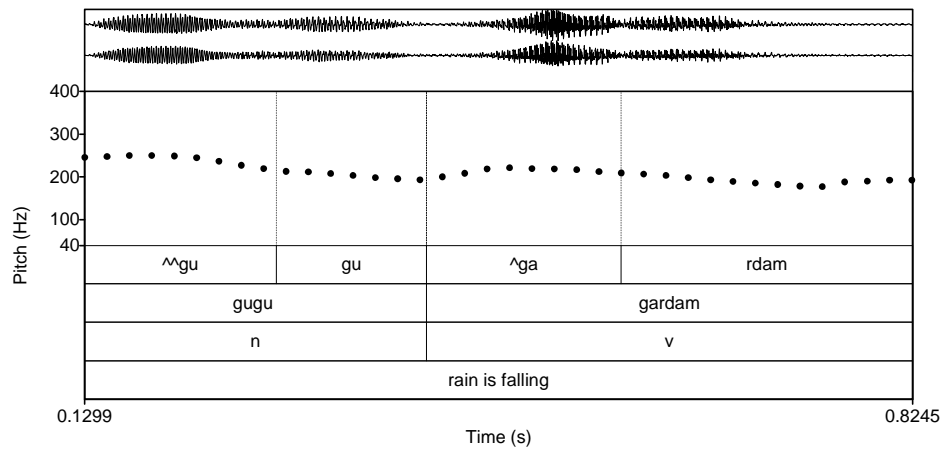
The main prominence is mostly heard on the first syllable of the constituent in initial position in the IU, thus on the first syllable of the focus domain, just as in comments. This position is usually occupied by an NP subject, but not exclusively. By receiving the main prominence, the subject NP in sentence focus constructions is marked differently to the topics in topic-comment structures, which agrees with Lambrecht's (1994:242) proposition that subjects are 'detopicalised' in sentence focus constructions.

Examples (70) and (71) illustrate thethetic sentences in this dataset. Both sentences refer to a weather event; descriptions of the weather are actually typical sentence-focus constructions in the languages of the world (Lambrecht 1994:140). Both pitch tracks are very similar, but the perceived prominences differ. In the first example, the main prominence is heard on the first syllable of the second constituent [^]*gardam* 'it falls', and in the second example, it is heard on the first syllable of the first constituent, [^]*gugu* 'rain'. This perceptual prominence is corroborated in the velocity curves, displayed underneath, which suggest different underlying targets for the syllables. These are discussed further below.

(70)

gugu ga-rda-m
water 3sg:fall-PRS

'Rain is falling.'
[IP:ES08_16_02]



(71)
 gugu ga-ruma-ny
 water 3sg-come-PST

'The rain came.'
 [BH:CS07_72_01]

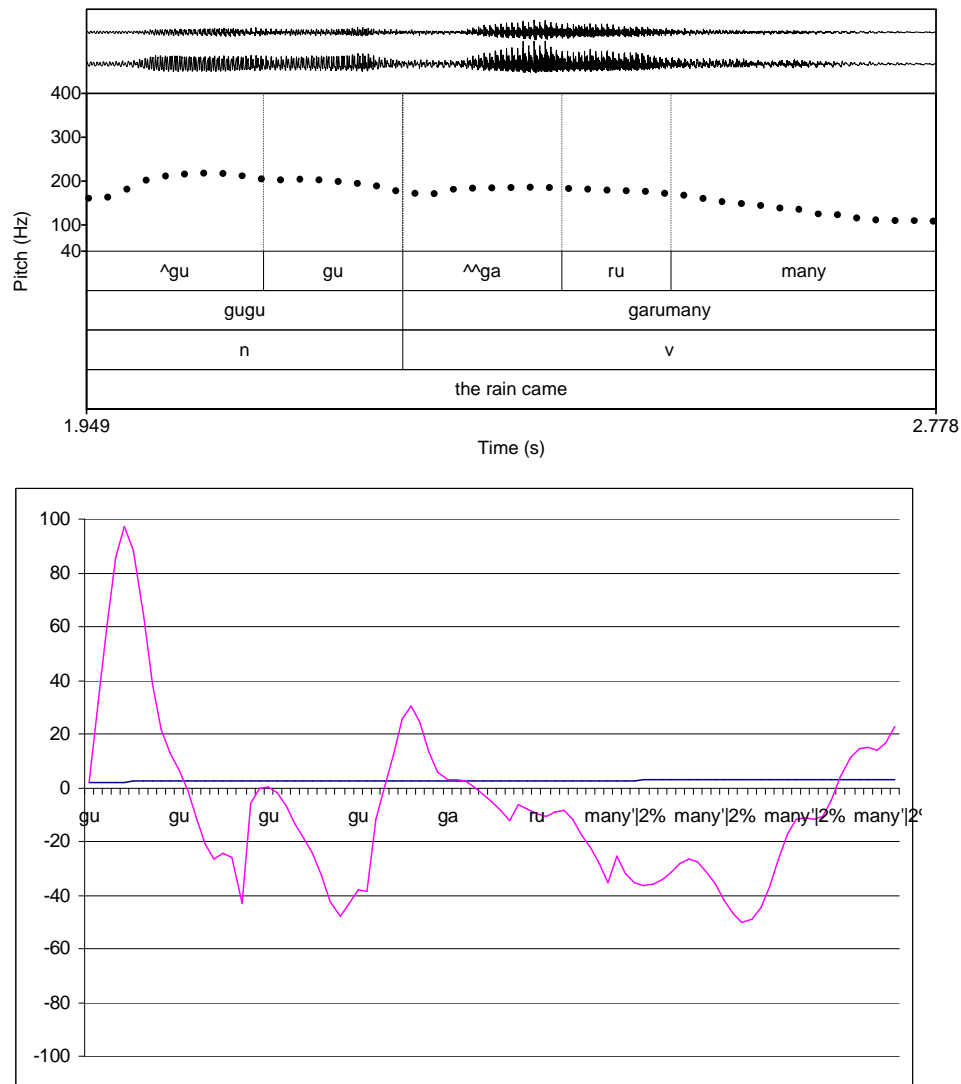


Figure 4-23 Pitch traces and velocity measurements illustrating patterns of prominences in thetic sentences. The top pane shows a main prominence heard on the first syllable of the second word, *gardam*; the bottom pane shows a prominence on the first syllable of the first word, *gugu*.

4.3.3 Argument focus

The domain of focus can also be restricted to a sole argument. If the focused argument is a subject/agent, it can not be the topic of the sentence;

focused arguments may also be objects or oblique. Focus is determined in reference to the context; an example (72) is given in Figure 4-24. One speaker is telling the story of a mythological being, and the listener is asking for some clarifications. The first part of the display shows the question, and the second part, on the right, is the answer to the question. The argument in focus, the object NP *nguyung bunyagna* ‘their husband’, follows the interjection *ya* ‘yes’. It is placed at the beginning of the IU, preceding the inflected verb *bunybardagarrany* ‘they followed’. There is no overt subject/agent in the IU, as is very common in Jaminjung.

(72)
 thanthu=biyang luba=yirram Jiniminy buny-bardagarra-ny
 DEM=now many=two ghost.bat 2du-follow-PST

ya nguyung bunyag =<na>? buny-bardagarra-ny
 yes husband 3DU.OBL =now 2du-follow-PST

‘Those two big ones followed the Bat?’
 [EH-IP:ES08_04_02]

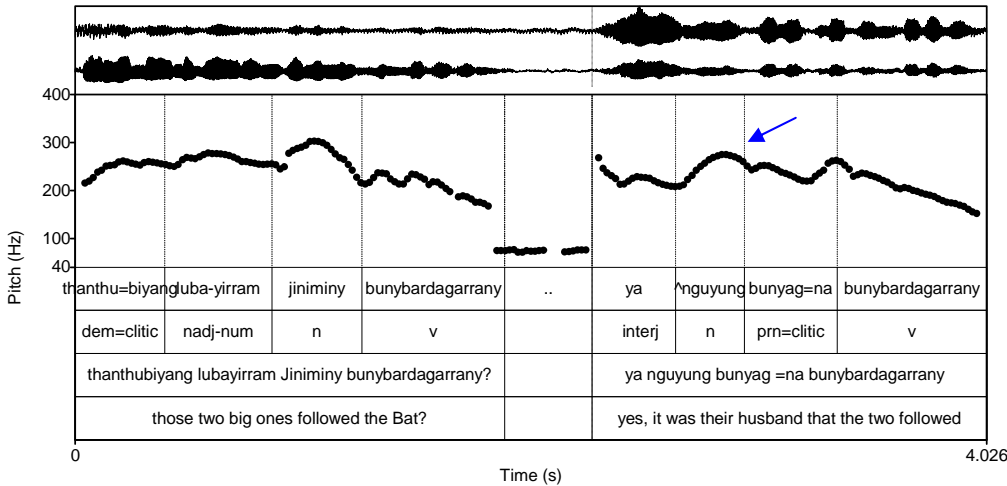


Figure 4-24 A question-answer pair, illustrating a focused argument in the answer, *nguyung bunyag=ma*.

Focused arguments can be subdivided into contrastive and non-contrastive subgroups; and each subgroup is differentiated along criteria defined as follows.

Non-contrastive argument focus is identified where:

- the elements in the IU are repeated from another IU, the argument being the sole new information;

- in question-answer pairs, where the argument appears as the answer to the question;

Contrastive argument focus is identified where:

- explicit contrast is expressed such as ‘W, then X, but Z, then Y’. These are restricted to tokens where only the NPs are contrastive; tokens where the whole predicates are contrasted have been excluded.
- in instances of corrective focus or replacing focus, including tokens with explicit corrections (also metalinguistic corrections, e.g. replacing a previous Kriol word, or structures like ‘not X, Y’). The selection is restricted to corrections involving NP arguments only.

For the purpose of comparing the three focus domains, only non-contrastive arguments in focus are retained. Contrastive arguments will be analysed in §4.6.1 below.

Argument focus is treated as an equivalent to *information focus*, a term used in some studies, and best illustrated in answers to constituent questions.

4.3.3.1 Dataset

The dataset used for this comparison consist of 35 tokens. 19 are sentences that repeat something that has just been uttered and add a new argument that is focused, and 16 are answers in a question-answer pair. Question words in interrogatives are also focused arguments, but only arguments occurring in declaratives are retained for the purpose of this analysis. Table 4-10 summarizes the criteria and number of tokens in the dataset.

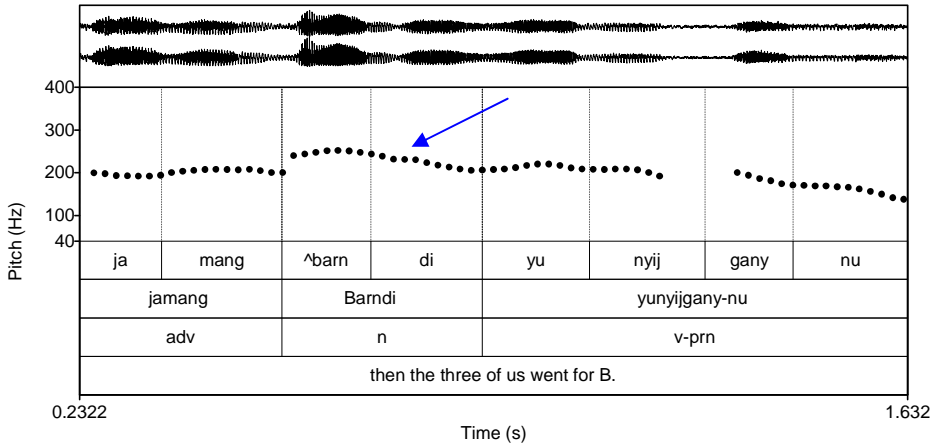
Type	Selection criteria	Number of tokens	Prominences
Non-contrastive	Repetition (but not a case of contrast) and only one argument is new	19	19/19 first syllable
	Clear cases of question/answer pairs	16	13/16 first syllable

Table 4-10 The dataset used for the analysis for the argument focus domain.

Examples are provided in Figure 4-25. Example (73) is an instance of argument focus. The speaker is relating the events of the previous day, mentioning a series of stopovers before setting off for a fishing trip, so that the ‘going’ and ‘stopping over’ are not new information, the only new information being the name of the person they picked up on the way. Example (74) is a question-answer pair. Two speakers are reminiscing about the tools people used to make. One of them asks what other kinds of hooks (for hunting or fishing) they used to make. The answer is from the second speaker, on the right-hand side of the pane. The first word of the answer *waya* ‘wire’ is the focused argument.

(73)
jamang B. yuny-ijga-ny=nu
finished B 13du:3sg-go-PST=3SG.OBL

'Then the three of us went for B.'
[EH-IP:ES08_04_02]



(74)

<en> nganthan=biyang <huk> burra-mila
and what=now hook 3pl:3sg-get/handle:IMPF

<waya> burra-linymana
wire 3pl:3sg-make-IMPF

'And what hooks did they make?' 'They used to make wires.'
[EH,IP:ES08_04_02]

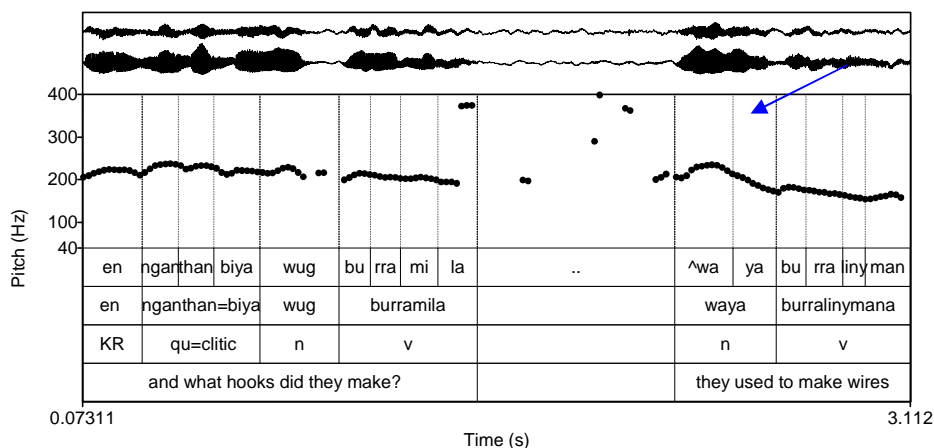


Figure 4-25 Example sentences illustrating different types of argument focus constructions. In the top pane (64), the comment repeats the information in the previous IU, the only new information is the focused argument, *Barrdi*. In the second pane (65), the argument in focus is contained in the answer, *waya*.

The structures of the IUs in this dataset are not dissimilar to those of the topic-comment structures described in §4.2.2.2, the main difference being that in the argument focus constructions, the domain of the focus is restricted to the argument, rather than comprising the entire comment.

Verbless clauses:

[NP]_{top} [NP_F]_{com}

Intransitive clauses:

[NP]_F [(coverb) V]_{Pred} or

[NP]_F [(coverb) V (OBL)]_{com}

Transitive clauses:

[NP]_{top} [NP_F (coverb) V]_{com} or

[NP]_F [[NP_F (coverb) V]_{com}

[NP]_F [(coverb) V NP]_{com} or

[NP]_{top} [(coverb) V NP_F]_{com}

The most common structure is of an argument preceding a complex verb, accounting for 57% of the IUs in this dataset, and spread over both

subtypes of information focus. The argument may also follow the complex verbs, as in 15% of tokens, which are all IUs where all is repeated from a previous utterance except the focused arguments. Another frequent structure, in another 15% of tokens, is a verbless clause (predicate NP), more often found in answers to questions. Finally, the complex verb in the comment sometimes has two arguments: in 11% of tokens, they are more likely to be found on each side of the verb, but there is one token where both arguments precede the complex verb (3%), the argument in focus may be either the first or the second NP. Table 4-11 below summarizes this information.

Structures of comments		type		
		Repeated IU-Arg	question-answer	Total
arg_complexV	Count	10	10	20
	%	52.6%	62.5%	57.1%
complexV_arg	Count	5	0	5
	%	26.3%	.0%	14.3%
NP-pred_arg	Count	1	4	5
	%	5.3%	25.0%	14.3%
arg_complexV_arg	Count	3	1	4
	%	15.8%	6.2%	11.4%
arg_arg_complexV	Count	0	1	1
	%	.0%	6.2%	2.9%
Total	Count	19	16	35
	%	100%	100%	100%

Table 4-11 The distribution of the subtypes of arguments in focus, related to their syntactic structure.

4.3.3.2 Prominences

The prominences in the IUs are labeled based on perception by the same two labelers as with the comment and sentence focus domains.

Table 4-12 gives a detailed account of the position of the prominent syllable within the argument, together with the position of the argument within the comment.

The most common pattern (66%) is for the prominence to fall on the first syllable of an argument in initial position in the comment. In fact, the prominence is usually heard on the first syllable of the focused argument,

wherever it occurs in the IU. For example, there are a few instances where the speaker chooses to emphasize the second word in an argument. There are only two tokens where the prominence falls on the second syllables. In short, regarding the prominences in the argument focus domain, the first syllables of an argument in comment-initial position are favoured.

Locus of prominence and Syllable position in comment		type		
		Repeated IU-Arg	question-answer	Total
1st syll_arg-CV	Count	8	10	18
	%	42.1%	62.5%	51.4%
1st syll_1st word_NPpred	Count	0	4	4
	%	.0%	25.0%	11.4%
1st syll_1st arg_arg-CV-arg	Count	1	0	1
	%	5.3%	.0%	2.9%
	subtotal	9	14	65.7%
1st syll_2nd word_arg-CV	Count	1	0	1
	%	5.3%	.0%	2.9%
1st syll_2nd arg_arg-arg-CV	Count	0	1	1
	%	.0%	6.2%	2.9%
1st syll_2nd word_NPpred	Count	1	0	1
	%	5.3%	.0%	2.9%
1st syll_2nd arg_arg-CV-arg	Count	2	1	3
	%	10.5%	6.2%	8.6%
1st syll_CV-arg	Count	4	0	4
	%	21.1%	.0%	11.4%
2nd syll_2nd word_arg-CV	Count	1	0	1
	%	5.3%	.0%	2.9%
2nd syll_CV-arg	Count	1	0	1
	%	5.3%	.0%	2.9%

Table 4-12 Prominences shown by syntactic structure and argument position.

The structures of the IUs and more specifically of the comments in this dataset vary. The focused arguments tend to be in IU-initial position in 2/3 of the tokens. Otherwise, they may occur either after another argument or after the complex verb. In all but a few tokens, the prominence is heard on the first syllable of either the first or the second word of the argument.

In comment focus and sentence focus domains, the prominent syllable is found mostly on the first syllable of the first word in the domain in focus. In argument focus, it is still on the first syllable of the focus domain, even if this word is not in IU – or even comment-initial position. A test is done to see if the correlates of mean F0, final velocity, extension size, and duration vary significantly when the prominence is on arguments that are not in IU-initial position. Out of all the comment structures described above, four subgroups are created to reflect the position of the prominence in the focused argument:

- when it occurs on the first syllable of an argument in initial position in the comment (1st_comment);
- when it occurs on the first syllable of an argument that appears elsewhere in the comment (1st_arg. elsewhere);
- when it occurs on the first syllable of the second word in an argument (1st_2nd word in arg);
- when it occurs on the second syllable in an argument (2nd syll in arg.).

The results show no significant differences between the four subgroups for all of the correlates tested. However, with a view to creating as homogeneous a dataset as possible for the comparison of the focus domains, only the IUs with a prominence in the first syllable in the argument in initial position are included.

4.3.4 Comparison: pitch target, final velocity and mean F0

The correlate of final velocity is examined first as being the best indicator of the syllables' underlying targets. The values for the three focus domains are shown below. They do not differ significantly between domains but the differences according to syllable positions differ significantly ($F(21, 727) = 1.866, p=.011$). Figure 4-26 illustrates the results.

The first syllables (PROM) have values of -9.65st/s in comment focus, 1.41st/s for sentence focus, and -5.54st/s in argument focus domain. In the second syllables, the values fall sharply in all three domains, with -18.33st/s, -19.24st/s, and -14.7st/s, respectively. The values of the prominent syllables in the comments and in the arguments, both negative, suggest a [fall] target, while the values slightly above zero in the utterances could be interpreted as a [high] pitch target and the [fall] is realized only in the second syllables in these constructions. This interpretation is supported by the visual inspection of the pitch tracks which display a 'peak' on, or slightly after, the initial syllable. The measurements of mean F0 are examined next.

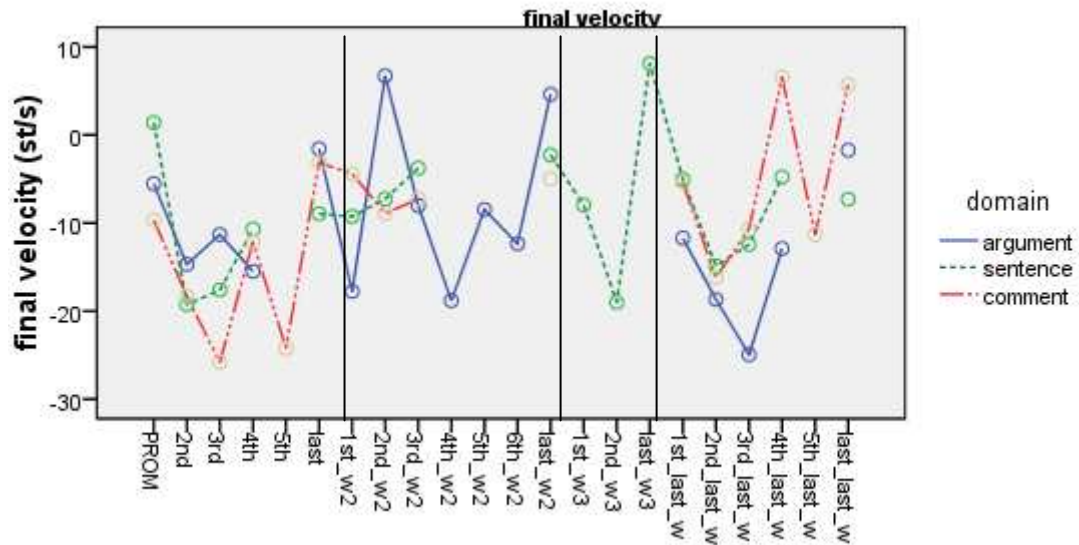


Figure 4-26 Final velocity measurements in the tree focus domains, according to syllable position.

The differences in the mean F0 values between syllable positions are significant ($F(22, 754) = 4.652, p=.000$), as well as the difference between the domains ($F(2, 754) = 48.327, p=.000$). The mean pitch of the syllables in each domain is shown in Figure 4-27 below.

The overall F0 patterns are interesting. Comment focus has the lowest mean pitch over all its syllables, while sentence focus has higher pitch. NPs in argument focus start with high pitch, which then diminishes in the succeeding words to reach levels similar to that of comments in IU-final syllables. This convincingly illustrates the role of high pitch in encoding focus, used in both argument and sentence focus, and in confirming the ‘unmarked’ status of the comments.

The values for the first and second syllables are 207.73Hz and 200.58Hz in comment focus, 224.2Hz and 232.89Hz in sentence focus, and 233.87Hz and 230.45 Hz in argument focus. While there is a slight decrease in comments and arguments, the values in the sentence focus domain rise from the first to the second syllables. This agrees with the final velocity values discussed above, where the peak associated with a [fall] target does not occur until the second syllable. It is thus suggested that focus is encoded by a [fall] target, the exact alignment of which may vary from the first to the second syllable.

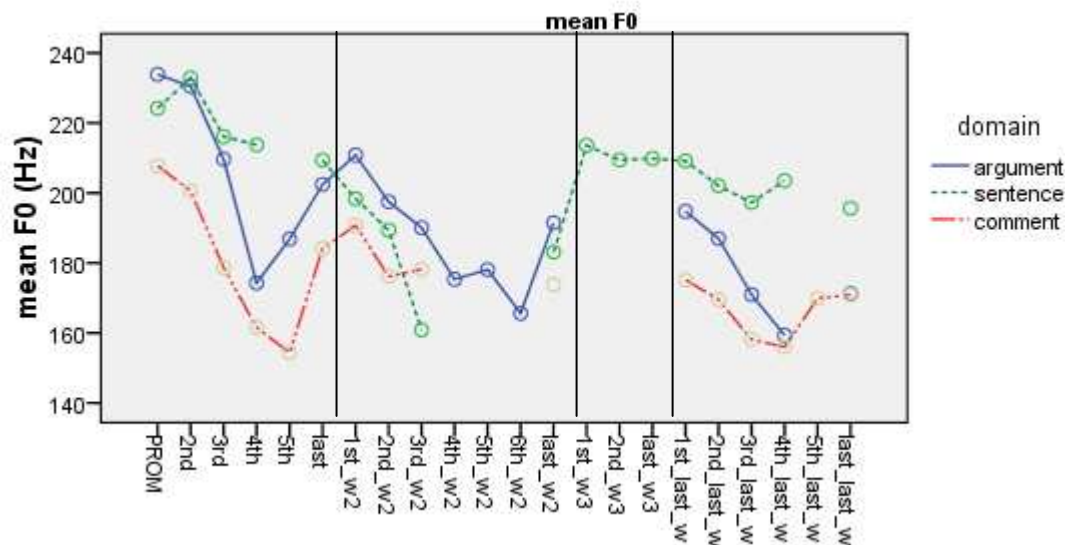


Figure 4-27 Mean F0 values of the syllables in the three focus domains. The words are separated by vertical lines.

4.3.5 Comparison: pitch excursion

The measures of the excursion size in the prominent syllables of each domain are shown in Figure 4-28 below. The difference between the syllables by positions is significant ($F(21, 739) = 2.834, p=0.000$), also that between the domains ($F(2, 739) = 10.716, p=0.000$). Again, the overall differences seem to be maintained throughout at least the word containing the prominence, and possibly during the entire IU.

Prominent syllables in focused arguments have the widest excursions, with 4.96st, followed by the comments with 4.03st, and the sentence focus domains with 3.15st. The erratic patterns of excursions in the first syllables of other words in the IUs (1st_w2, and 1st_last_w) highlight the patterns observed in the prominent syllables, further evidence that wider excursions are used as cue for focus.

For the correlate of pitch excursion, sentence focus is the least marked, followed by comments and arguments, supporting the general hypothesis that the larger the domain of focus, the less marking it receives.

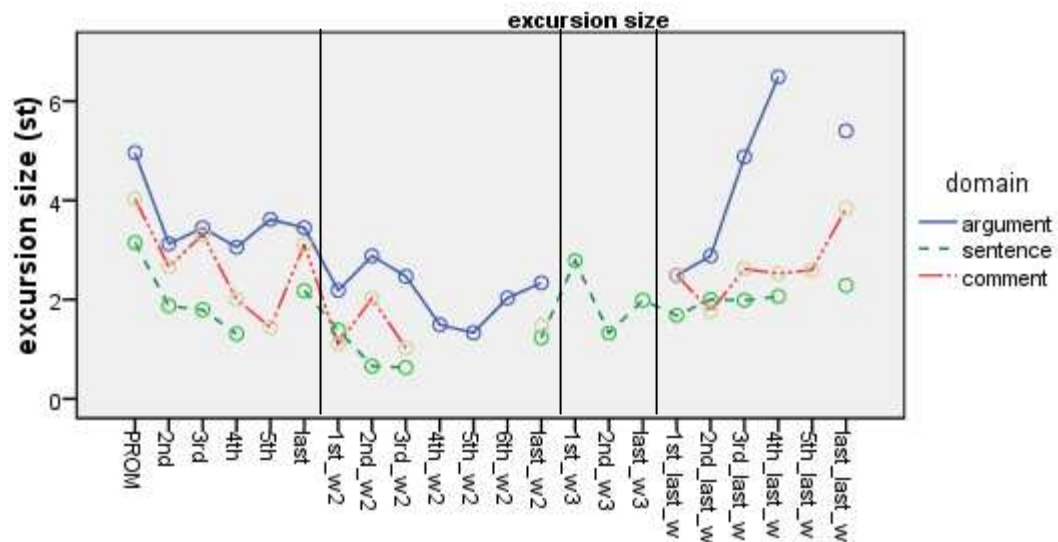


Figure 4-28 Mean excursion size values of the syllables in the three focus domains.

4.3.6 Comparison: duration

The duration of the syllables in the IUs are shown in Figure 4-29. The differences are statistically significant regarding the position of the syllables ($F(21, 764) = 4.863, p = .000$), but not regarding the domains. The prominent syllables in focused comments have longer durations than those in focused arguments with 220.40ms to 186.33ms, and sentence focus constructions have the shortest prominent syllables with averages of 168.68ms. The durations of the second syllables in all three domains show a decrease compared with the first syllables. The overall pattern, contrary to those observed for the other correlates, is only sustained throughout the first word. After the first word in the focus domain, the patterns become more variable, and no pattern (except possibly final lengthening) is easily discernable.

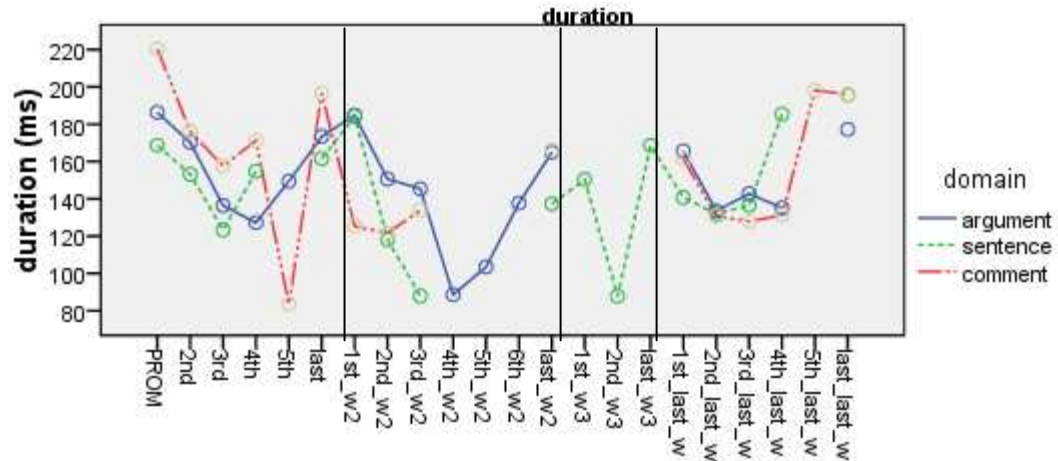


Figure 4-29 Durations of syllables in the prosodic words for each focus domain.

In summary, final velocity measurements in the first two syllables point to a [fall] target for comments and argument focus, and a [high] followed by a [fall] in sentence focus. The three domains vary significantly in the mean pitch over all their syllables, where arguments and sentence focus have higher values, and comments have substantially lower values throughout because they are IU initial. The differences in final velocity between the prominent and second syllables observed in the final velocity are reflected by the mean F0 values, with sentence focus having a higher mean F0 in the second than in the first syllables. The results of the final velocity and mean F0 measurements suggest that focus is encoded by a [fall] underlying target realized at the left edge of the focus domains, the exact alignment of which may vary from the first to the second syllable.

As to pitch excursion, the widest are found in arguments, followed by the comments, and finally by the sentence focus constructions.

Finally, the prominent syllables in comments have significantly longer durations, followed by the argument and the sentence focus domains. As mentioned in the discussion of the comments (§4.2.3), these longer durations may be explained by the high degree of occurrences of monosyllabic coverbs in this position. However, as longer durations are found in all domains, this correlate is retained as a valid parameter for encoding focus.

4.4 Conclusion and Discussion: focus domains

In all three domains, the primary prominence is generally heard on the first syllable in the focus domain, i.e. the first syllable of the entire IU for sentence focus, the first syllable of the comment for comment focus, and the first syllable of a focused NP for argument focus. Since the position of focused argument within the comment may vary potential effect of this position were tested. The results show that the correlates of the prominent syllables in argument focus do not differ significantly with the position of the focused NP in the comment.

In the case of sentence focus, the perception of prominences is not as unambiguous as in the other focus domains, and, although a prominence is usually heard on the first syllable of the first word, other syllables may also be perceived as prominent. These secondary prominences usually coincide with the first syllables of the other constituents in the IU. A test was performed to assess whether primary and secondary prominences in sentence focus really encode focus or simply serve to mark grouping. The measurements show that the correlates of the syllables with primary prominences differ from those with secondary prominences, notably in final velocity and excursion size. They also differ from those of topics in IU-initial position which are described in the following section (4.5). This suggests that primary prominences encode focus, and secondary prominences are associated with the boundary marking of major constituents and thus the grouping function. This interpretation is contrary to Féry and Ishihara (2009:38), who state that “(i)n an all-new sentence, that is, in a sentence without topic-comment organization, the formation of prosodic phrases as well as the tonal pattern and scaling depend entirely on the morpho-syntactic structure”.

In Jaminjung, the position of the constituents in focus is determined by the principle that the first position of a larger domain is preferred for the focused constituent. This does not fully accord with Lambrecht’s prediction that in *focus-first languages* the initial position is exclusively reserved for focal material (1994:199-201) as, for example, a comment in focus is not necessarily IU-initial – it can be preceded by a non-detached overt topic –

the same can be said about an NP argument in focus; as to the sentence focus, the principle of initial prominence requires IU-initial prominence, syntactically, the constituent placed there tends to be the subject, resulting in one of the cross-linguistically attested patterns for distinguishing sentence focus from comment focus, that of subject accenting.

Jaminjung favours moving constituents around instead of shifting the accent, a strategy that is not unexpected in non-configurational languages. In configurational languages, the syntactic functions of subject and object can be deduced from their position in the sentence. Hale (1983) was the first to describe Australian languages as *non-configurational* in reference to the relative free ordering of words at the clause level, implying that they allow more than one possible ordering of constituents. Thus, in such languages, the ordering of major constituents is not based upon grammatical function but reflect some other functionality. Word order may not be constrained grammatically, but may be conditioned by information structure, hence the placement of the focus material in IU-initial position in Jaminjung, where it receives the main sentence accent, understood in the sense of Kohler (2006), as “refer[ing] to the perceptual salience of some words over others in utterances “. This conforms with Van Valin and La Polla (1997:213), who state that in languages where word order is highly constrained, prominence placement is very flexible, as in English, but in languages where word order is very flexible, focus placement is very constrained, as in Italian – and in Jaminjung. They further propose that a typological categorization could be established based on the interaction of syntax and focus structure.

The results of our measurements and statistical tests strongly suggest that focus in Jaminjung is encoded by a [fall] target which is realized on the first syllables on the focus domains, the exact alignment of which may vary between the first to the second syllable. The differences observed in the three domains may be explicable by factors such as the position of the focused element in the prosodic sentence, or its part-of-speech. The results presented here, including the detailed analysis of the comments in §4.2.2, suggest that the longer the unit in focus, the more likely it is for the [fall]

target to occur in the second syllable or even later. In this case, the first syllables receive a [high] target.

I propose that while the encoding of focus does not differ in the basic parameters used in all three domains, the specific applications of the parameters vary from one domain to another. Arguments in focus have wider pitch excursions; comments have lower pitch throughout and are more likely to realise the [fall] underlying target within the boundaries of the first syllables (at least for monosyllabic coverbs), leading to longer durations; and sentence focus constructions are noteworthy for realizing the [fall] target on the second syllable and having less salient pitch excursions and durations on the prominent syllables and throughout the IU. To sum up the characteristics of sentence focus constructions, the overall contour in the IUs is of an initial prominence followed by a fairly sharp decline in pitch, until the next constituent, which receives a prominence, albeit reduced in scale, the subsequent syllables continuing to decline until the end of the IU. It is suggested that the steep decline of the overall contour serves to enhance the monolithic structure of the IUs.

The pattern of underlying targets [high] and [fall] described for Jaminjung can be compared to that of a ‘delayed high’ described by Bishop and Fletcher (2005) for Bininj Gun wok. That is, based primarily on the observation of the pitch tracks, a delayed high may well resemble the pattern of a [high] on the first syllable and a [fall] on the second as described for Jaminjung. Targets in AM and ToBI are not anchored to syllables, which is why the problem of ‘alignment’ is so often raised in analysis using these tools.

Other features include a wider excursion size on the prominent syllable, which is most notable in focused arguments. Another feature is longer duration. The longer durations of the prominent syllables of focused elements in Jaminjung, particularly in arguments again, are also found in other Australian languages such as Warlpiri. Butcher and Harington (2003:324) report that “in phrase-initial position and in a focused context, the initial rhyme of the focused word is lengthened relative to the same word in an unfocused and phrase-medial context”. Interestingly, the authors have also found some evidence that the consonant in the rhyme could be

hyperarticulated, and not the vowel as in English. This would serve the purpose of enhancing the greatest number of contrasts in their own phonemic system, as Australian languages tend to have a wide consonantal and a more limited vocalic inventory. Whether the same phenomenon occurs in Jaminjung remains to be tested.

		Focus domain		
		Comment	Sentence	Argument
Prominence	Location	First syllable, initial word in comment	IU-initial syllable, also secondary prominences in first syllables of other main constituents	First syllable in first constituent in comment, but not exclusively
	Morphosyntax	Coverb, other constituents also possible	NP (nominal) subject/agent, other constituent also possible	NP usually first position in comment, but not exclusively
Correlates	Mean pitch	lowest	higher	higher
	Pitch excursion	Middle (still higher than all others except those at the final boundary)	Less (still higher than all others except those at the final boundary)	Most (higher than all others except those at the final boundary)
	Pitch target	[fall] [mid] mostly on monosyllabic and bisyllabic coverbs [high] [fall] [mid] in trisyllabic words	[high] [fall]	[high] [fall] or [fall] [mid]
	Duration	Yes, prominent syllable longer than all others in	No	Yes, all syllables in NP argument in focus longer

Table 4-13 Summary of correlates in the three focus domains tested in this section.

4.5 Topics

The analysis presented in this section is centered on the notion of *sentence topics* or *aboutness topics*⁴⁶ and not on what has been termed discourse topics (§4.1.1.1). The principal criteria used to identify the subtypes of topics analysed in this section are presented here. Cross-linguistically, sentence topics tend to be fronted, and correspond to an

⁴⁶ It is important to remember that in Jaminjung, sentence topics can be omitted altogether.

expression in the sentence. The term *topic expression* (Lambrecht 1994:129) is sometimes used to distinguish between the notion of what constitutes the topic and its linguistic substantiation in the sentence. Although topic referents are usually associated with presupposed or old information, this can not be a defining factor, as new topics - when the speaker introduces some new participant or entity - are an important subgroups of topics. A further distinction is useful in the context of this analysis. Chafe (1976:50) defines topics which are not integrated into the predicate-argument structure of the clause as: “[a topic that] sets a spatial, temporal, or individual framework within which the main predication holds (the frame within which the sentence holds)”. These topics I will refer to as *framesetting topics* (see Maslova and Bernini 2006 for an overview). The most typical examples of frame-setting elements are spatio-temporal elements; they are often placed sentence-initially. The spatial or temporal locations denoted by these topics are often already part of the shared background of the discourse participants, or can be easily inferred. Finally, there is never a direct predication relation between the framesetting topic and the clause following it (Dipper et al. 2007).

4.5.1 Dataset

The distinctions introduced above are necessary as the dataset comprises several utterances spoken during storytelling sessions based on pictorial stimuli. An example is shown below (75), taken from a retelling of ‘the Frog Story’. The first element *ngiya=biyang*, which can be translated as ‘here now’, is the framesetting topic, the speaker having just turned the page. The sentence or aboutness topic (marked [ab] in the example) is *munurr-ni*, ‘the bees’.

(75)

<i>ngiya=biyang</i> [fs]	<i>munurr-ni</i> [ab]	<i>yurl</i>	<i>ganuny-birri-m</i>
PROX=now	bee-ERG	pursue	3sg:3du-bite-PRS

‘Here now the bees are chasing the two’

[DP:ES96_07_01]

This analysis will make use of the *given/new* distinction (§4.1.4). It refers to the status of the information within the discourse and does not apply exclusively to topics. While speaking, the speaker continually assesses whether the information has already been introduced into the discourse and is known to the hearer, thus *given*, or whether it is information introduced for the first time and thus *new*. The topics analysed in this dataset are categorized according to this distinction. Topics are *given*, when their referent is mentioned in the preceding utterances and is easily accessible. Given topics usually occur at the left edge of the IU, but they are also found in post-verbal positions, either integrated in the IU or as a separate element. These are considered the least marked subtypes of topics. Given topics in Jaminjung may host the clitic =*marlang*, glossed as ‘GIVEN’, these will be tested separately.

Topics are called *new* when their referent has no antecedent in the discourse, and is not accessible through some relational reasoning. They are usually found at the left periphery of the utterance, either integrated in the IU or as a separate element (left detached). The interpretation of such elements as topics, rather than as (presentational) sentence focus constructions, can be rather problematic in Jaminjung. Lambrecht (1994:184) proposes that the two constructions can be formally distinguished in that “presentational NPs may not normally be pronouns and that detached NPs may not normally be indefinite”. The distinction on formal grounds is difficult in Jaminjung as there is no overt marking for definiteness. Another consideration regarding new topics is the very notion of what is considered to be ‘new’. In the spontaneous materials uttered by people who live very closely together, from which our dataset is drawn, it can be difficult at times to assess whether something is really judged as ‘new’ by speakers and hearers. Because of this lack of clear criteria for categorization, the analysis will only make a general set of observations about *new* topics and all interpretations will have to be treated with caution. Finally, prosodic correlates of the position of the topics in the utterance, either at the left or right edge of the clause, for both the given or new subtypes, will be tested. At either edge, they may be separated by a pause or integrated in the IU.

The left periphery in Jaminjung tends to be associated with constituents that are discourse-new entities, in which case they are either followed by a pause or integrated in the single IU with the predicate. The left periphery is not solely associated with discourse new topics; given topics most often occur as integrated topics in this position.

The right edge is associated with given topics that are either integrated in the IU or separated by a pause. NPs at the right edge of the IU with a co-referent pro-form inside the IU are referred to in the literature as *right dislocations* or *antitopics* (Lambrecht 2001:1043 and 1994:118). They tend to be well-established topics and, in English at least, the right-dislocated elements do not receive an accent. Functionally, these NPs serve to reactivate referents present in the discourse situation but not mentioned for some time, or to mark the referent of the right-peripheral NP as the subject for the following discourse segment (reactivation).

To sum up, the sub-categories of topics tested in this section are:

- Given left integrated
- Given left with clitic =*marlang*
- Given left without clitic =*marlang*
- Given right integrated
- Given right dislocated
- Given left with clitic =*marlang*
- Given left without clitic =*marlang*
- New left integrated
- New left detached

Examples are shown below. For given topics, example (76) in Figure 4-30 is from a retelling of the Frog Story, about a boy and his dog looking for a frog. The prosodic sentence contains several IUs, and is quite typical of Jaminjung constructions. At this point in the story, the boy has been picked up by a deer, while the dog is left to follow. The topic here is the dog, *wirib*; it is given and integrated in the IU with the comment *gangga* ‘goes’. Its status as a topic is confirmed in the following IU, a relative clause

introduced by the subordinating clitic =*ma*, giving more information about this referent.

(76)

wirib ga-ngga dang-dang=mang <mijelb>
 dog 3sg-go.PRS RDP-tail.up=SUBORD himself

...'the dog goes, with his tail up,'
 [DP:ES96_07_01]

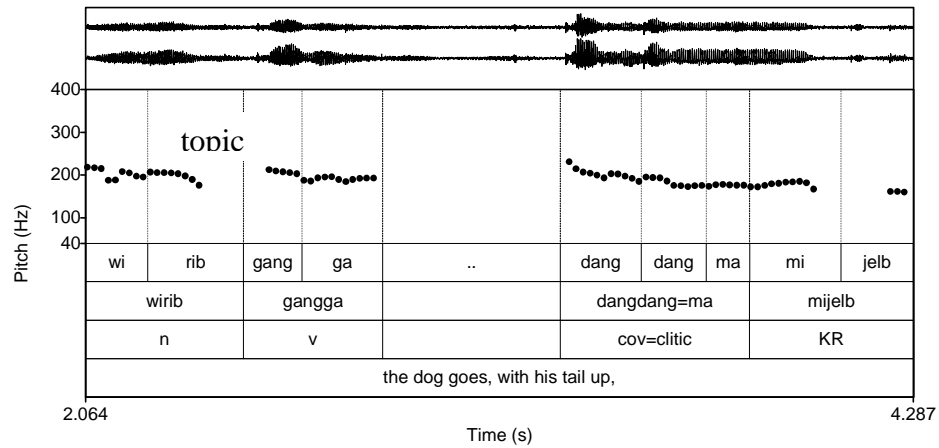


Figure 4-30 The topic in this sentence is uttered in the first IU, *wirib*. It has been referred to many times in the course of the storytelling, and is thus a given topic.

Example (77) in Figure 4-31 is also a given topic integrated at the left of the IU, from the same Frog Story retelling. The topic in this case is the owl, which has been mentioned before. The clitic =*marlang*, signaling the given status of the referent, is attached to the topic.

(77)

<that> ngayin=marlang bul gani-ma bunyag
that meat/animal=GIVEN emerge 3sg:3sg-hit.PST 3DU.OBL

'That animal came out to the two.'

[CP:ES96_18_02]

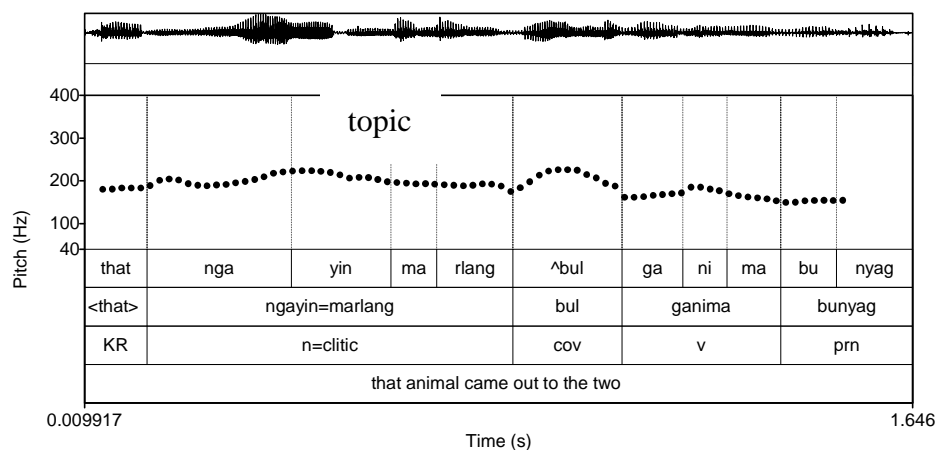


Figure 4-31 A given topic at the left of the IU, this time occurring with the clitic =marlang.

In example (78), the speaker reiterates what another speaker has just said, adding the topic under discussion, the dog *wirib*, at the right edge of the IU.

(78)

yinthu=biyang yugung gan-unggu-m wirib-di
PROX now run 3sg:3sg-do-PRS dog-ERG/INSTR

'Here he is running, the dog.'

[DP:ES96_07_01]

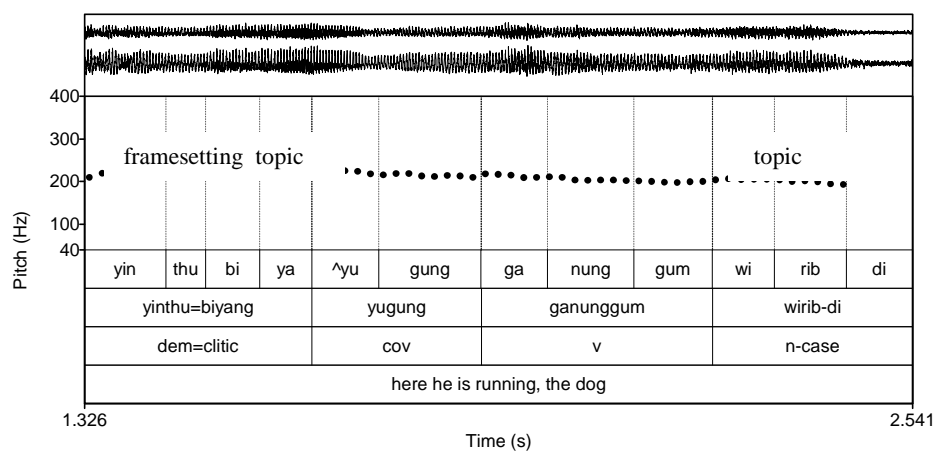


Figure 4-32 A given topic occurring at the right edge of the IU, *wirib-di*.

In example (79) again from a Frog Story, the speaker has just turned the page and thus uses the framesetting expression *yinju=biya* 'here now'.

The sentence topic, *jarlig 'n wirib* 'the child and the dog', is found at the right edge of the utterance, separated by a pause. It is given, as the child and the dog are the two main protagonists in the story, thus part of the shared knowledge of the speaker and hearer, but may not have been mentioned in the last utterance.

(79)

yinju=biyang mugurn buny-agba jarlig'n wirib
 PROX=now lie 3du-be.PST child.and.dog

'Here now, the two are sleeping, the child and the dog

[CP:ES96_18_02]

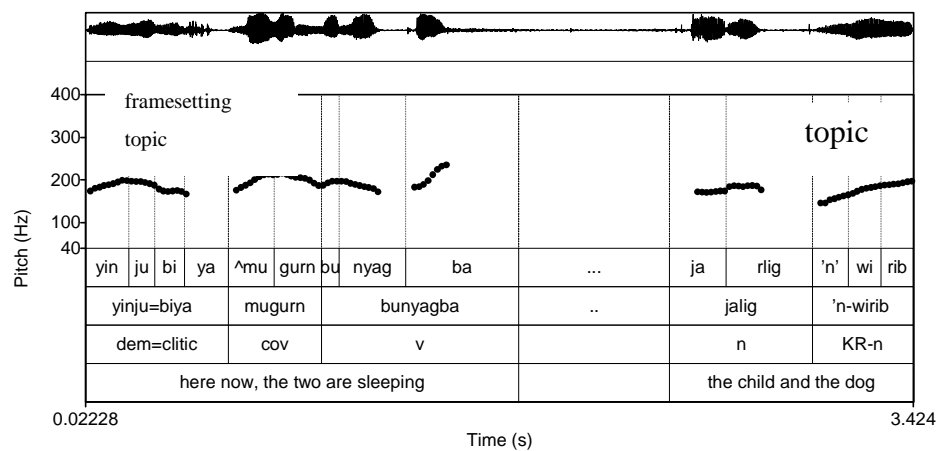


Figure 4-33 The framesetting topic, *yinju=biya* is at the left edge of the sentence, and the given topic *jarlig n-wirib* is found at the right edge, after a pause.

Lastly, in example (80), the speaker is talking about bush food. He has been talking about mussels, *naribu*, and he ends his comments by reiterating the topic, adding the clitic =*marlang*, so that *naribu=marlang* occurs in right-dislocated position, after a pause.

(80) (also in (95))

gurrany yanjungarna jarlig... naribu=marlang
 NEG IRR:2sg:3sg-give child mussel=GIVEN

'You should not give them to the children, those mussels.'

[BH:CS07_72_01]

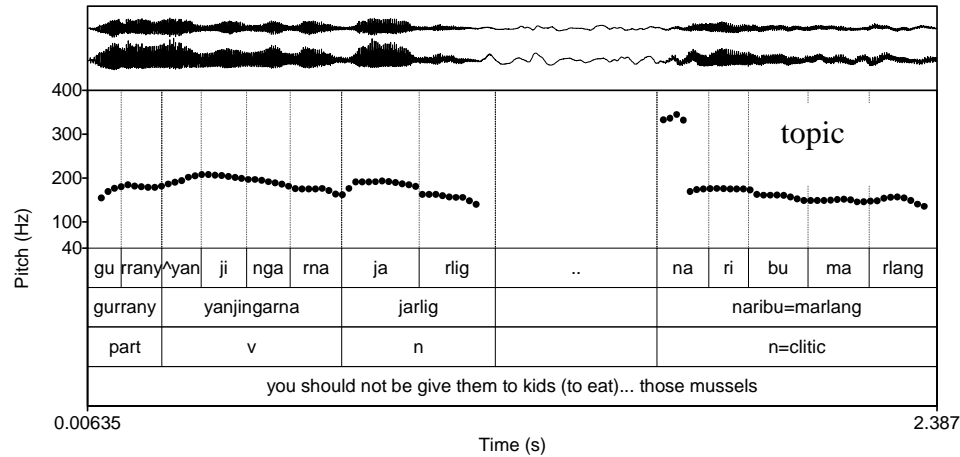


Figure 4-34 A right dislocated detached topic with the clitic =marlang

Example (81), Figure 4-35 shows a new topic, taken from a story about the speaker being bitten by a centipede. At this stage, she tells how hot sand was applied to the bite to relieve the pain. The new topic, *buyud=biyang*, 'sand', forms an IU with the comment and is followed by another IU, an afterthought expressed with the NP *buj-mawu buyud* 'sand from the bush'.

(81) (also in (36))

buyud=biyang <jabul >-ni burr-angu=rrgu=rndi
 sand=now shovel-ERG/INSTR 3pl:3sg-get/handle.PST=1sg.OBL=FOCUS

<buj>-mawu buyud
 bush-HABITAT sand

'They got sand for me with a shovel, the bush kind of sand.'

[IP:ES97_03_02]

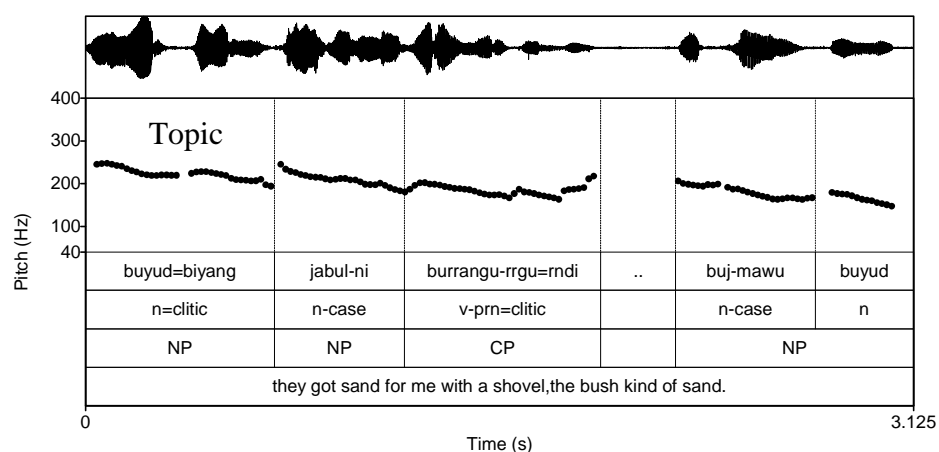


Figure 4-35 A new topic, *buyud=biyang*, at the left of the prosodic sentence.

Example (82) is from an elicitation session with a sequence of cartoons representing short stories about a man walking with his dog. The topic *warladbari* 'old man' is new, as this is the very first cartoon in the sequence. It is separated by a pause from the main predication.

(82)

warladbari wirib wuju nuwina gan-anja-m warlŋginy
 old.man dog small 3SG:POSS 3sg:3sg-bring-PRS walking

'The old man, he is walking his small dog.'

[JM:CS07_083_07]

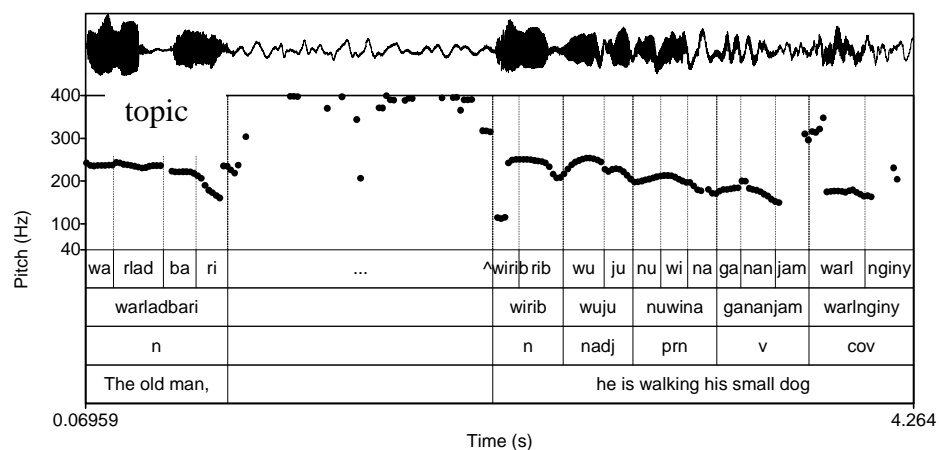


Figure 4-36 A new topic, *warladbari*, separated by a pause from the rest of the prosodic sentence.

4.5.2 Hypotheses

The subtypes of topics discussed and illustrated in the previous sections will provide the materials to test the encoding of topics in Jaminjung. The main hypothesis is that topics that are part of the shared knowledge, or *given*, are less marked prosodically than the *new* or *contrastive* topics which are functionally more salient, following Givón's basic principle of iconicity, that "the more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it" (Givón 1983:18). It is further hypothesised that topics form a phrase of their own, and that the position and integration/dislocation of the topic in the IU influences its encoding. It is further suggested that the presence of discourse-related clitics i.e. =*marlang* impacts on the encoding.

4.5.3 Methodology

This section continues with a methodology similar to that used in the previous section on focus domains. The IUs in the dataset are selected and organized into subcategories; they are segmented into syllables and labeled according to their position. Their mean pitch, excursion size, final velocity and duration are measured. The values of the measurements provide evidence to test the hypotheses.

The topics in constructions formed of a given-topic and comment are examined first. Their correlates are analysed in some detail, as they will serve as points of comparison in the subsequent analysis of the other subtypes. The first test will distinguish between the topics that take the clitic =*marlang* and those that do not. The results of the measurements are compared with those of the syllables in the adjoining comments.

A comparison will then be shown between all the subtypes of given topics, according to whether they are integrated in the IU but in post-verbal position, or right dislocated, and with or without the clitic =*marlang*.

Subsequently, new topics are examined. Finally, a comparison of the encodings of all the subtypes of topics is effected.

4.5.4 Given topics

4.5.4.1 *Given topics at the left periphery*

Integrated given topics at the left of the comment are treated first, continuing with the assumption, based on Lambrecht (1994), that a construction with a given topic and a comment focus is the least marked. The dataset for the IU-initial given topics consists of 58 IUs, out of which 17 have the clitic =*marlang*, glossed as ‘given’. Importantly, the syllables in these topics are not perceived as prominent.

The measures of final velocities are used as indicators of the underlying pitch targets. The differences in the topics with and without the clitic =*marlang* are not significant. The differences in the values according to the position of the syllable in the IU, however, are significant ($F(18, 548) = 2.807, p=.000$). The values are shown in Figure 4-37 and Table 4-15 below. Syllable positions in topics are labeled ‘topic-first, t2, t3’, etc. Those of comments are labeled ‘comment-first, c2, c3’, etc. The standard deviations can be fairly large, as some of the tokens in the dataset displayed uncharacteristic patterns; they were not removed, however, as there were no obvious reason to do so (such as ambient noise, other speakers, etc). The averages are still considered to be representative of the data, but a wider study with a larger, and more controlled, dataset would be needed to determine this with more certainty.

The average velocity for the first syllables of the topics is 7.04st/s, while the second and subsequent syllables in the topics have negative values between -5 and -10st/s. The values decrease from the first to the second syllables, a pattern that is the reverse of that found in comparable syllables in comments, whose values increase (§4.2.2.3.3).

The pattern suggested for given topics is of a [high] target on the first syllable of the given topic, and [mid] targets on the other syllables.

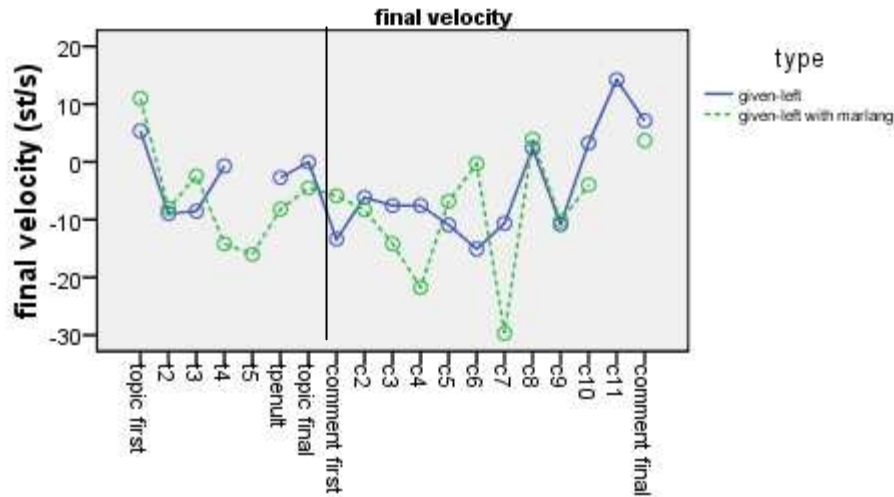


Figure 4-37 Average final velocity in the syllables of given topic/comment constructions.

syll_position	Mean final velocity (st/s)			Std. Deviation			N		
	given-left	given-left with marlang	Total	given-left	given-left with marlang	Total	given-left	given-left with marlang	Total
topic first	5.36	10.98	7.04	16.91	18.66	17.47	40	17	57
t2	-8.93	-8.10	-8.56	23.16	41.12	31.96	21	17	38
t3	-8.54	-2.44	-5.75	18.44	23.16	20.50	13	11	24
t4	-0.71	-14.20	-9.70	0.56	11.29	11.18	2	4	6
t5		-16.00	-16.00		.	.		1	1
antepenult	-2.71	-8.18	-4.61	20.40	13.39	18.32	32	17	49
topic final	-0.04	-4.53	-1.40	24.14	16.23	21.99	37	16	53
comment first	-13.38	-5.87	-11.14	23.34	29.58	25.32	40	17	57
c2	-6.14	-8.30	-6.78	18.18	25.44	20.37	38	16	54
c3	-7.58	-14.17	-9.55	12.32	16.13	13.75	35	15	50
c4	-7.54	-21.76	-12.52	12.60	16.48	15.47	26	14	40
c5	-10.90	-6.82	-9.65	16.74	18.48	17.03	18	8	26
c6	-15.07	-0.29	-10.72	24.11	9.80	21.72	12	5	17
c7	-10.63	-29.67	-15.39	18.55	28.70	21.01	6	2	8
c8	2.38	3.91	2.89	9.69	16.08	10.43	4	2	6
c9	-10.86	-10.20	-10.64	0.66	.	0.60	2	1	3
c10	3.26	-3.98	-0.36	.	.	5.12	1	1	2
c11	14.24		14.24	.		.	1		1
comment final	7.15	3.69	6.16	26.19	25.41	25.78	40	16	56

Table 4-14 The average final velocity values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.

The mean F0 of the syllables in the topics and the comments are then measured. The difference between the subgroups (with or without the clitic) is significant ($F(1, 553) = 16.518, p = .000$), as is the difference between the

syllables according to their positions in the IU ($F(18, 557) = 5.718$, $p=.000$). The values are shown in Figure 4-38 and Table 4-15 below.

The most striking observation is that topics with the clitic =*marlang* have lower F0 values than the topics without. Interestingly, this lowering of F0 values seems to spread to the syllables in the comments, so that comments after a topic with =*marlang* have lower mean pitch than comments that follow a topic without a clitic.

In all tokens, the overall pattern is that of a gentle decrease in values over all the syllables of the topic and their comments, with a reset on the first syllable of the comment. However, in the topics followed by the clitic =*marlang*, there is a slight rise in the mean pitch from the penultimate to the final syllable in the topic, suggesting a slight continuation rise.

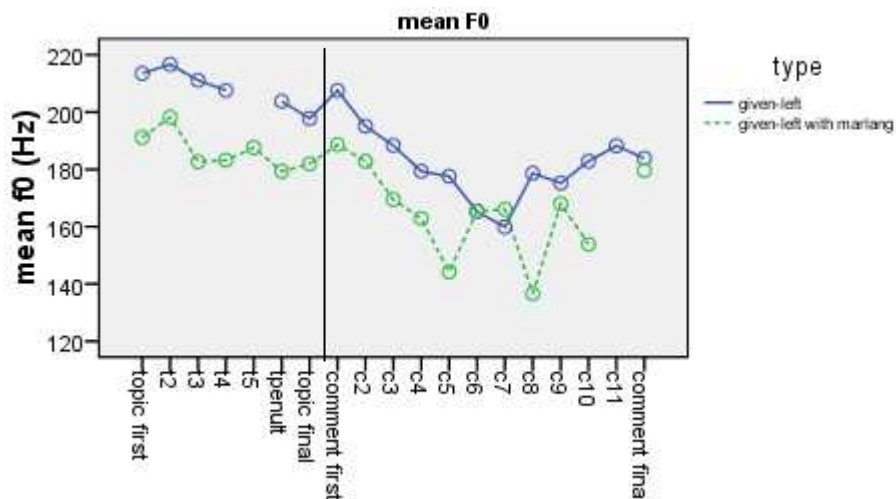


Figure 4-38 Graph showing the average pitch values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =*marlang* (dotted line), and those that are not (full line).

Syllable position	Mean F0 (Hz)		St dev		N		
	given-left	given-left with marlang	given-left	given-left with marlang	given-left	given-left with marlang	Total
topic first	213.51	191.15	23.95	11.64	41	17	58
t2	216.67	198.20	30.96	9.63	21	17	38
t3	211.13	182.71	36.96	25.06	13	12	25
t4	207.55	183.22	6.03	7.32	2	4	6
t5		187.64		.		1	1
topic penult	203.75	179.33	29.25	11.80	32	17	49
topic final	197.79	181.92	25.90	29.40	38	17	55
comment first	207.56	188.62	28.21	8.92	41	17	58
c2	195.09	182.85	22.20	15.11	37	17	54
c3	188.39	169.50	25.48	31.46	36	15	51
c4	179.34	162.83	25.79	17.22	27	14	41
c5	177.61	144.23	18.98	34.79	18	11	29
c6	165.26	165.42	33.72	17.29	12	4	16
c7	159.91	166.08	42.68	25.22	6	2	8
c8	178.67	136.59	5.59	59.33	4	2	6
c9	175.22	168.02	7.49	.	2	1	3
c10	182.88	153.81	.	.	1	1	2
c11	188.23		.		1		1
comment final	183.91	179.62	44.75	65.39	39	13	52

Table 4-15 Average pitch values for IUS with given topics and comments, distinguishing topics that host the clitic =*marlang*, and those that do not.

The next correlate measured is the excursion size, as an indicator of pitch range expansion. The statistical tests show no significant difference between the subgroups with and without the clitic =*marlang*. The differences for the factor syllable position, however, are significant ($F(18, 561) = 2.060, p = .006$). The average values are illustrated in Figure 4-39 and Table 4-16 below. The overall pattern is of declining values in the syllables of the topics, except in the second syllables which have slightly wider excursions than initial syllables. The average values of the excursions on the first syllables of the topics are just under those of the comments where focus is most likely to be realized (§4.2.2.4).

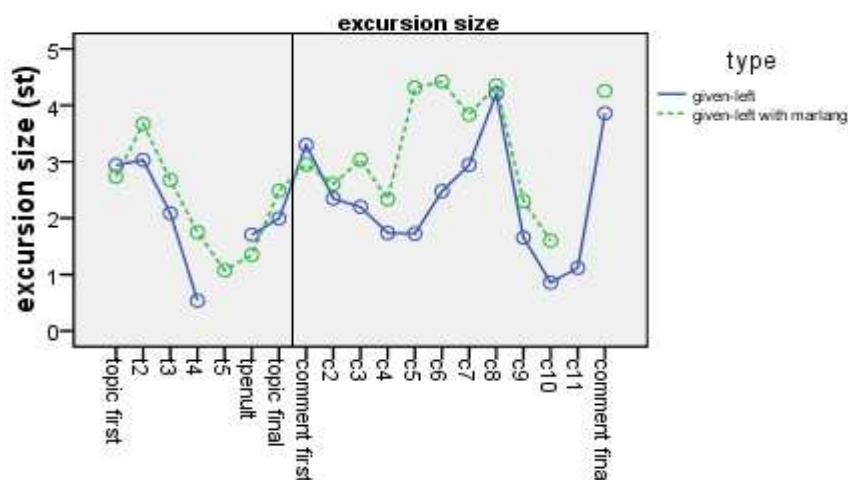


Figure 4-39 Graph showing the average pitch excursions in the syllables of given topic/comment constructions.

syll_position	Mean excursion size (st)			Std. Deviation			N		
	given-left	given-left marlang	Total	given-left	given-left marlang	Total	given-left	given-left marlang	Total
topic first	2.94	2.74	2.88	2.76	2.13	2.58	41	17	58
t2	3.03	3.67	3.32	1.69	1.26	1.53	21	17	38
t3	2.08	2.67	2.37	1.36	2.21	1.80	13	12	25
t4	0.54	1.75	1.35	0.16	0.98	0.99	2	4	6
t5		1.07	1.07		.	.		1	1
topic penult	1.71	1.35	1.58	1.07	1.06	1.07	32	17	49
topic final	1.99	2.49	2.15	1.85	4.24	2.78	38	17	55
comment first	3.30	2.95	3.20	2.64	1.72	2.39	41	17	58
c2	2.35	2.61	2.43	1.52	2.00	1.67	38	17	55
c3	2.20	3.03	2.44	2.37	3.71	2.82	36	15	51
c4	1.74	2.34	1.94	2.32	1.95	2.19	27	14	41
c5	1.72	4.32	2.71	1.46	4.79	3.33	18	11	29
c6	2.48	4.42	3.05	2.17	5.82	3.54	12	5	17
c7	2.94	3.84	3.17	1.96	1.99	1.87	6	2	8
c8	4.21	4.35	4.26	5.70	1.62	4.47	4	2	6
c9	1.65	2.30	1.87	1.34	.	1.02	2	1	3
c10	0.86	1.60	1.23	.	.	0.53	1	1	2
c11	1.12		1.12	.		.	1		1
comment final	3.86	4.26	3.98	3.82	5.07	4.18	41	17	58

Table 4-16 The average pitch excursion values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.

Finally, the duration measurements have a similar pattern with no significant differences between the subgroups (with/without =marlang), but significant differences between the syllable positions ($F(18, 525) = 11.013$, $p = .000$). The first syllables of topics are slightly shorter than the second

syllables, with respective values of 157.19ms and 177.61ms. The subsequent syllables decrease in duration until the final syllable, which is longer, probably as a constituent boundary encoding. The first syllables of the comments remain the longest in the IU, as discussed in §4.2.2.5.

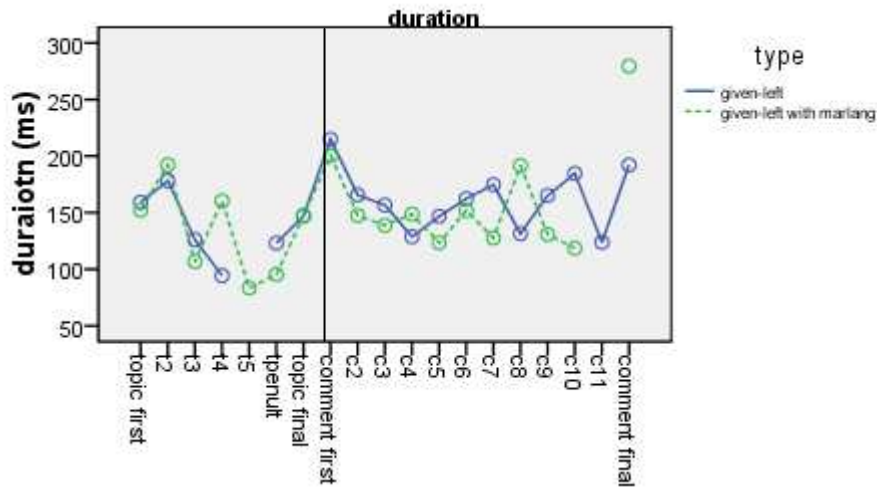


Figure 4-40 Average duration in the syllables of given topic/comment constructions.

syll_position	Mean duration (ms)			Std. Deviation			N		
	given-left	given-left with marlang	Total	given-left	given-left with marlang	Total	given-left	given-left with marlang	Total
topic first	159.23	152.26	157.19	57.54	30.27	50.90	41	17	58
t2	178.22	192.19	184.47	70.64	48.40	61.32	21	17	38
t3	126.06	106.77	116.80	40.34	36.52	39.01	13	12	25
t4	94.38	160.39	138.39	38.32	59.42	59.79	2	4	6
t5		83.38	83.38		.	.		1	1
topic penult	123.05	95.27	113.41	43.79	25.44	40.41	32	17	49
topic final	146.96	147.40	147.10	47.72	60.36	51.38	38	17	55
comment first	214.78	200.26	210.52	78.07	48.20	70.52	41	17	58
c2	165.60	147.50	160.00	45.55	42.45	45.02	38	17	55
c3	156.67	138.85	151.43	57.30	51.32	55.70	36	15	51
c4	128.76	148.71	135.57	43.54	67.27	52.86	27	14	41
c5	146.75	123.28	137.85	54.03	23.65	45.90	18	11	29
c6	162.33	152.05	159.31	60.99	48.56	56.31	12	5	17
c7	174.70	127.60	162.92	58.71	1.69	54.20	6	2	8
c8	131.40	191.55	151.45	18.08	81.47	49.88	4	2	6
c9	165.17	130.68	153.67	46.67	.	38.54	2	1	3
c10	184.62	118.48	151.55	.	.	46.77	1	1	2
c11	123.78		123.78	.		.	1		1
comment final	192.12	279.49	217.73	57.81	118.07	88.70	41	17	58

Table 4-17 The average duration values for IUs with given topic and comments, distinguishing topics that are followed by the clitic =marlang, and those that are not.

As a summary, the final velocity measurements suggest that the first syllables have a [high] pitch target and the subsequent syllables in the topic a [mid] target. Mean F0 is interesting in that it is the only correlate that distinguishes between given topics with and without the clitic =*marlang*: an IU with a topic followed by =*marlang* has lower pitch values over all its syllables. This would indicate that the clitic intensifies the *givenness* property associated with the topic.

For the correlates of pitch extension, final velocity and duration, the presence or absence of the clitic does not produce significant differences. The second syllables have slightly wider pitch excursions and durations than the first syllables: the values decrease gently over the subsequent syllables. The parameters of pitch range and duration appear not to be used to encode given topics.

4.5.4.2 Comparison: subtypes given topics

In this section, a comparison between the measurements of all the subtypes of given topics is presented. The comparison comprises five subtypes of given topics: given topics integrated at the left periphery of the IU; given topics integrated at the left periphery with the clitic =*marlang*; integrated given topics at the right periphery⁴⁷; right dislocated given topics; and right dislocated given topics with the clitic =*marlang*. The calculations only involve the syllables of the topics. Some of the average values of syllables in third, fourth or fifth positions should be taken with much circumspection, as the number of tokens of longer topics is often much reduced. For ease of comparison, the results of the measurements are shown in graphic form for each of the correlates below.

Final velocity measurements are presented first (Figure 4-41). The differences in the final velocity values are not significant between the subtypes of given topics, but they are significant regarding their relative position ($F(7, 368) = 2.796, p = .008$). In all subtypes of given topics, the first

⁴⁷ There are no integrated given topics at the right periphery with the clitic =*marlang*, in this dataset, which does not mean that they may not exist in natural language, however if the clitic is to be interpreted as signaling lesser accessibility of the referent, it may make sense for it not to be found in this position.

syllables have higher final velocities, followed by a sharp decrease in the second syllables to rise slightly over the third syllables. The values for the first syllables for the given topics at the left, and the right dislocated topics with the clitic =*marlang* are positive and strongly suggest a [high] underlying target. Those of the integrated and right dislocated topics without the clitic are just under zero and thus less evidently indicate a [high] target. In both cases, the [high] target is still auditorily perceptible, albeit in a reduced manner.

At the right edge of the topics, the velocity rises slightly from the penultimate to the last syllables, which all end very near the zero mark, except for the post-comment integrated topics.

The contour in topics is thus of a [high] on the initial syllable followed by a succession of [mid] targets. The differences between the first, second and third syllables possibly help in highlighting the change in targets.

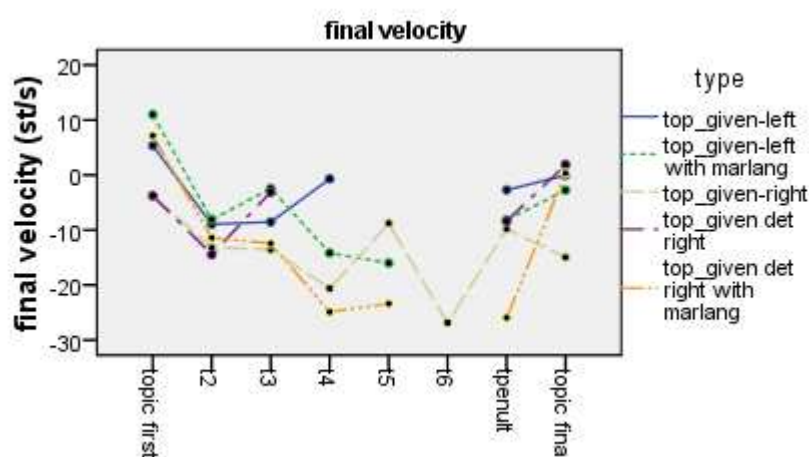


Figure 4-41 A comparison of final velocity average values in all subtypes of given topics.

The pattern of targets suggested by the final velocity measurements is corroborated in the measures of the correlate of mean F0. They are found to vary significantly between the subtypes ($F(4, 372) = 17.476, p = .000$), and according to subtypes syllable position ($F(7, 372) = 2.672, p = .011$).

The position of the topic at the left or right of the IU, integrated or detached, quite understandably affects the values of the mean F0 over the syllables (Figure 4-42).

The mean pitch tends to rise slightly from the first to the second syllables in given topics, except in the right-dislocated topics with the clitic =*marlang*, which have mean F0 values that decrease sharply from the first until the last syllable, indicating that the presence of the clitic negates the need of a prosodic marking.

In both cases of =*marlang*, either at the left or in right dislocated topics, the overall mean F0 is lower than when the topic is not followed by a clitic. It is as if there is a scale in the *givenness* of the topic, which manifests itself in the mean pitch of the entire topic, the clitic =*marlang* having the effect of accentuating the givenness of the referent. Alternatively, the variation in the overall mean pitch could be interpreted in terms of accessibility, lower pitch corresponding to the more accessible stage.

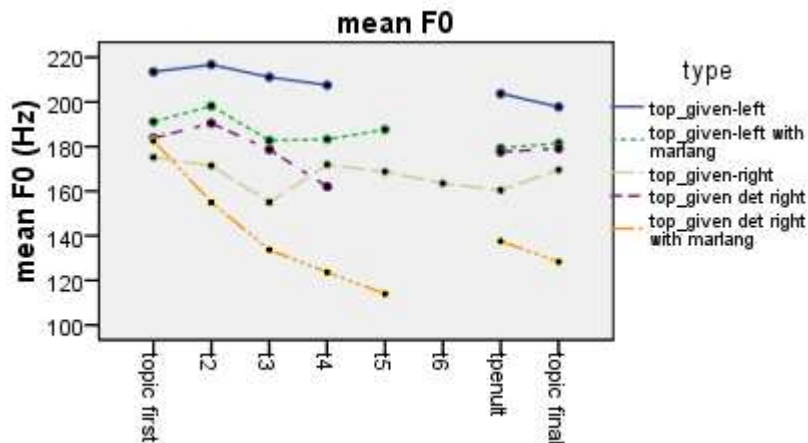


Figure 4-42 A comparison of mean F0 values between all subtypes of given topics.

Excursion size does not vary significantly for subtypes nor for syllable position (Figure 4-43). The values for the excursion size in the first syllable are close to 3st for most subtypes, except for the right dislocated topics with =*marlang*, which have a slightly wider excursion, an effect of the steep fall in pitch described above. Excursion sizes then tend to decrease until the final syllables where they may get wider again, most noticeably in the syllables of IU-final topics. The excursion sizes at the right edge are interpreted as boundary-marking signals. At the left edge, however, at around 3st, they are slightly less than that of the prominent first syllables of comments (around 4.5st).

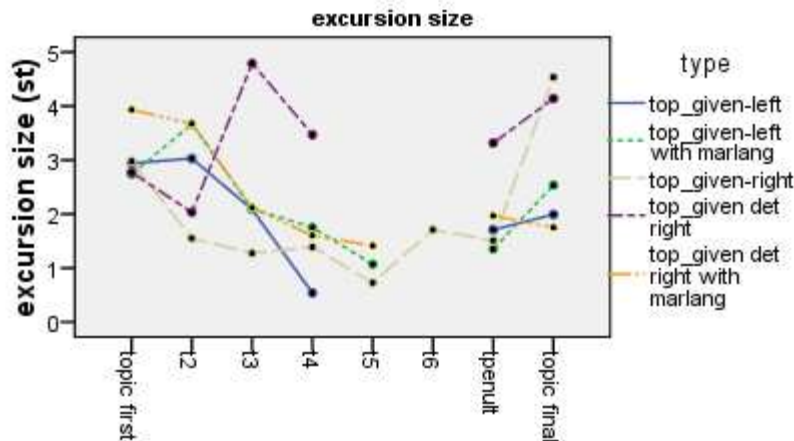


Figure 4-43 A comparison of the excursion sizes in the syllables in all subtypes of given topics.

Finally, the differences in duration values are not significant between the types of given topics, syllable position, however, is significant (($F(7, 374) = 5.125, p = .000$) (Figure 4-44). The duration of the first syllables is very similar, between 160ms and 175ms. It is shorter than that of the second syllables, with a decrease over the subsequent syllables until later in the topic, where the final syllables are longer. Obviously, the longer durations of the final syllables are associated with final lengthening, and it is not surprising that the topics at the left periphery, as they are not IU-final, should have shorter durations. Here, interesting patterns are noted in right dislocated topics. The patterns in the durational values in both subtypes are very similar until their final syllables, the length of the preceding pauses also differ. Final syllables in topics with *=marlang* are substantially longer with 231.97ms to 160.38ms. The pause preceding the dislocated topic is markedly longer in the topics with *=marlang*, 818.41ms compared to 532.63ms in topics without *=marlang*. Furthermore, the final syllable in the preceding comments is longer when the following dislocated topic has *=marlang* than in those that do not. This is taken to be an indication that when a speaker uses the clitic *=marlang*, he/she may have intended to end the utterance at the end of the comment and then decided to reiterate the topic, having judged that it was not accessible in the hearer's mind any longer.

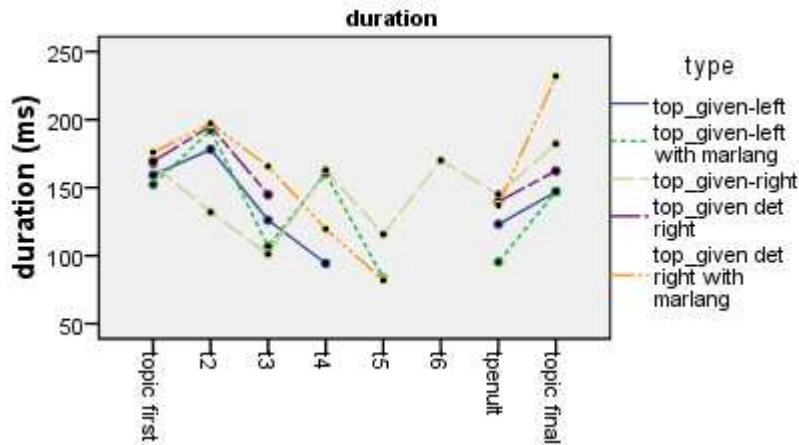


Figure 4-44 A comparison of the duration values in all subtypes of given topics.

4.5.4.3 Summary: given topics

The comparison for the given topics shows no great variation in the encoding between the subtypes tested, and the observed variations can easily be correlated with their relative positions in the IU or prosodic sentence. Given topics have gentle falling contours, marked by a [high] underlying target on the initial syllables, followed by a succession of [mid] targets. Topics followed by the clitic =*marlang* have significantly lower mean pitch than those without the clitic, both when they are at the left and at the right edge of the comment. At the right-edge, the final syllable of the comment, and the pause that precedes topics with =*marlang* are longer than those that do not. Right-dislocated topics with =*marlang* seem to have a much steeper falling contour than all other subtypes of given topics.

Altogether, apart from the [high] initial target, given topics are not marked by distinctive prosodic means. The next section will explore the encoding of new topics.

4.5.5 New topics

New topics are either integrated in the IU with the comment, or they occur as a detached element to the left of the comment, separated by a pause. The two subtypes are analyzed in this section. The dataset comprises 48 utterances, 30 are new integrated topics and 18 are detached from the

comment. Both can be interpreted as introducing new referents to the discourse.

The differences between the two subtypes (detached, integrated) are found not to be significant, except for those of the mean F0 ($F(1, 464) = 6.659, p = .010$). The factor syllable position is significant for all the correlates (mean F0 ($F(19, 450) = 3.103, p = .000$; excursion size ($F(19, 467) = 4.264, p = .000$), final velocity ($F(19, 456) = 2.510, p = .000$); duration ($F(19, 467) = 2.348, p = .001$)). The values are shown in graph form for each of the correlates. The adjoining comments are also included, for ease of comparison.

The final velocities of the first syllables are .38st/s for the detached new topics, and 11.94st/s for the integrated topics, a fair difference (Figure 4-45). Those of the second syllables are much less divergent, with respective values of -12.03st/s and -17.47st/s. Both have [high] underlying targets on the first syllables, which are more salient in the integrated topics. At the right edge of the topics, there is a steeper rise in the velocity values of the left dislocated topics, which go from -15.65st/s to -3.34st/s, while the integrated topics have almost no change, from -9.89st/s to -9.67st/s. This is interpreted as an indication that left dislocated topics form an IU of their own and not an encoding of topics as such. Furthermore, the change in the velocity in the final syllables of the detached topics suggest that instead of receiving a final [low] target, they are possibly marked with a [rise] target.

The pattern of underlying targets suggested is [high], followed by [mid] targets.

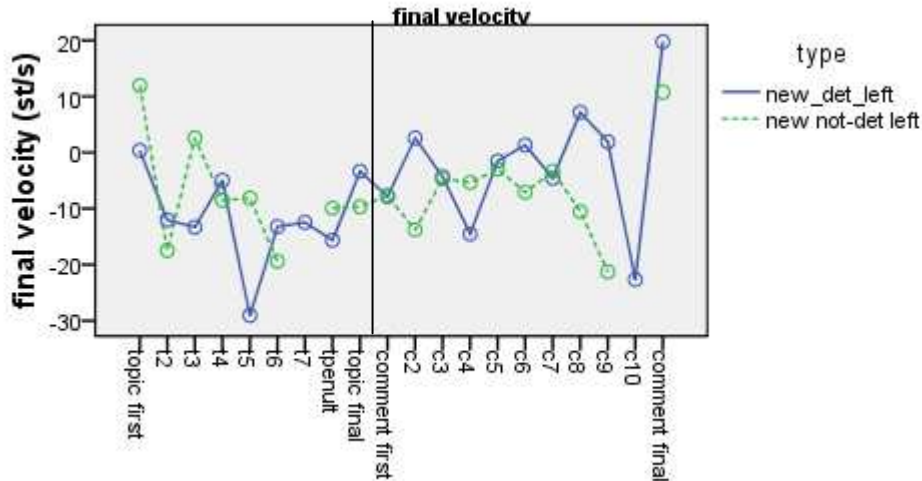


Figure 4-45 Comparative values of final velocity in new topics, whether they are integrated in a clause or separated by a pause.

As for mean F0 (Figure 4-46), the difference between the two types of new topics is significant, but this is understandable, given their positions within the prosodic sentence. Both have diminishing values over the syllables of the topics. There are not many topics with more than three syllables, so the averages over the intervening syllables may not be over-reliable; however, at the right edge of the topics, where all tokens are taken into account, the values appear to rise. This can be treated as a continuation rise: the speaker having just introduced a new referent, he/she signals that there is more information to come.

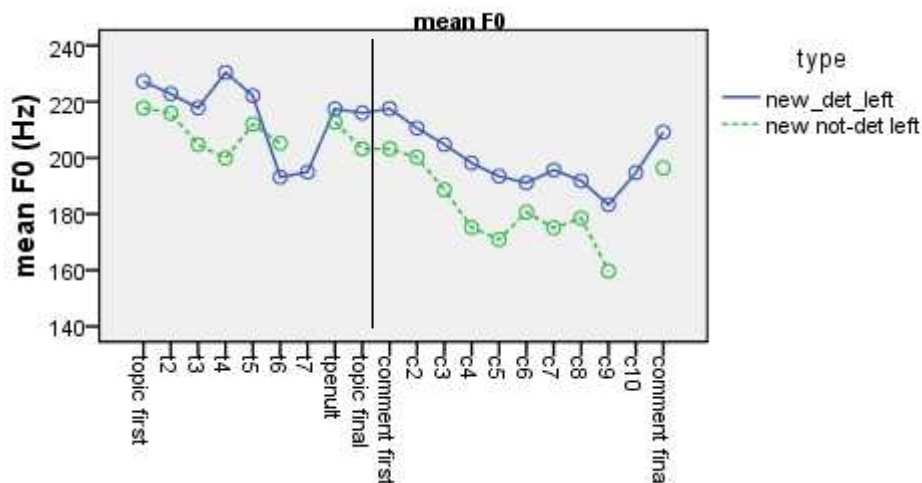


Figure 4-46 Comparative values of mean F0 in new topics, whether they are integrated in a clause or separated by a pause.

The excursion size measurements show a remarkable similarity between the two subgroups, until the final syllables where the detached topics have much wider excursions (Figure 4-47). At the left edge, the detached topics also have higher values, with 4.60st to 3.77st in the integrated topics. Their second syllables have much less wide excursions, and the values then decline until the end of the topic. The more extreme values of the detached topics at both edges confirm their status as distinct IUs.

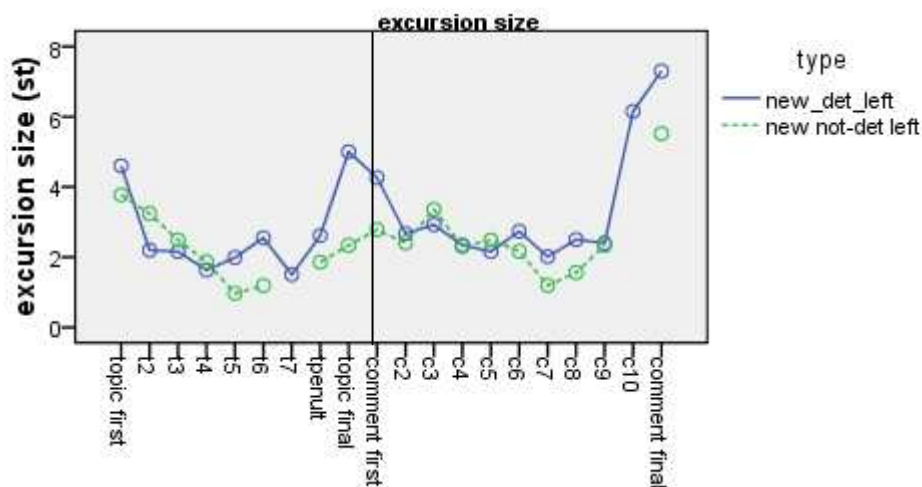


Figure 4-47 Comparative values of excursion size in new topics, whether they are integrated in a clause or separated by a pause in detached and non-detached new topics.

The patterns in duration values are not dissimilar, with the syllables at the boundaries longer than those intervening (Figure 4-48). Detached topics have noticeably longer initial and final syllables. They are followed by fairly long pauses, at 886.10ms. These values are also congruent with the IU status of detached new topics.

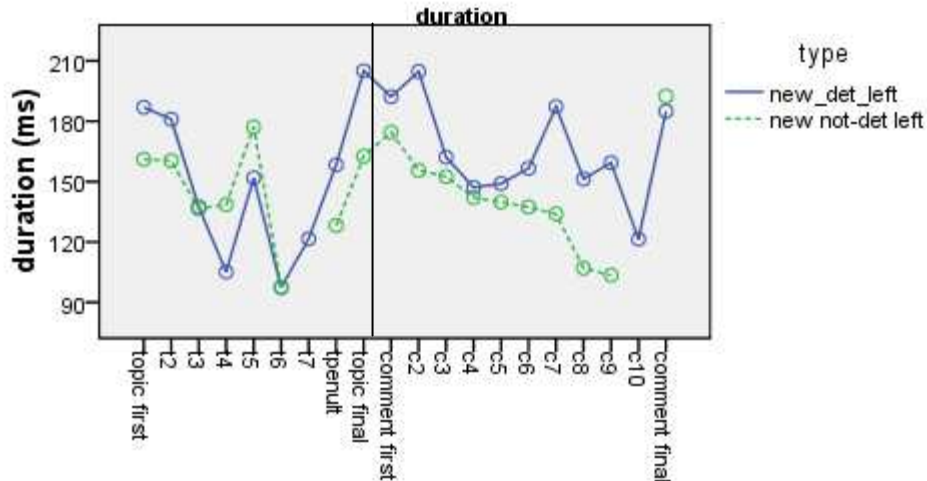


Figure 4-48 Comparative values of duration in new topics, whether they are integrated in a clause or separated by a pause in detached and non-detached new topics.

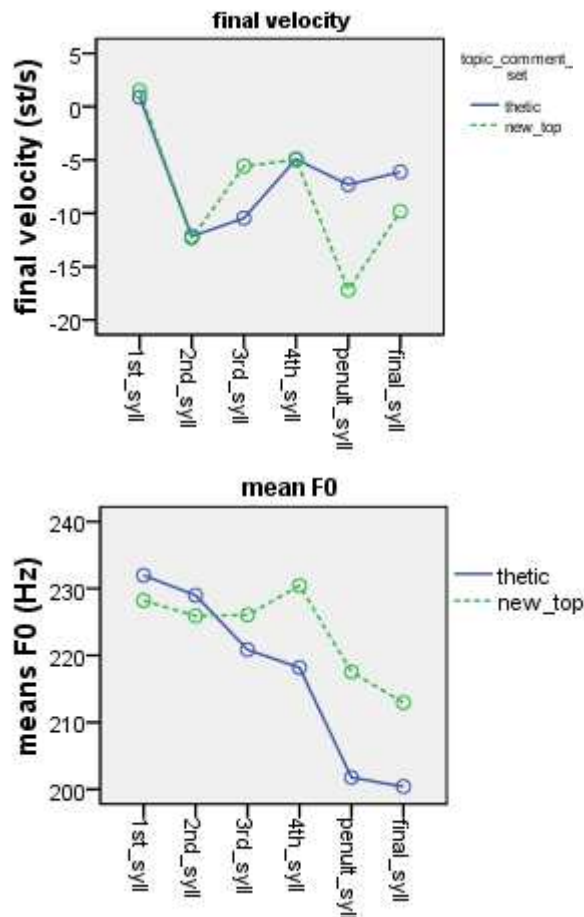
4.5.5.1 *Thetic and/or new topic*

The difficulty in distinguishing a new topic from a single-word thetic sentence has been mentioned (§4.5.1). This test aims at comparing the values of the correlates for both constructions, with the aim of ascertaining if and how they differ. The dataset comprises the tokens of sentence focus domain and new detached topics. Only the 6 syllables are taken into account in the calculations, the first four, and the last two.

The comparison of the values between thetic sentences and new topics show how alike they are, as the graphs in Figure 4-49 illustrate. The new topics differ in their mean pitch which does not fall as sharply as in thetic sentences. Their excursion size is more pronounced at both edges, but most notably in the initial syllable. The final velocity values, however, are very remarkably similar in the first and second syllables. These syllables also have longer durations.

These measurements show that new topics cannot be equated altogether with thetic sentences, even if the values of final velocity at the left edge are arresting in their similarity, suggesting a pattern of a [high] followed by a [fall] target. On the other hand, the excursion sizes on the first syllables of new topics are much wider. Also, at the right edge, the overall values of the correlates indicate that thetic sentences are self-contained

entities, while those of the left dislocated new topics suggest that they are not self-contained, but an IU that is part of a bigger sentence, as is particularly evidenced by the mean pitch, and excursion size. The values of both mean pitch and excursion size in thetic sentences and new detached topics can be interpreted as signaling continuation versus final boundary at their right edge.



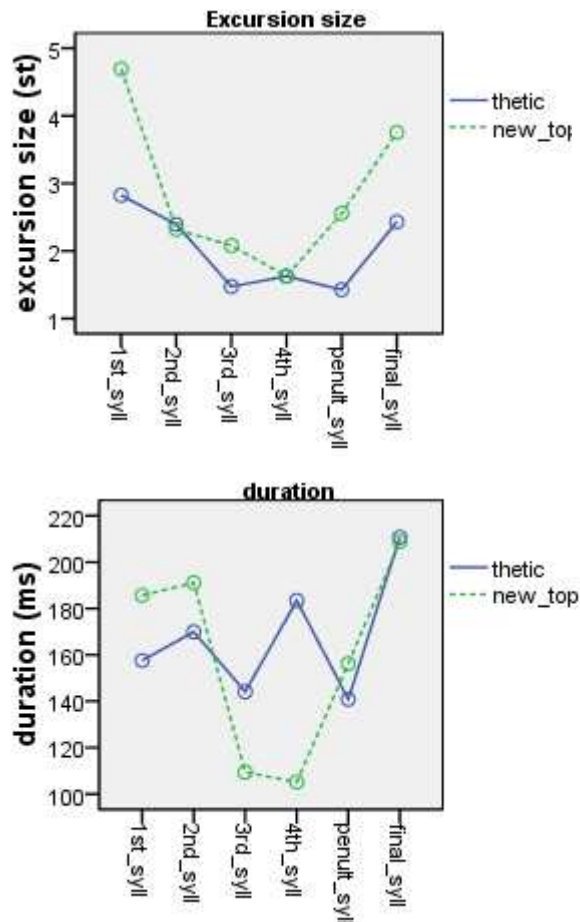


Figure 4-49 A comparison between thetic sentence and new dislocated topics, including correlates of final velocity, mean F0, excursion size, and duration.

To conclude, the measurements for excursion size, final velocity and duration do not vary significantly in detached or IU-integrated new topics. However, a closer inspection of the two average F0's does show variance, which is predictable given the positioning of the topical elements in the prosodic sentence. The detached new topics have higher mean F0 averages than integrated topics, which is in accordance with their sentence-initial positions.

The final velocity measurements suggest a pattern of [high] targets on the initial syllables, followed by [mid] targets.

The larger excursion sizes and durations of the syllables at the left and right edges of the detached topics highlight their status as a separate IU.

4.5.6 Comparison: topics

In this section, I will compare the values of the correlates in all new and given topic subtypes in order to address the question of what constitutes topic encoding, and highlight similarities and differences between the subtypes, to finally propose an explanation for these differences.

The comparison concerns all the subtypes of topics tested so far:

- Given left without clitic =*marlang*
- Given left integrated with clitic =*marlang*
- Given right integrated
- Given right dislocated
- New left integrated
- New left detached

For the purpose of this comparison, only topics with four syllables are retained, as there are enough tokens in each subtype to provide reasonable averages, and it is thus possible to distinguish the patterns at the left and right periphery.

The subtypes vary significantly in the correlates of mean F0 ($F(6, 150) = 10.579, p = .000$). The factor syllable position is significant for all the correlates (mean F0 ($F(3, 150) = 7.834, p = .000$; excursion size ($F(3, 150) = 2.693, p = .049$), final velocity ($F(3, 149) = 4.844, p = .003$); duration ($F(3, 152) = 10.728, p = .000$)).

Firstly, final velocity measurements are compared, to check patterns in underlying targets (Figure 4-50). Overall, at the left edge the patterns are similar, with higher values in the first syllables which decrease in the second. Whether this encodes similar underlying targets is difficult to decide. I suggest that the values well above the zero mark in the first syllable indicate a [high] target, that is, for all subtypes except the new detached topics, and the right dislocated topics. At the right edge, it seems that the final velocity values rise from the penultimate to the final syllables in all but the given right topics, with a steeper rise in right given detached topics and right given detached topics with =*marlang*, reflecting their sentence-final position.

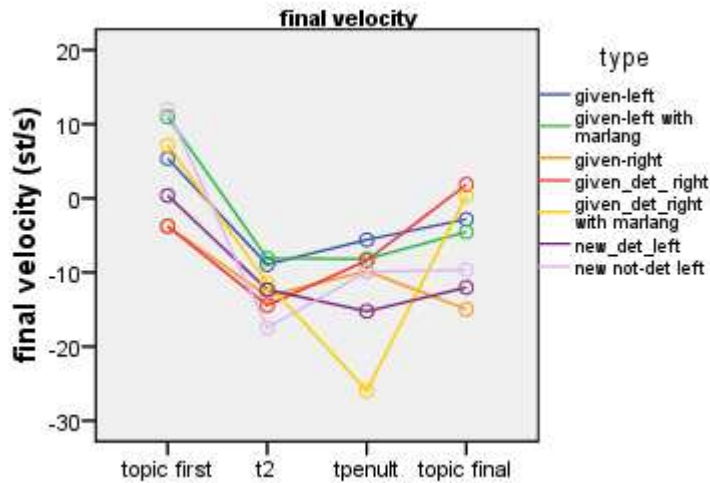


Figure 4-50 A comparison of the final velocity measurements in all subtypes of topics.

There is a significant difference between all the subtypes of topics in the overall mean F0 values, reflecting their position at the left or right edge of the comment or the prosodic sentence (Figure 4-51). The left dislocated new topics have the highest values in their first syllables, reflecting the fact that they form a separate IU altogether. For all subtypes, the overall pattern is of declining values. At the right-edge, the values are lower. Topics with the clitic =*marlang* both at the left and at the right, have lower mean F0 values, particularly in their initial syllables. Rather surprisingly, integrated topics at the right have lower values than right dislocated given topics. At the right edge, the pattern is of a continuing declination from the penultimate to the final syllables. The decline is steeper in both left dislocated new topics, and in right dislocated given topics with and without =*marlang*, highlighting their IU status.

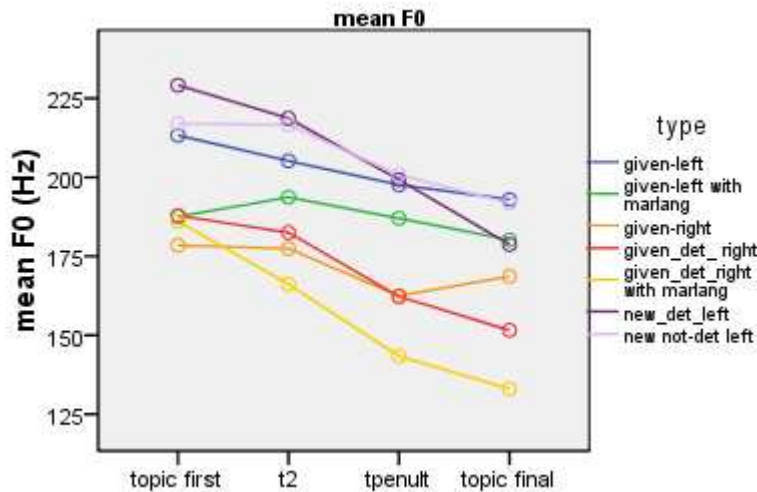


Figure 4-51 A comparison of the mean F0 measurements in all subtypes of topics.

Topics have fairly reduced excursion sizes; in fact, the first syllables of most given topics have values under 3st (Figure 4-52). New dislocated and integrated topics and integrated left given topics have wider excursion sizes. The excursion sizes in subsequent syllables are reduced, until the final syllables, where the topics that are unit-final have wider excursions, probably an effect of final lowering.

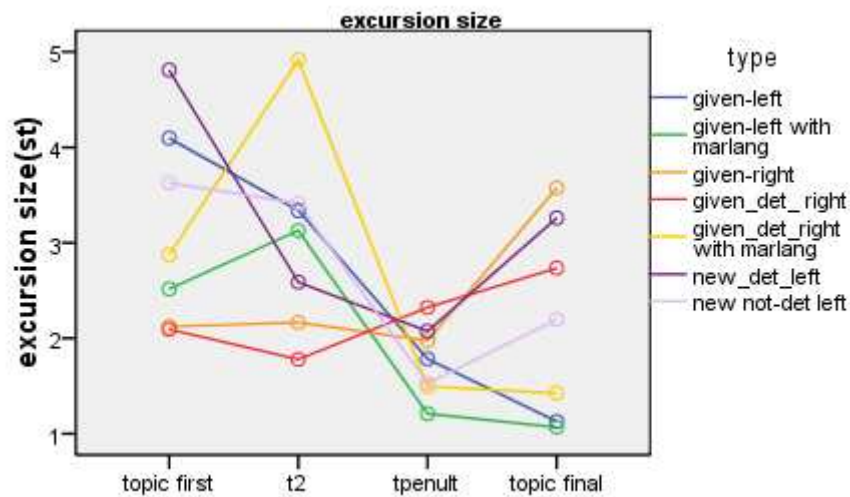


Figure 4-52 A comparison of the excursion size measurements in all subtypes of topics.

The patterns for the values of duration do not suggest a clear distinction between the subtypes. All have longer syllables in initial positions, and have longer final than penultimate syllables, except for the given right and given right detached topics with =*marlang*. Syllables at the

edge of an IU, such as the right dislocated given topics with *=marlang*, the left dislocated new topics, given integrated topics at the right of the comment, are markedly longer than those that are at the edge of a prosodic phrase.

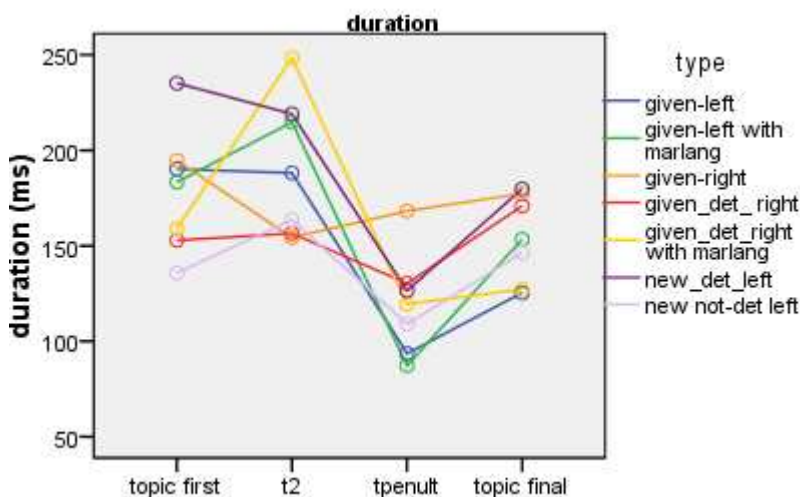


Figure 4-53 A comparison of the duration measurements in all subtypes of topics.

In conclusion, it is important to remember that in Jaminjung, the favoured strategy is to omit continuous referents altogether. Hence, the mere presence of a referent NP is already an indication that the speaker wishes to bring it to the fore. For example, when NPs occur in pre-verbal position, they tend to be either topical, in argument focus, or part of athetic construction.

The encoding of topics (when they occur) can be described as follows. The left given topic is confirmed as an unmarked category, in line with cross-linguistic findings - it has a less steep contour, low excursion sizes, and fairly low durations. Its final velocity values are about midway through the averages of all the others topics types tested.

Given topics followed by the clitic *=marlang*, either at the left or at the right of the comment, have lesser mean F0 values than their equivalent without the clitic, as if the degree of 'givenness' is somewhat increased. This indeed supports our early hypothesis that discourse-related clitics influence prosodic encoding, at least for the clitic *=marlang*.

As to the right dislocated given topics, those with the clitic *=marlang* have a much steeper falling contour, reflected in their excursion

size. The length of the pause is notable here, that preceding topics with =*marlang* is longer (818ms) than when there is no clitic (533ms).

Concerning the new topics, integrated topics are marked by high mean F0 values and wide excursion sizes in their first syllables. As to the left dislocated new topics, I will contend, in light of the results presented so far, that it could be treated as a sentence in focus, given the closeness of the measurements of its correlates to that described for sentence focus (§4.4).

Topics differ mainly from comments in that topics do not have prominences, which are usually associated with the first syllables of the element in focus. Speakers in Jaminjung either integrate a topic into an IU, or choose to isolate it in a separate IU, thus making it more salient. The results of the measurements for the topics can be interpreted in terms of a *givenness* scale (Figure 4-54). Indeed, the findings in this section are in accordance with the basic hypothesis, after Givon's (1983:18) principle of iconicity, that "the more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it". If we posit the *given topic + comment* structure as unmarked for Jaminjung, occupying a mid-point in our dataset, as shown in Figure 4-54 below, then other subtypes can be placed along a 'more given' to 'less given' hierarchy. The 'more given' topics are shown on the left of the axis; in speech, they can be either integrated or detached, and be positioned either at the left or at the right of the comments. The 'least given' topics are shown at the right of the axis. They include new topics which are more likely to occur at the left edge of the IU. The most marked topics are left detached new topics which have a [rise] target at their right edge indicating a continuation. The most prosodically marked topics are the contrastive topics which are examined in the next section.

The findings reported here for the encoding of topics in Jaminjung are thus in line with cross-linguistic generalizations of topic marking in spontaneous speech. Further, as noted by Wichmann (2000:39) (for English), the most salient features of topic shift in spontaneous speech are often the high pitch resets and acceleration in tempo associated with beginnings, "the previous topic does not end, it fades away. The voice become quieter, slows down, and the pitch range narrows. So the transition

from one topic to another in this kind of speech is often a gradual one and it is difficult to identify a precise boundary”. This of course refers to topic shifts, but nevertheless, an important point is made: prosodically, the function of marking topics overlaps with the structural organization of the speech, hence its prosodic encoding may also share the same parameters.

<div style="display: flex; justify-content: space-between; align-items: center;"> + givenness in topics - </div>							
given	given	given	given	given	new	new	contrastive
detached	detached	integrated	integrated	integrated	integrated	detached	both
right	right	right	left	left	left	left	
= <i>marlang</i>			= <i>marlang</i>				

Figure 4-54 The givenness scale in Jaminjung. On the left, the topics subtypes that are ‘more given’ and on the right those that are ‘less given’.

4.6 Contrast

In this section, I will test the encoding of contrast, which, in many theories of information structure, is claimed to be essentially marked by prosodic means. For the purpose of this research, contrasted elements are defined as being ‘in opposition’ with other alternatives in a limited set. Contrast applied to both focused elements, here arguments in focus, and to topics will be examined. In both cases, the contrastive constructions are compared with non-contrastive ones, thus arguments in contrastive constructions to those in non-contrastive ones; and contrastive topics to given topics.

A comparison between the correlates of contrastive arguments and contrastive topics is undertaken. The effect of the clitic =*gun*, glossed as ‘contrastive’, after both focused arguments and topics is evaluated. Finally, all of the contrastive constructions are compared in order to assess whether they make use of the same prosodic encoding.

4.6.1 Contrastive argument focus

4.6.1.1 Dataset

Contrastive argument focus is identified:

- where explicit contrast is expressed such as ‘W, then X, but Z, then Y’.
These are restricted to tokens where only the NPs are contrastive; tokens where the whole predicates are contrasted have been excluded. There are 18 tokens in the dataset.
- in instances of corrective focus or replacing focus; including tokens with explicit corrections (also metalinguistic corrections, e.g. replacing a previous Kriol word, or structures like ‘not X, Y’). The selection is restricted to corrections involving NP arguments only. There are 18 tokens in the dataset.

These two constructions are illustrated in examples (83) and (84) in Figure 4-56 and Figure 4-56. In (83), an explicit contrast is made. The speaker is retelling the events of the previous day, she indicates what they ate, specifying the kind of meat they had as *julag=biji* ‘only chicken’. The use of ‘only’ is typically associated with contrastive focus (König 1991).

In (84), a corrective focus is illustrated. The speaker corrects herself for using a Kriol word, replacing it with *langinybina*.

(83)

yawayi	jurlag=biji	yurru-minda-ny
yes	bird=ONLY	1pl.incl:3sg-eat-PST

‘Yes, we just ate bird (i.e. chicken).’

[IP:ES08_04_01]

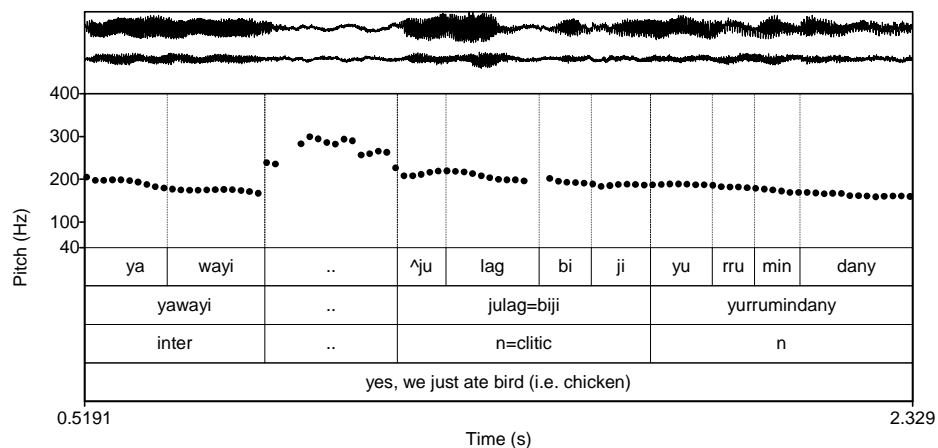


Figure 4-55 A contrastive argument focus is shown, *julag=biji*.

(84)

yinju=biyang liwu buny-ijgany <log-bina> ah langiny-bina
 PROX=now swim 3du-go-PST wood-ALL ah wood-ALL

'Here now, the two swam to the log, ah, to the tree.'

[CP:ES96_18_02]

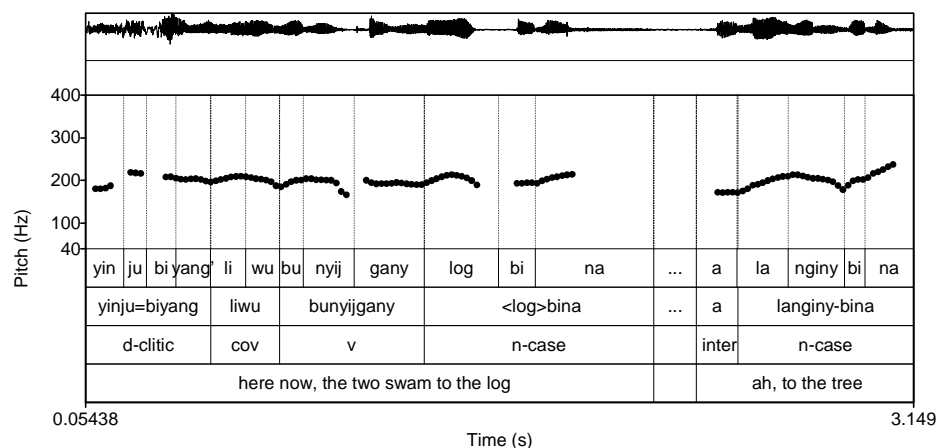


Figure 4-56 A contrastive argument focus is shown, *julag=biji*. The bottom pane shows a corrective focus, the speaker providing a Jaminjung equivalent, *langiny-bina*, to a word first uttered in Kriol, *log-bina*.

4.6.1.2 Comparison between contrastive and non-contrastive focused arguments

This section reports on a test conducted in order to check whether the two subtypes of argument focus, information focus and contrastive focus, receive a different encoding. The arguments in the dataset are labeled as either *non-contrastive*, comprising basic and question-answer focus or *contrastive*, comprising non-corrective and corrective/replacement focus. A summary of the subtypes is presented in Table 4-18.

Type	Selection criteria	Number of tokens	Prominences
Non-contrastive	Repetition (but not a case of contrast) and only one argument is new	19	19/19 first syllable
	Clear cases of question/answer pairs	16	13/16 first syllable
Contrastive	Explicit contrast of the type "W, then X, but Z, then Y", only when NPs are contrastive, not included: whole predicate is contrasted	18	13/15 first syllable
	Explicit correction (also metalinguistic correction, e.g. replacing a previous Kriol word, or structures like "not X, Y"), restricted to correction involving NP arguments	18	16/17 first syllable

Table 4-18 Subtypes of argument focus tested.

The correlates of mean F0, excursion size, final velocity, and duration are used as dependent variables and syllable position and contrastiveness as independent factors. The differences between the subtypes are not significant, but syllable position is significant for mean F0, ($F(5, 230) = 5.011, p = .000$), excursion size ($F(5, 232) = 2.348, p = .042$), and duration ($F(5, 234) = 3.722, p = .003$). The averages are shown in graph form in Figure 4-57 below. The results are discussed as, even if the differences do not reach statistical significance, patterns seem to emerge between the subgroups.

The final velocity measures show no great variation in the first and second syllables, the locus of the perceived prominences, with respective values for the first and second syllables of -8.12st/s , and -16.21st/s in information focus and -2.82st/s and -18.71st/s in contrastive focus. There is variation at the right edge, not in the word-final but in the penultimate syllables. The mean F0 values are interesting in that contrastive arguments have higher mean pitch, not in the syllables at the edges, but in all intervening syllables. The excursion sizes support this, they are not higher in the prominent syllables, but are consistently higher in the remainder of the word. Duration values, however, are very similar across subtypes.

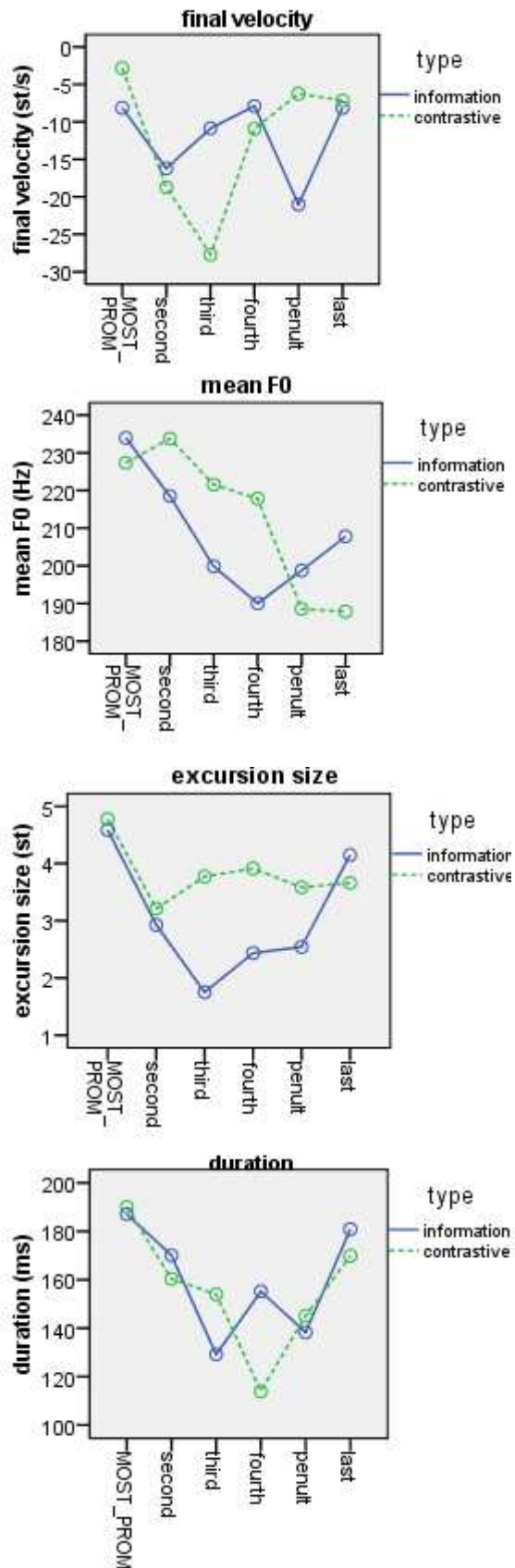


Figure 4-57 The mean values of the measurements of the correlates of final velocity, F0, excursion size, and duration, distinguishing between contrastive and non-contrastive argument focus.

Finally, I would suggest, based on auditory perception, that the encoding of contrastive focus makes use of the correlate of *intensity*, which has not been included in this analysis for methodological reasons. Further research including this parameter would probably yield some interesting results.

One explanation is proposed to account for the observed patterns in contrastive and non-contrastive focus and their lack of statistical significance. It is possible that the encoding of focus is gradient rather than categorical, as proposed by Baumann et al. (2006), in which case, categorical differences would be much harder to discern. A more controlled dataset with a larger number of tokens could possibly clarify this issue. Our findings are in line with those reported for other Australian languages, notably Bishop (2003:17), who describes an extension of pitch range to mark contrastive focus in Bininj Gun Wok and Singer (2006) who describes a similar pattern for Mawng.

4.6.2 Contrastive topics

Topics can also be contrastive. They are defined as topics, that is what the speaker wishes to talk about, but their contrastive quality indicates that there exist alternatives the speaker wants to talk about. The contrastive topics are compared to the given topics-comments constructions.

Example (85) and Figure 4-58 show a contrastive topic, where the speaker recounts where she stayed the day before. She contrasts ‘the house’ *jarriny-gi* where she stayed to ‘outside’, *balarrgu=mang*. Note that the most prosodically salient parts of these IUs are the verbless predicates which bear contrastive focus, *bundurrwari* ‘warm’ and *garrij* ‘cold’.

(85)

jarriny-gi bundurrwari ... balarrgu= mang garrij
hole-LOC warm-QUAL outside=SUBORD cold

'In the house (it was) warm ... outside on the other hand (it was) cold.'
[DB:ES97_01_03]

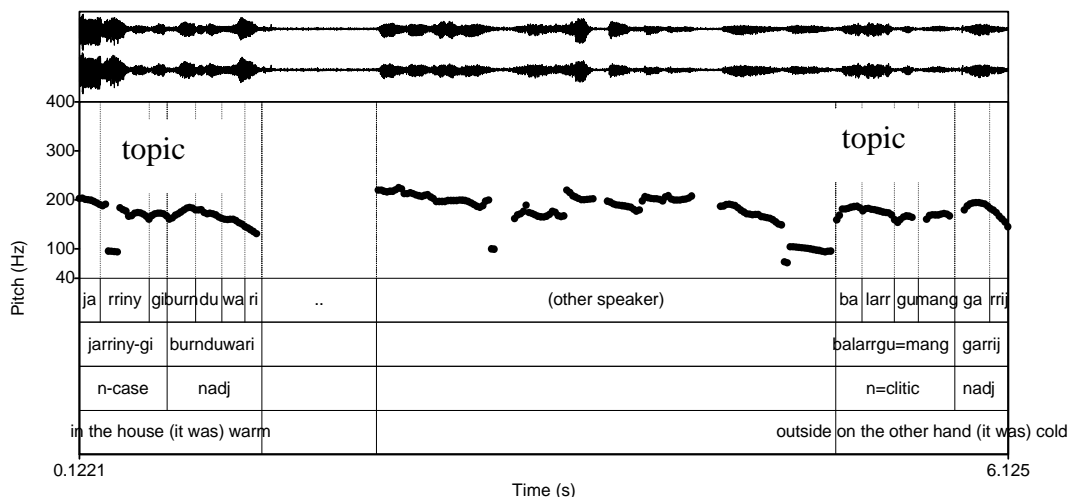


Figure 4-58 The contrastive topics *jarriny-gi* and *balarrgu=mang*.

The dataset is reduced with only 16 tokens. In 14 of the 16 tokens, the topic occurs in initial position in the IU, in the other 2 tokens, it is placed at the right edge of the comment. The differences between the two types of constructions are not significant. When syllable positions with less than 3 tokens are eliminated from the statistical analysis (to avoid under-representation), the factor *syllable position* is found to have a significant effect on the measurements for mean F0 ($F(8, 403) = 5.232, p = .000$), final velocity ($F(8, 397) = 3.468, p = .001$), and duration ($F(8, 403) = 4.840, p = .000$). Figure 4-59 illustrates the results for each correlate.

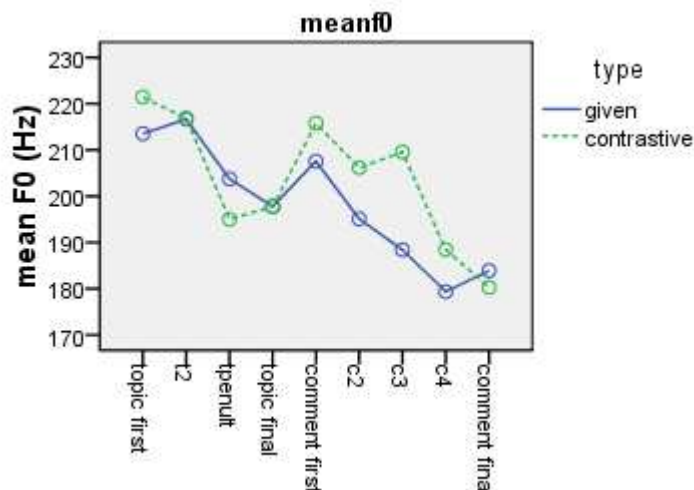
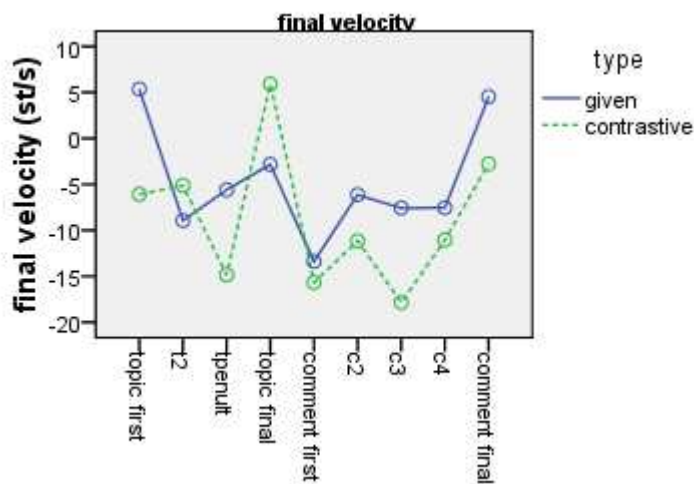
The final velocity values in the first and second syllables of the contrastive topics are -6.09st/s and -5.11st/s, as opposed to 5.36st/s and -8.93st/s in given topics, suggesting a [fall] target realized over the first two syllables. The mean F0 values of the first syllables are also markedly higher in contrastive topics. It is also interesting to note that while the overall pitch on the remaining syllables of both types of topics is very similar, the comments after a contrastive topic are uttered with higher overall pitch.

The excursion sizes in contrastive topics have high values not only in the initial syllables, but throughout the topical NP. This is probably one of the most striking parameters of the encoding of contrastive topics.

The duration measurements also vary: overall, the syllables in contrastive topics tend to be longer. The pattern over the first and second syllables are somewhat reversed with the first syllables in contrastive topics substantially longer than the second and subsequent syllables with values of 177.68ms and 137.23ms, to 159.23ms and 178.22ms in given topics.

These measures suggest that contrastive topics have a [fall] target on their initial syllables, an encoding that, in fact, resembles that of other elements in focus.

contrastive topic
 1st syll 2nd syll 3rd syll
 [fall] [fall] [mid]



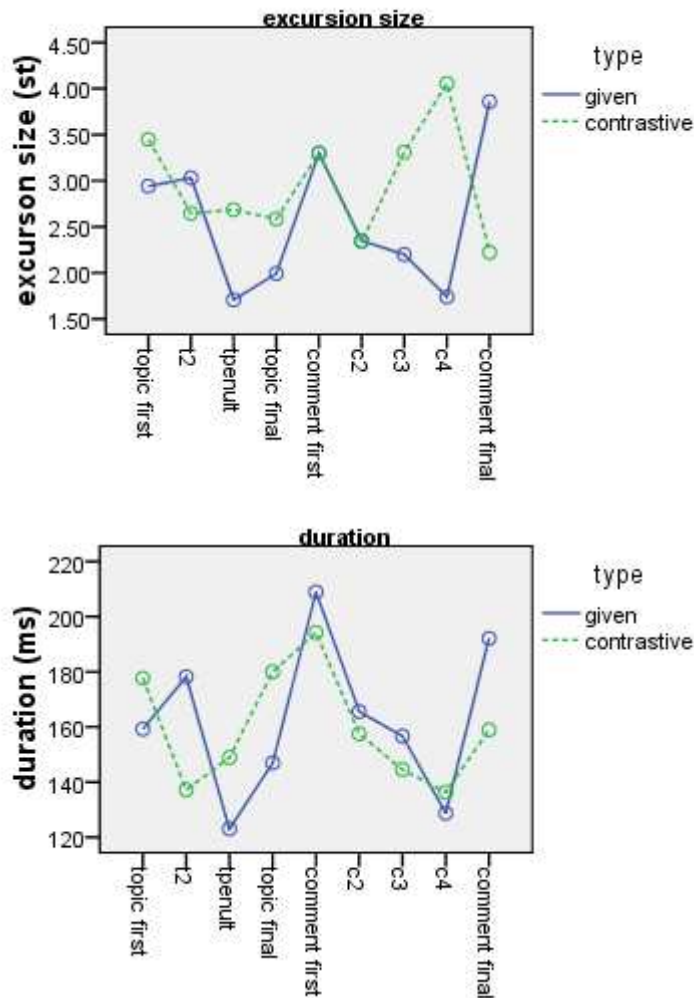


Figure 4-59 The values of final velocity, mean F0, excursion size, and duration of contrastive topics and the following comments, compared to those of the given topic-comment constructions.

To sum up, the contrastive topics display an uncharacteristic pattern amongst topics. Their average pitch values are high, comparable to that of the comments; an expanded pitch excursion is maintained throughout the syllables of the topic. Final velocity measures suggest a [fall] target in the initial syllables, suggesting an overlay of the encoding of focus over this syllable. The duration of the first syllables is markedly longer than that of second syllables, and durations of the topic overall are longer than those of given topics.

4.6.3 Contrastive clitic =*gun*

This section will examine the effect of the clitic =*gun*, a general marker of contrastiveness in Jaminjung, which can occur either after contrastive NPs or verbal constructions. A small dataset is examined here to assess the effect of the clitic on the encoding of contrast. The analysis is limited to NPs, and consists of 16 tokens of contrastive argument focus and 7 of contrastive topics. The values of the correlates of mean F0, excursion size, final velocity and duration are measured in order to assess whether their encoding differs.

The IUs containing the contrastive arguments in this dataset display the same variation in structure and constituent order as arguments without the clitic. The clitic is found at the right-edge of the NP which has a prominent first syllable. Example (86) is an illustration of a prosodic sentence consisting of three IUs. The first one is *wiribni wagurra nguyungnguyung gananggamindi* ‘the dog is sniffing the rock’; the second one *majani nguyung gananggam wirlgagun*, ‘maybe he smells the tracks’; and the third is a tag, *ngih* ‘isn’t it?’. The contrastive argument is found at the end of the second IU, in *wirlga=gun* ‘footprints’; the speaker wishes to emphasize that the dog is sniffing about for something special, here the tracks of the animal it is trying to find, rather than the rock. The prominence of its first syllable is evidently illustrated in the pitch track.

(86)

wirib-ni wagurra nguyu-nguyang gan-angga-m=mindi
dog-LOC rock smell:RDP 3sg:3sg-get/handle-PRS=12DU

majani nguyang gan-angga-m wirlga=gun ngih
maybe smell 3sg:3sg-get/handle-PRS footprint=CONTR TAG

'The dog is sniffing the rock (on you and me), maybe he smells the tracks, doesn't he? '

[IP:ES97_03_01]

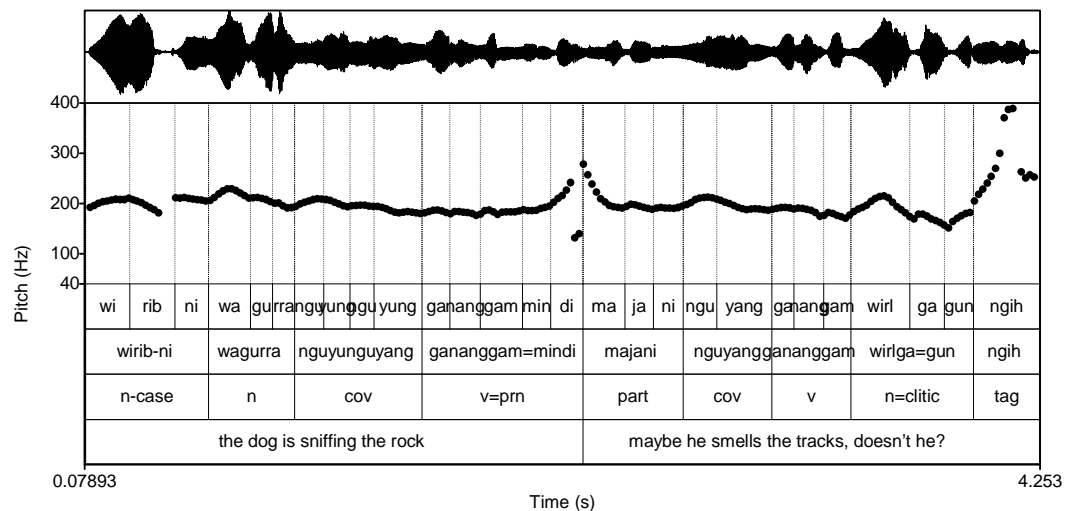


Figure 4-60 A sentence showing a contrastive argument followed by the clitic *wirlga=gun*.

The measures for the contrastive argument focus and contrastive topics which host the clitic =*gun* are represented in graph form in Figure 4-61.

The measures are found not to differ significantly for the correlates of pitch excursion and final velocity, but they are significant for mean F0 ($F(1, 163) = 9.801, p = .002$) and duration ($F(1, 165) = 11.287, p = .001$). This is not unexpected, as both mean F0 and duration clearly reflect the positioning of the element in the IU, the topics being more likely to be IU-initial. The prominent syllables in the contrastive argument focus have a velocity value of -6.59st/s and the contrastive topics of -0.81st/s . The second syllables have respective values of -28.28st/s and -9.80st/s . These patterns suggest that both contrastive arguments and topics receive a [fall] target which is more marked in arguments than in topics. These differ mostly from the focused arguments and topics that do not have the clitic in the values of the mean pitch, as shown in the next test.

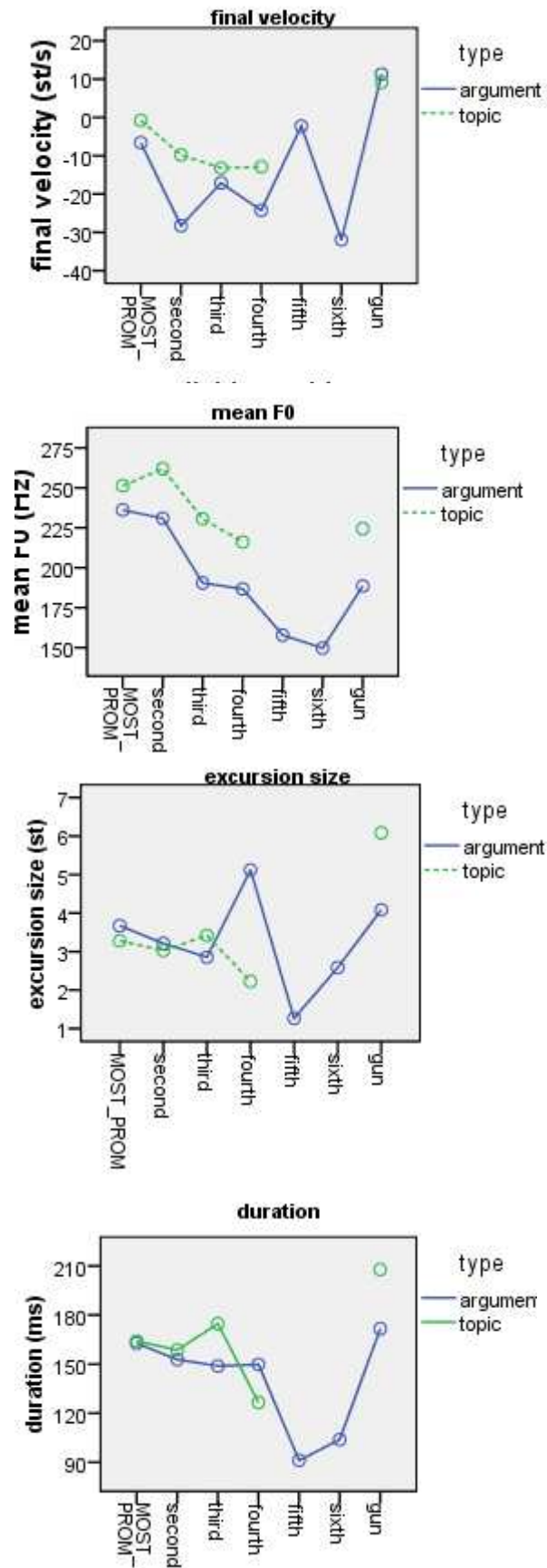


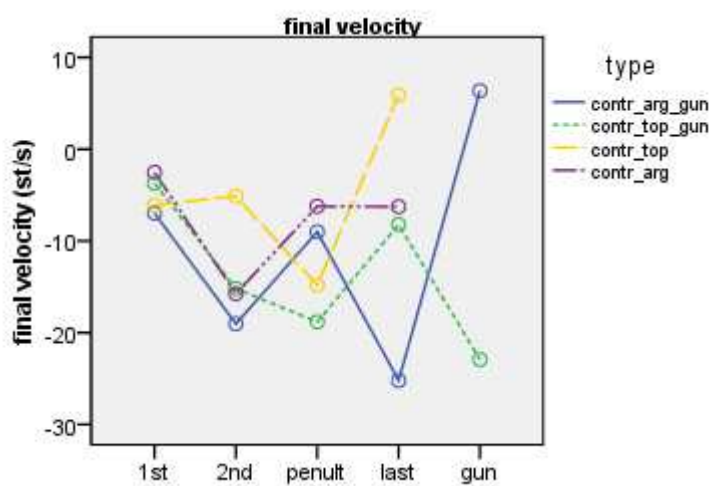
Figure 4-61 Graphs of the values of mean F0, excursion size, final velocity and duration of the contrastive arguments in focus and topics which host the clitic =*gun*.

4.6.4 Comparing contrastive constructions

In this final test, the correlates of all contrastive elements are compared, with the aim of showing whether contrast is encoded differently in focused arguments and in topics, and whether the clitic =*gun*, glossed as ‘contrast’, influences the encoding. The first syllable, in all cases heard as prominent, the second, the penultimate and final syllables are included, as well as the clitic =*gun*, when present.

The differences in mean F0 are significant, for both types ($F(3, 243) = 4.709, p = .003$) and syllable position ($F(4, 243) = 4.484, p = .002$).

Otherwise, all correlates show no significant difference between the subtypes tested. The examination of the results, summarized in graph form in Figure 4-62 below, confirms that the patterns are indeed similar in all subgroups. All have final velocity values that suggest a [fall] on the initial syllables, and all have fairly wide and sustained excursion sizes. The least marked appears to be the contrastive topic with the clitic =*gun*, which has the highest mean pitch throughout, but also the lowest values for excursion, indicating a fairly flat high contour. The most marked is the contrastive focused argument, evidenced by its wider excursions, and longer duration.



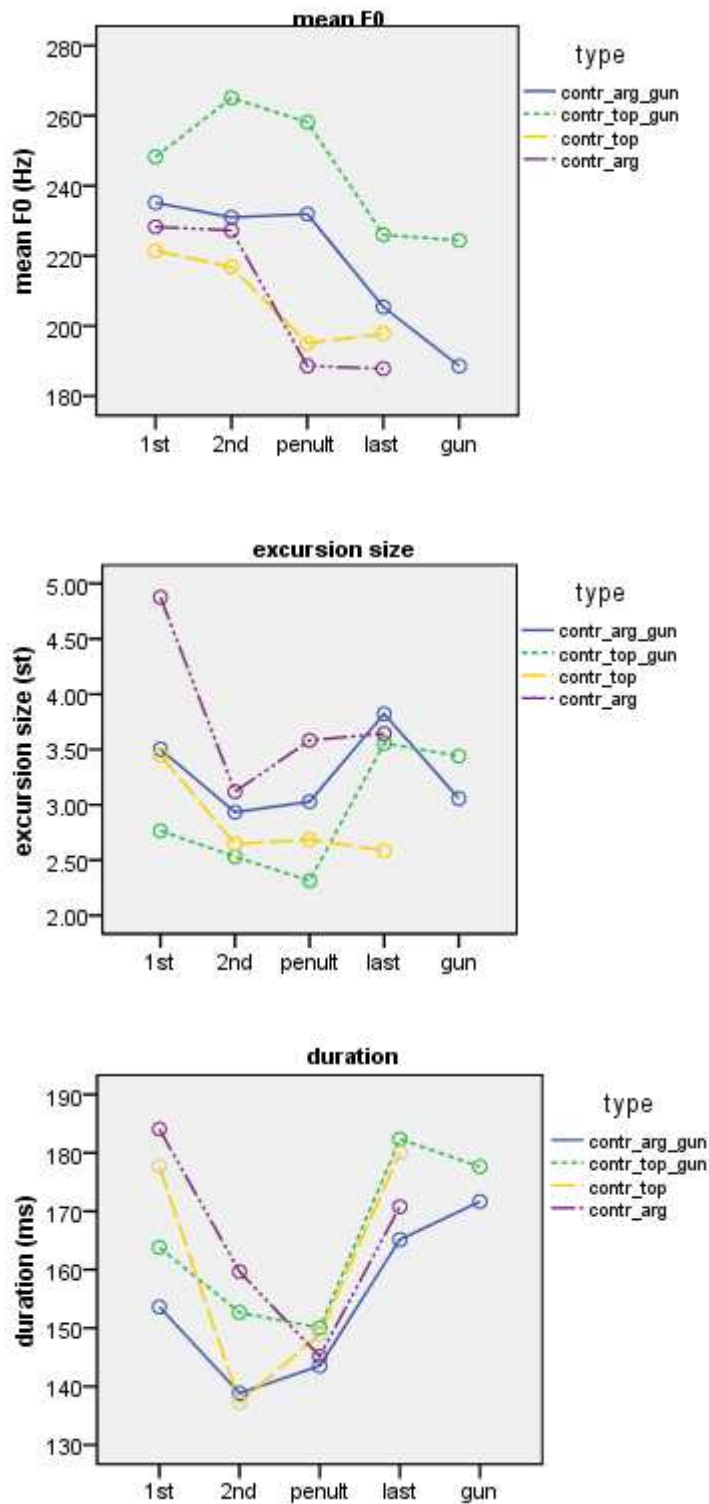


Figure 4-62 A comparison of the contrastive constructions, including the correlates of final velocity, mean F0, excursion size, and duration.

This test confirms that a single encoding scheme can account for the patterns of contrastive elements in this dataset. They are marked by a [fall] target on their first syllables, and by an extended pitch range throughout.

Thus contrast can be seen as a notion that is related to focusing, in that contrasted elements are always focused, as demonstrated by the common [fall] underlying target and longer duration of the prominent syllables, but which is also a feature of its own, as indicated by its being applied to both arguments in focus and topics. It is specifically encoded by a wider pitch excursion throughout the entire contrastive element. This interpretation of contrast is in accordance with other research which has demonstrated that, as contrast can be superimposed on and combined with the notions of topic and focus, it is a linguistically relevant phenomenon and should be added to the inventory of relevant information structural categories (Chafe 1976, Molnar 2002 *inter alia*).

4.7 Conclusion: marking information structure

In this chapter, the encoding parameters of some major information structural categories have been defined. Three focus domains and two major topic subtypes, and the notion of contrast have been examined. Some attention was given to information structure-related clitics such as =*marlang* ‘given’ and =*gun* ‘contrast’.

Overall, a prominence is heard on the first syllables of the initial word of the construction in focus. In comments, it is often the coverb, thus signaling its greater semantic weight. A [fall] pitch target is associated with the first syllables of focus constructions; its alignment may vary with the length of the focused element. Prominences are also marked by wider excursions and longer durations. The findings show a gradation in the prosodic correlates from less marked to more marked, according to the length of the focused element, thus supporting Baumann et al.'s (2006) prediction that prominence-lending cues increase as the focus domain narrows, so that in Jaminjung the argument focus domain is the most marked and the sentence focus domain the least.

Topics do not receive a prominence and are marked by a [high] target on their initial syllables. The encoding of topics seems to follow a scale of ‘givenness’, where the more given topics are less marked prosodically than the less given topics. The category originally termed ‘new

left dislocated topics' is better described as a sentence in focus, or more appropriately, as an IU in focus. Its status as a member of a wider unit, the prosodic sentence, is shown by it receiving a [rise] at its right-edge, signaling non-finality.

The clitic =*marlang* has the effect of lowering the mean pitch of the entire topics.

The results of the measurements for contrastive focused arguments and topics show that contrast is encoded with the [fall] target on the initial syllable associated with focus, and with a wider pitch excursion on all the syllables of the contrastive element. This specific encoding persuades me to consider contrast as an independent information structure category to focus and topic.

Finally, contrastive elements which host the clitic =*gun* do not have more marked values in their prosodic correlates. As the clitic =*marlang* seemed to attenuate the prosodic encoding parameters in the topics hosting it, the clitic =*gun*, as a marker of contrast, could have exacerbated the prosodic encoding of the elements associated with it. However, the results do not point to such a conclusion, highlighting the need for further research to better define the interplay of morphological markers and prosody in Jaminjung.

Chapter 5 Distinguishing sentence types

In many languages, the meaning of sentences can be changed through their intonational patterns, that is to say, whether a sentence is a statement, a question or a command is often discriminated by intonational means. Or, it could be said that a speaker chooses an intonation contour to express the communicative type of the sentence he/she is uttering and is aware of the illocutionary force associated with it. This chapter will examine the intonational patterns of three types of sentences (declarative, interrogative and imperative) in Jaminjung.

Sentence type involves at least the distinction between declarative and interrogative. The latter is usually further subdivided into two types: Wh-questions, also referred to as information questions, usually containing a question word, and yes/no-questions or polar questions.

In a study of the prosodic realization of utterances such as this one, it is useful to interpret the differences between sentence types in terms of speech acts (see Austin 1962, and Searle 1975 in Saeed 2003). As Couper-Kuhlen (2001:15) points out, “those who wish to see intonation as part of grammar will now assume that intonations are illocutionary-force-indicating devices and distinctive in the way they pair with different illocutions “. In this framework, language not only provides the listener with information about the world (the statement function), it is also about getting things done. In speech, different *speech acts* such as assertions, questions, or contradictions are associated with specific prosodic shapes. By way of example: in English, declaratives end with a falling tone while questions (particularly polar questions) are generally associated with higher pitch in the sentence, most commonly near the end, manifested as a final rising contour, or a higher/expanded pitch range near the end of the sentence. In Russian, polar questions end in a rise if the accented syllable is phrase final; if any unstressed syllables follow the accented syllables, they are pronounced with a fall, the pitch movement becoming a rise-fall (Makarova 2001). In Chinese questions, pitch range is expanded starting from the first accented word to the end of the sentence, while lexical tones are maintained

(Liu and Xu 2005). Indeed, the pattern of declaratives ending in a fall and interrogatives ending in a rise is so common in the world's languages that it is sometimes suggested as one of the rare language universals. Viewed from a wider perspective, sentences bearing the meaning of completion, termination, finality or assertion are associated with low or falling pitch, and those bearing the meaning of inquiry, uncertainty, question and non-finality with high or rising pitch (Bolinger, in Ladd 1996:113, see also Cruttenden 1997: 151-164 for a cross-linguistic overview of the patterns). There are, however, some counterexamples to this generalization about high and low endings. Examples from African languages are discussed in Rialland (2009); Gordon (1999) claims that typical declarative utterances in Chickasaw end with a high boundary tone instead of a low. Ladd (1996:115-118) also provides evidence against the universalist position, most notably from Hungarian polar questions which end in a high-low sequence of tones rather than a rise or a simple high tone.

The challenges in a description of the contours associated with sentence types reside not only in ascribing meanings to them, but also in the characterization of the patterns themselves, especially when considering languages that are dissimilar to the better studied West-Germanic languages. As an example, consider the 'neutral' intonational pattern for declaratives. In West-Germanic languages, the neutral pattern is often described as the 'hat pattern' consisting of a rise early in the phrase, an optional plateau, and a fall at or near the right edge of the phrase (in ToBI terms, H*L-L%). In the AM framework, it is usual to assign a nuclear pitch accent on the last accented word. This is a source of difficulty for the description of languages that do not conform to the same patterning. For instance, in a recent paper on Egyptian Arabic, El-Zarka (to appear) cites not only Egyptian Arabic, but also Danish and Estonian as languages with different patterns. "What EA shares with Danish and other languages of this type, is the high number of pitch accents per phrase which exhibit a downdrifting declination, especially in the type of utterance that is often called 'neutral declarative'". This note of caution is relevant to the study of Jaminjung, particularly as it was shown in chapter 4, that the patterns of accenting focused elements (placement of pitch accents) differ from those found in English.

Furthermore, the characterisation of a *declarative neutral contour*, based on the position of accents, mixes the encoding of two functions: that of marking focus and that of differentiating sentence types, a pitfall which the PENTA model attempts to avoid. Each function receives its encoding, the parameters of which are specific.

This chapter is based on the following assumptions. Intonation contours constitute an integral part of the meaning attributed to sentences (the signifier side), a view widely accepted, at least since Bloomfield (1933:159ff.), who counted *modulation* as one of the four ways of arranging linguistic forms, the other three being: the ordering of the constituents, their phonetic alterations, and finally the selection of the forms themselves. It is also assumed that the scope of the contour is the intonation unit, therefore the overall shape of the contours is considered as being potentially meaningful. On the other hand, it is assumed that using intonational means to distinguish sentence types is only one of the possible strategies⁴⁸ available in any given language, a view that is congruent with construction-oriented grammars (Fried and Ostman, 2004:13ff.).

This chapter is structured around the three main sentence types: declaratives, interrogatives and imperatives. Section 5.2 starts with an examination of simple declarative sentences (formed of one IU) in order to obtain a basic understanding of the default intonation for declaratives. Then, complex declarative sentences are analysed to determine how the default patterns behave as the complexity of the sentence increases. Specific contours of prosodic sentences formed of 2 IUs will be described, according to whether the second IU is another verbal clause, a non-verbal clause consisting of a coverb as a non-finite predicate, or an NP functioning either as an afterthought or a reactivated topic. To complete this inventory, constructions where the first IU is ‘iconically lengthened’ and vocatives will

⁴⁸ Other means including word order variations and morphological marking.

be examined⁴⁹. Section 5.3 examines interrogatives and their subtypes, information, polar, and tag questions (the tag is used to seek hearer's response).

Section 5.4 examines imperatives/commands, and finally, section 5.6 compares the three main sentence types. In sections 5.7 and 5.8, special attention is given to vocatives and quotation speech. The hypotheses underpinning the analysis of each of the sentence types are presented at the beginning of each subsection.

5.1 Methodology

The overall F0 contour in declarative, interrogative and imperative utterances is taken into account as well as the specific characteristics of the syllables at the left and right boundaries of the IUs. The quantitative analysis consists of measurements of:

- the mean pitch in the last syllable of the IU relative to the other syllables in the IU;
- excursion size: variation in minimum and maximum pitch within each syllable and between the syllables in the IU;
- final velocity of the syllables at the boundaries (first, second, penultimate and final) as an indicator of the underlying pitch target;
- alignment of pitch targets from the most prominent syllable to the final syllable of an IU;
- syllable duration of final syllables relative to the other syllables in the IU.

The IUs are labeled according to their subtypes as well as the number of words and their positions in the IUs; first, second, penultimate and last syllables in IUs are labeled. The number of syllables in the second IUs is also noted. In some of the tests, because the number of syllables in IUs

⁴⁹ Complex sentences consisting of two IUs can also have an NP introducing a new participant/referent to the discourse in a separate IU at the beginning of the sentence followed by the verbal clause, but this construction will not be analysed here (§4.5.5).

varies from 2 up to 15 syllables, but with very few tokens of more than 6 syllables, creating many subgroups with very few tokens that complicate the interpretation of statistics, only the syllables at the left (first and second) and right (penultimate and final) periphery are considered. A statistical analysis is conducted to validate the results.

5.2 Declaratives

In this section, I will seek to establish the prosodic encoding of declarative sentences. I will explore simple declarative sentences, and then complex declarative sentences, including a subtype to which I will refer as the ‘lengthened construction’. Complex sentences are formed of a first IU, usually with a falling contour, but possibly also with a fall-rise or level contour, and a second IU, which may be a verbal clause, a non-verbal clause or an NP. The three contours found in the first IUs are compared, as well as that of the lengthened constructions. Then second IUs are examined, starting with IUs that consist of verbal clauses, distinguishing between sentences with IUs that are juxtaposed and those that are separated by a pause. This is followed by a comparison of the correlates of non-verbal clauses (non-finite predicates) and NPs (afterthoughts or reactivated topics). The IUs that follow an iconically lengthened first IU are compared with those of the non-lengthened sentences.

Detailed descriptions and examples are given at the beginning of each subsection.

5.2.1 Hypotheses

It is expected that the prosodic encoding of declaratives conforms to the patterns described cross-linguistically, of a falling contour. The overall pattern is also expected to reflect, particularly in complex sentences, the contribution of the elements in question to the sentence, either as new information or given information. The prosodic characteristics of clauses, verbal and non-verbal, and afterthoughts should resemble that of simple declaratives, while reactivated topics, which do not contribute new

information, are expected to have different intonation contours, for a start not to receive any of the features associated with the encoding of focus on their initial syllables (§4.5). It is also hypothesised that the degree of syntactic integration to the main clause (corresponding to the first IUs in this dataset) is reflected in the prosodic encoding of the IUs, hence the encoding of verbal clauses and non-finite predicates is expected to differ from that of NPs, both in afterthoughts and reactivated topics.

I also hypothesize that the patterning of the sequences of IUs in complex sentences is ‘compositional’, each IU receiving a specific encoding according to its function. Thus, with regard to the constructions with iconic lengthening, I expect the contour of the first IU to have the same encoding as that of the level contour in non-lengthened constructions, except for the durational cue. I further expect the second IU to bear the same encoding as those of non-lengthened constructions.

5.2.2 Simple sentences (1 IU)

In this section, I will examine the contours and the prosodic cues associated with single IU declaratives, that is, these constitute a prosodic sentence of their own. As this overlaps in part with the examination of IUs and prosodic sentence in chapter 3, and with the marking of information structure in chapter 4, I simply aim to confirm the patterns revealed by the findings of the previous chapters and associate them specifically with the declarative sentence type, which will then be used as a default pattern for the subsequent comparisons with other sentence types.

Simple declaratives in Jaminjung consist of a succession of phrases, each of which is marked with a slight pitch reset (§3.3.4.1); it usually contains a comment which contains an accented or prominent syllable, functionally associated with the focused element and usually, but not exclusively (§4.2) found in comment-initial position. They end in a [low] pitch target, as shown in chapter 3, thus presenting a gently terraced falling contour. This contour is the unmarked contour in statements.

5.2.2.1 Dataset

There are 69 tokens in the dataset, from 9 different speakers, all topic-comment or comment-only constructions. They are verbal clauses, except for one token that consists of a nominal predicate and 2 non-finite predicates.

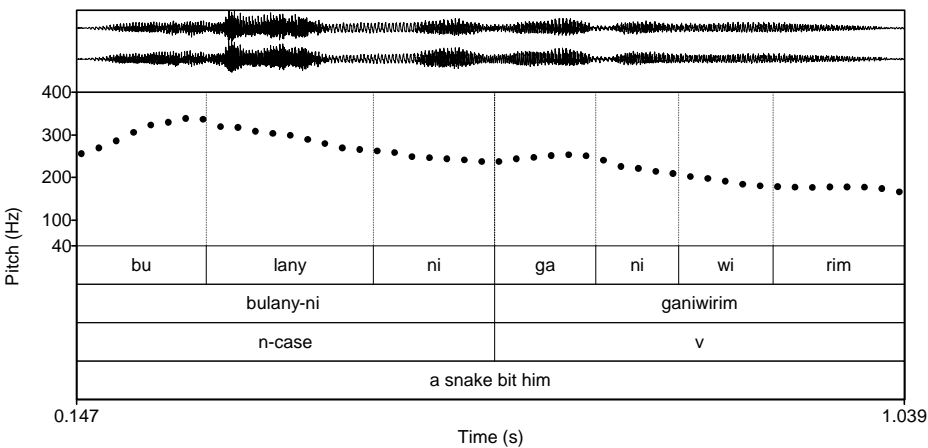
Figure 5-1 shows examples of simple declaratives. Example (87) is an IU with an explicit topic-comment structure, and example (88) has a topic followed by a nominal predicate. Both IUs display the falling contour punctuated by the slight pitch [high] target associated with the first syllables of prosodic phrases (see ch. 3).

(87)

bulany-ni	gani-wiri-m
snake-ERG/INST	3sg:3sg-bite-PRS

'A snake bit him'.

[IP:ES08_16_02]



(88)

nginju jarlag wirib
PROX good dog

'This one is a good dog.'
[BH:CS_72_01]

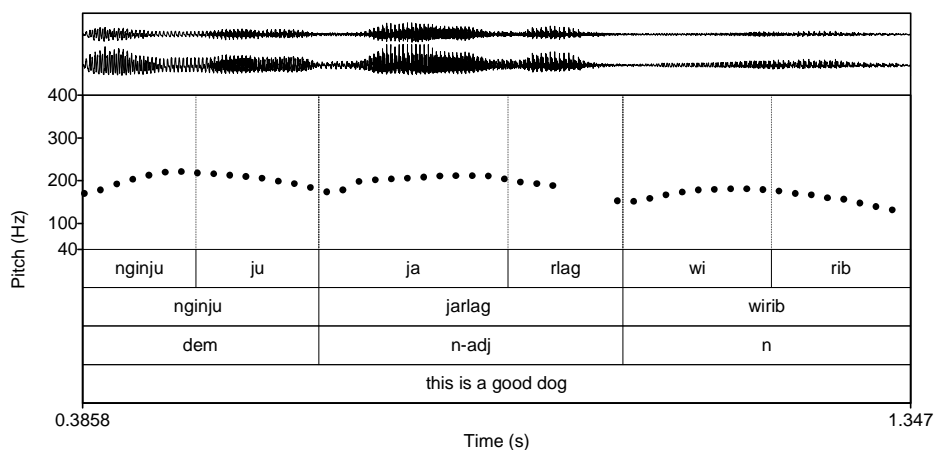


Figure 5-1 Simple declaratives, the first example is an IU with an explicit topic; the second consists of a topic and a nominal predicate.

5.2.2.2 Average pitch, final velocity and pitch targets

IUs in this dataset are more likely to have 6, 7, or 8 syllables and 3 words, but can have up to 5 words and 15 syllables. 64 of the 69 tokens have an overall falling contour, 3 have a fall-rise contour and 2 have a flat contour.

The average mean F0 for each syllable, taking into account its position within the word and the relative position of the word in the IU, are shown in boxplot form in Figure 5-2. The prosodic words are separated by vertical lines. Only the 66 IUs ending in a non-rising contour are used in this illustration. A multifactor ANOVA test, with mean F0 as dependent variable and speaker and syllable position as factors, shows that, although there are significant differences between speakers, these differences do not impact on the overall measures by syllable position. Another test with mean F0 as dependent variable and syllable position and relative position of word in the IU as factors yields significant results ($F(19, 569) = 10.067, p = .000$). The decrease of mean F0 is apparent from the first word (leftmost section of the graph) to the final word (rightmost section), confirming the ‘gentle fall’ suggested by impressionistic observations. There is also a slight reset in the

first syllable of each prosodic phrase as evidenced by the rise in values when compared to the last syllable of the preceding word.

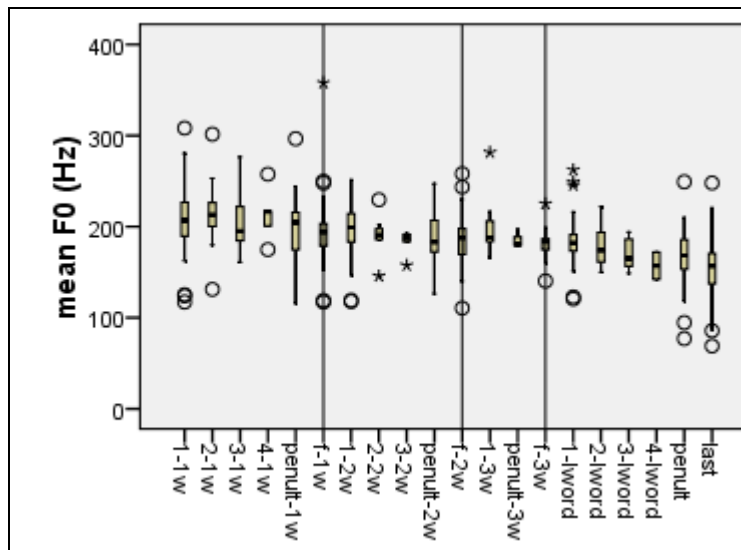


Figure 5-2 Average mean F0 values of the syllables of simple declaratives.

The measurements of final velocity are examined in order to determine the underlying pitch targets of final syllables in declarative sentences. 38% of the IUs in this dataset end in creaky phonation (a marker of the right boundaries of prosodic sentences §3.3.3.4.3⁵⁰), the penultimate syllable is thus also considered.

The patterns in the measurements over the entire IU are shown in Table 5-1. For ease of comparison, the first and second syllables of words are highlighted in light grey, and the penultimate and final syllables in darker grey. The differences are found to be statistically significant ($F(19, 554) = 3.213, p = .000$). Post hoc tests show that the first syllable of the IU differs from the others, but that the other values do not vary significantly. The standard variations are again large in some cases, and the widest values have not been removed from the dataset. It is still considered that the averages are representative of the data, the large variations indicating that values that are not necessarily closely clustered near the mean, a situation

⁵⁰ Creaky phonation can influence the final velocity measurements taken 30ms before the syllable boundary.

that is not unusual with ‘natural’ data. The average value of all final syllables of the IU is -5.74st/s. That of the penultimate syllables is -17.80st/s, suggesting a [low] target at the right edge. At the left edge, the first syllables have values averaging 5.11st/s, suggesting a [high] pitch target. I will not associate this pitch target with the encoding of sentence types, however, as it was shown in chapter 4 that the first syllables of topics and of the focused elements receive specific encodings. Thus, this value is a very general average, keeping in mind that the IUs in this dataset may present different information structure characteristics.

The second syllables have negative values, fairly close to the zero mark. When taking into account the declining values of mean F0 observed in the syllables, the pattern that is suggested is of a succession of [mid] targets until the end of the IU, where there is perceptually a fall in pitch. It is suggested that a [low] target is actually realized in the last syllables, where the velocity of pitch changes to values of -5.74st/s, corresponding to the general description of pitch targets in chapter 3.

Syllable position in word, position of word in IU		Mean final velocity (st/s)	Std. Dev	N
1st word	1	5.11	20.43	66
	2	-8.35	25.06	33
	penult	-9.19	14.03	45
	final	-11.54	20.95	60
second word	1	-2.65	19.99	49
	2	-16.45	16.15	7
	penult	-4.09	14.11	19
	final	-7.70	25.17	39
3rd word	1	-13.18	17.02	16
	penult	-18.43	6.76	3
	final	-9.61	17.40	12
Final word	1	-11.05	16.43	62
	2	-15.89	12.19	17
	penult	-17.80	17.53	41
	final	-5.74	24.17	53

Table 5-1 Final velocity values of the syllable in simple declaratives, shown according to their position in the word and in the IU.

The values in the table above conflate all IUs, not taking into account the length of the IUs, either in terms of the number of words they contain or the number of syllables in the words. A test is conducted,

isolating the final words in the IUs and organizing them by the number of syllables each of them contains so as to compare sets that are as homogeneous as possible. The results are shown in the graphs in Figure 5-3 below, representing 2-, 3- and 4-syllable words, as there are enough tokens of each to appreciate the patterns (24, 25 and 12, respectively).

The 2-syllable words have average values of -18.11 and -10.11st/s; 3-syllable words of -7.52st/s, -19.14st/s and -.09st/s; and the 4-syllable words have values of -9.81st/s, -17.80st/s, -12.55st/s, and -7.19st/s. In interpreting the results, it is necessary to take into account possible effects from the encoding of other functions. Thus, the penultimate in 3-syllable words may share properties with both the second and with the penultimate in 4-syllable words. As a generalization, the pattern of lower values in the penultimate syllables veering towards zero in the IU-final syllables is repeated in all words to a varying degree, thus confirming the interpretation of a [low] final target.

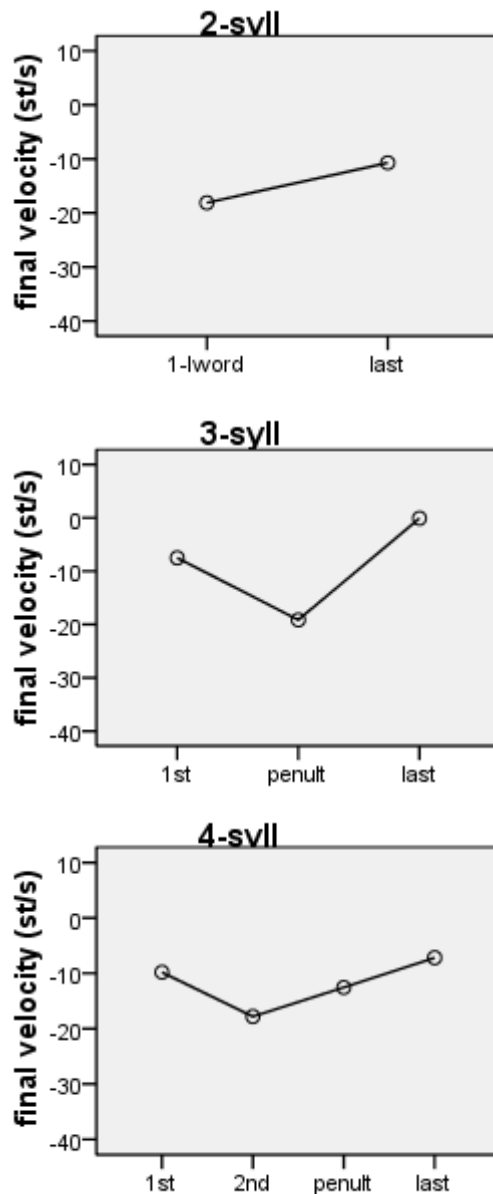


Figure 5-3 Final velocity values in the IU-final words of varying length.

In conclusion, the values for the first syllables will not be discussed further here as they have been investigated in the sections on grouping and information structure (chapters 3 and 4) and are understood to encode different functions than sentence type. It suffices to restate that the higher mean F0 found in the first syllables is the result of a pitch reset marking a new IU.

The encoding for declarative sentences can be summarized as follows: they present an overall falling contour, characterized by a [high]⁵¹ or [fall] pitch target on the first syllable, followed by a sequence of [mid] tones, and demarcated by a final fall, realized as a [low] pitch target on the final syllable.

Similar descriptions have been made for many other Australian languages, among others Dalabon (Fletcher 2007), Alawa (Sharpe 1972:36), Kuku Yalanji in Northern Queensland (Patz 2002:40), and Dyirbal in Central Australia (King 1994:597).

5.2.3 Complex sentences (2 IUs)

In this section, complex sentences are investigated. The correlates of the contours of the first IUs are compared first, based on the shape of the contour, and comprising the prevalent falling contours, contours with a fall-rise, and level contours. The first IUs of the lengthened constructions are also included in this comparison.

The analysis then concentrates on the second IUs in complex sentences and lengthened constructions. Firstly, verbal clauses are examined in some detail, establishing the impact of the presence or absence of a preceding pause on the correlates, and comparing them with simple declaratives.

A comparison is then made between the correlates of the various subtypes of second IUs, i.e. non-verbal IUs (or semi-independent predicates, the two terms are used interchangeably hereafter), NPs (afterthoughts and reactivated topics). Finally, the second IUs of the lengthened constructions are compared to the other subtypes of IUs in order to establish whether the second IUs in both lengthened and non-lengthened sentences have the same encoding.

⁵¹ The first syllables could also receive a [fall] target if they were at the beginning of a focused constituent.

5.2.3.1 Dataset

Prosodic sentences consisting of two IUs are retained for this analysis. They are subdivided into 3 main subtypes. They all have a first IU with verbal clause, and a second IU that is either i) an independent or subordinate verbal clause ii) a non-verbal clause consisting of a non-finite predicate, usually a coverb iii) an NP which serves different functions⁵², either as afterthought, or reactivated topic. They are described in more detail hereafter.

Independent clause (32 tokens in this dataset), or a subordinate clause such as a relative clause marked by the clitic *=ma* (2 tokens) can be contiguous with the first clause (11 tokens) although they are more likely to be separated by a pause (21 tokens). Both clauses/IUs usually have falling contours, but the first IU may also have a flat or falling-rising contour. Most of the contours described as ‘fall-rise’ have a falling contour until their last syllables, where the rise occurs. These contours can also be found in IUs preceding non-verbal clauses and NPs. The fall-rise contour is attested for all speakers and in different genres (conversation, elicitations and narratives). This contour often coincides with the Kriol particle *na* ‘now’ but not exclusively so, and not with its Jaminjung equivalent *=biyang*. This is the ‘default’ fall-rise contour, but in 3 out of the 14 tokens described as having a ‘fall-rise contour’ in this dataset, the rise starts earlier, that is, on the first syllable of the IU-final word. There are no instances of fall-rise contours starting from the initial prominence (at least in this dataset). First IUs with a level contour are also attested, but only in narratives. 6 out of the 9 tokens are uttered by the same male speaker, so, in order to avoid skewing the results, some of the tokens from this speaker have been filtered out. There are no obvious syntactic, semantic or pragmatic reasons why speakers use the level or rising contours in the first IU.

Examples of sentences comprising two verbal clauses with the three contours in the first IUs are shown below. Examples (89-92) are instances of

⁵² The functions analysed here may not represent all possible functions of NPs in this position. They are those most often encountered in the Jaminjung data.

sentences made of two verbal clauses, all from elicitation tasks in which speakers are asked to describe scenes from pictorial stimuli, except the bottom pane which is extracted from a mythological narrative. In example (89), IUs are juxtaposed without a pause. The first IU ends after *garumany* ‘he/she came’; both IUs have a characteristic overall falling contour. Examples (90) and (91) show prosodic sentences with a fall-rise contour on the first IU, one coinciding with the particle *na* ‘now’, and the other where the rise coincides with the beginning of the inflected verb *ganamany* ‘he put/did it’. Example (92) shows a first IU with a level contour.

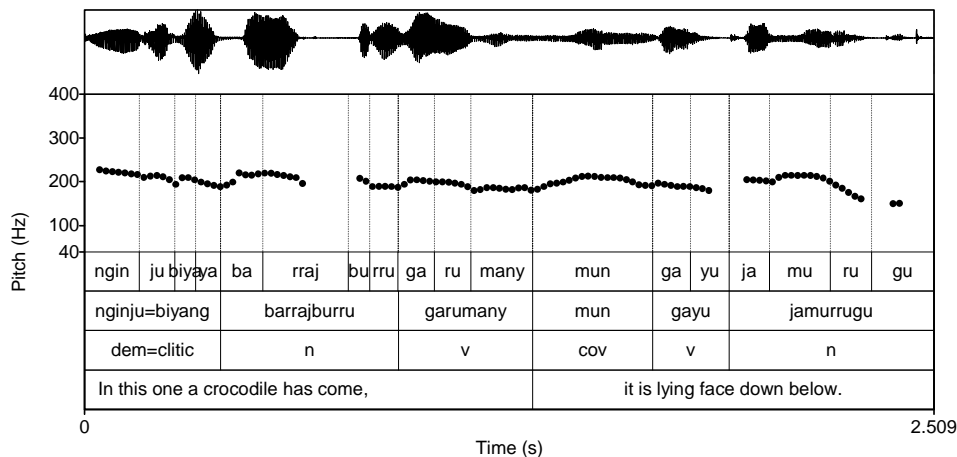
(89)

nginju=biyang barrajburru ga-ruma-ny mun ga-yu
 PROX=now saltwater.crocodile 3SG-come-PST belly.down 3SG-say/do

jamurrugu
 below

‘In this one a crocodile has come, it is lying face down below.’

[JM:CS07_62_01]

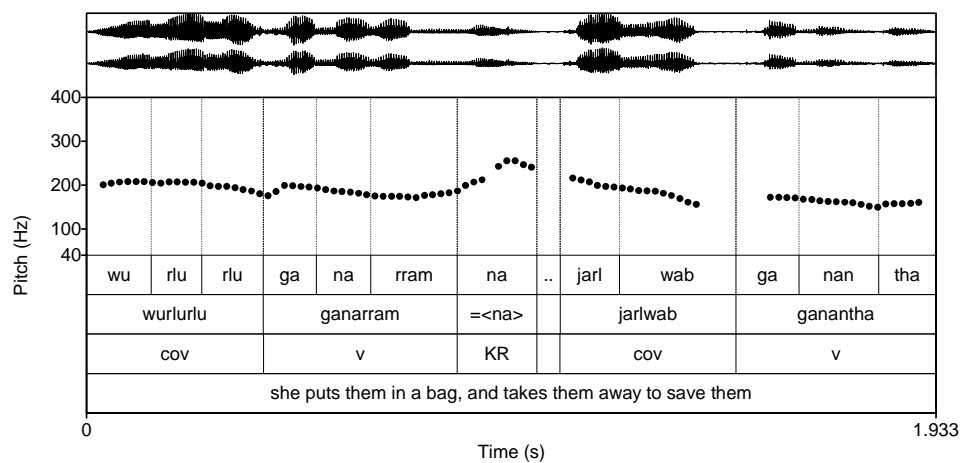


(90)

wurlurlu gan-arra-m=<na> jarlwab gan-antha
enter.of.many 3sg:3sg-put-PRS=now saved 3sg:3sg-take.PRS

'She puts them in a bag, and takes them away to save them?'

[IP:ES96_08_02]

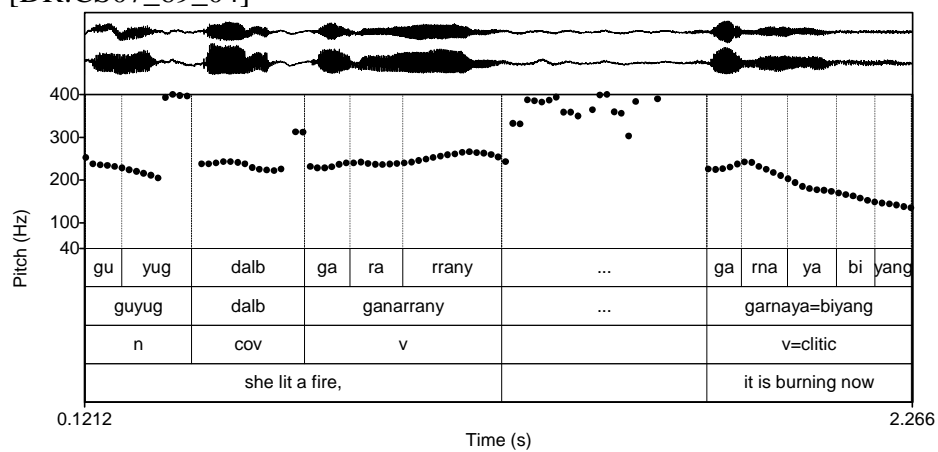


(91)

guyug dalb gan-arra-ny... ga-rna-ya=biyang
fire light.fire 3sg:3sg-put-PST 3sg-burn-PRS=now

'She lit a fire, it is burning now.'

[DR:CS07_69_04]



(92)

buru=biyang ga-jga-ny... yagbali birdij gan-arra-ny
 return=now 3sg-go -PST place find 3sg:3sg-put-PST

'She went back then, she found a camp.'

[DM:MH96_19_01]

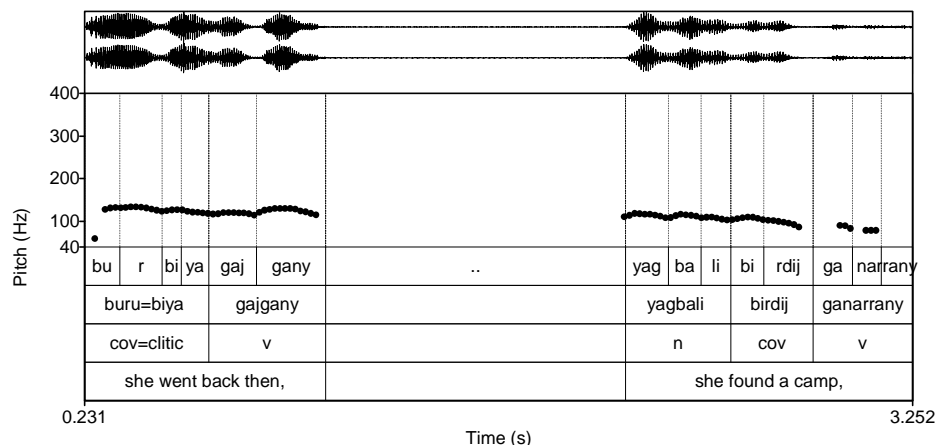


Figure 5-4 Sentences consisting of two clauses. The top pane shows a first IU with a falling contour; the middle panes show the ‘fall-rise’ contour, the first with the rise coinciding with the clitic =na ‘now’, and the second with a rise from the first syllable of the final word. The bottom pane illustrates a level contour.

The second IU may be a non-verbal clause (17 tokens), which consist of a coverb with or without an argument. I will prefer the term non-finite predicate, although they are also referred to as semi-independent predicates in the following analysis. In Schultze-Berndt’s (2000:136) definition, coverbs as semi-independent predicates introduce new events or express a different aspect of the same event (simultaneous or resultative sub-event). They may have the same semantic relationship with the predicate in the preceding main clause as this coverb and an inflecting verb in a canonical complex verb (§3.2.3.2.2 and §3.2.3.2.3). Alternatively, a coverb as semi-independent predicate can occur without any semantic relationship with the preceding predicate. While applying to most tokens in this dataset, this definition cannot be applied to all, as Schultze-Berndt reserved it to bare coverbs and this datasets includes coverbs with markers of subordination such as case markers.

In example (93), the coverb *waga* ‘sit’ in the second IU predicates on one of the arguments of the predicate in the preceding intonation unit, the elided topic/subject ‘the boy’, and receives a depictive reading with respect to this predicate.

(93)
 nginju wagurra-ni ga-yu... waga
 PROX rock-LOC 3sg-be.PRS sit

'Here he is on the rock... sitting.'

[JM:CS06_35_01]

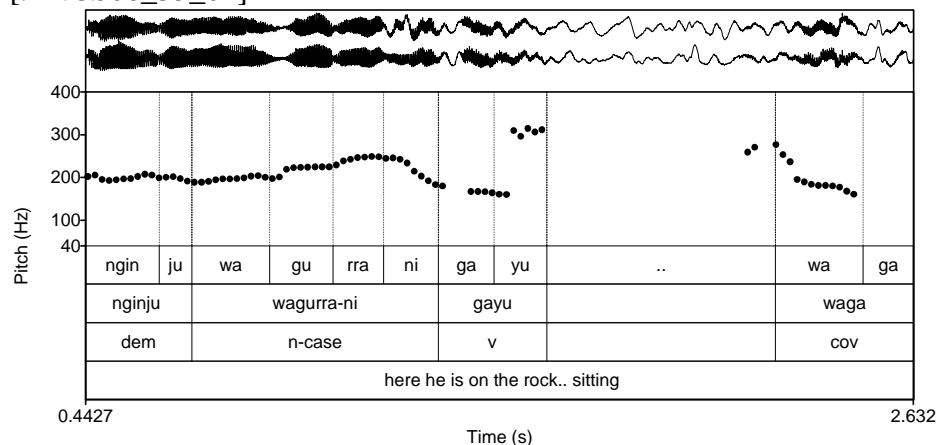


Figure 5-5 A prosodic sentence made of a main verbal clause, followed by a coverb *waga* ‘sit’ as a non-finite clause.

Finally, the second IU may consist of an NP which serves different functions⁵³, either as afterthought, or reactivated topic. Afterthoughts, usually non-verbal elements, such as NPs or adverbials, are used to supply additional information. I follow Auer (1996), who defines them as structures in which something is added after a sentence is completed as addition or modification, and Chafe (1994:142), who states that “it is common for speakers to complete the scanning of a center of interest, indicating completion by falling pitch, and then supplement the information already conveyed with a brief additional focus... “. The interpretation of NPs as afterthought is to a large extent dependent on the linguistic and extra-linguistic context – afterthoughts are often treated as the result of lapses in speech planning. NPs interpreted as afterthoughts in the prosodic sentences analysed in this dataset are always preceded by a pause, never topical, and often circumstantial elements.

⁵³ The functions analysed here may not represent all possible functions of NPs in this position. They are those most often encountered in the Jaminjung data.

Example (94) (Figure 5-6) is of a sentence where the second IU is an afterthought, used by the speaker to add information about where the fight took place.

(94)

buny-mama-ji-na <jiya>... bala-ni
 3du-hit-REFL-IMPF there plain-LOC

'They were fighting there, on the plain.'

[DM:MH96_19_01]

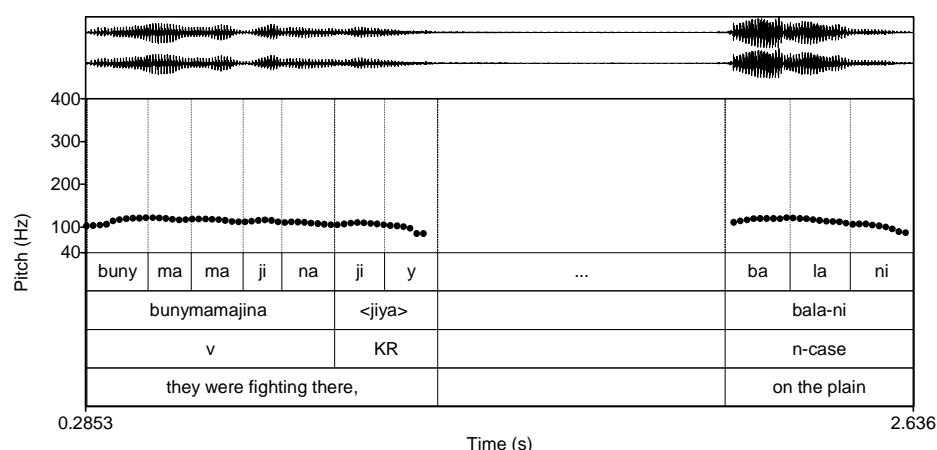


Figure 5-6 A prosodic sentence with an NP functioning as an afterthought in the second IU.

Reactivated topics (right dislocation, see also §4.5.4.2) for their part help to solve potentially unclear referents, and are often found to be the topic referent for the following sentence. Restating the topic is a fairly common strategy used by Jaminjung speakers, to palliate the lack of clarity brought about by the frequent topic elision in the main clauses. The term *antitopic* is found in the literature for similar constructions. Chafe (1976) defines it as a pragmatic category which functions to ‘confirm established information’, noting that it frequently occurs at the right periphery. Lambrecht (1994:203ff) defines *antitopics* in contradistinction with afterthoughts. For him, the propositional information is put on hold temporarily until the referent is fully named, in other words, the referent is accessible although not yet an established topic: “the presuppositional structure of the antitopic construction involves a signal that the not-yet-active topic referent is going to be named at the end of the sentence”. This definition is based on the notion that the topic is not yet established, and that this is somehow signaled in the first part of the sentence. This does not

correspond to the constructions in this analysis, in which the first IU is an independent, ‘complete’ clause, after which the speaker decides to reiterate the topical referent. This motivates the choice of the term reactivated topics over that of antitopics.

A reactivated topic is shown example (95). The speaker has been talking about getting mussels, and this has been the topic of the preceding prosodic sentence. He continues talking about them now and restates the topic at the end of this sentence with *naribu=marlang* ‘mussels=GIVEN’

(95) (also in 80))

gurrany	yanji-ngarna	jarlig...	naribu=marlang
NEG	IRR:2sg:3sg-give	child	mussel=GIVEN

'You should not give them to the children, those mussels.'

[BH:CS07_72_01]

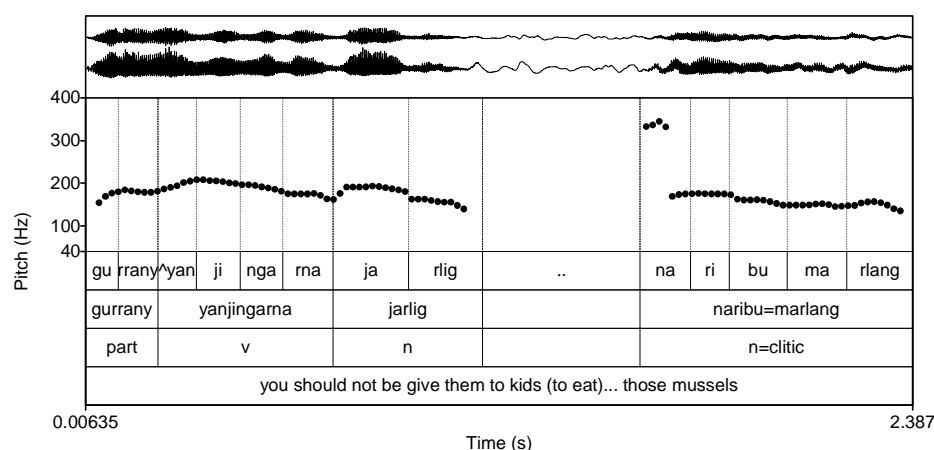


Figure 5-7 A prosodic sentence with an NP functioning as a reactivated topic in the 2nd IU.

A last construction is investigated under the complex sentences label, usually involving two (or more) IUs. It will be referred to as the *iconic lengthening* construction. Lengthening is a device used in Jaminjung by which the durational aspect of an action is iconically represented by lengthening the vowel in the coverb or verb in the first IU⁵⁴. It is usually bounded by a second IU which may be a verbal clause (10 tokens), a non-verbal clause such as a coverb as non-finite predicate (12 tokens), or an NP

⁵⁴ The lengthening may occur elsewhere in the IU; this analysis is restricted to the more common lengthening of the final vowel.

(8 tokens) functioning as an afterthought. Schultze-Berndt (2002:290) defines this construction, which she terms *event delimiting construction* as follows:

...This is a bipartite construction, consisting of a main predicate followed by a “delimiting” predicate. It is mainly defined by a particular prosodic contour: the final syllable of the main clause is lengthened, often considerably, with level intonation, and the delimiting predicate has a sharply falling boundary intonation.

This appears to be the privileged manifestation of iconic lengthening in Jaminjung, but it is not exclusive. There are examples in the data of lengthening being used in a simple declarative (2 tokens), and a few other prosodic sentences that are longer, with at least 3 IUs (7 tokens), where a second, independent clause occurs between the two types of clauses described earlier. There are also some instances of a succession of clauses all containing a lengthened element.

A similar construction has been described in various Australian languages, sometimes referred to as a ‘narrative high monotone’, ‘stylized high sustained intonation’, ‘continuous action’ and ‘prolonged action’ e.g. Warlpiri (Chapman 2007), Mawng (Hellmuth et al. 2006:3), Bininj Gunwok (Bishop⁵⁵ 2003:89ff), Iwaidja (Birch 2003), Gooniyandi (McGregor 1990), Nunggubuyu (Heath 1984), and Alawa (Sharpe 1972).

Examples of the construction are shown in Figure 5-8, example (96) showing a non-finite predicate in the second IU, example (97) showing an NP in the same position, and example (98) a verbal clause.

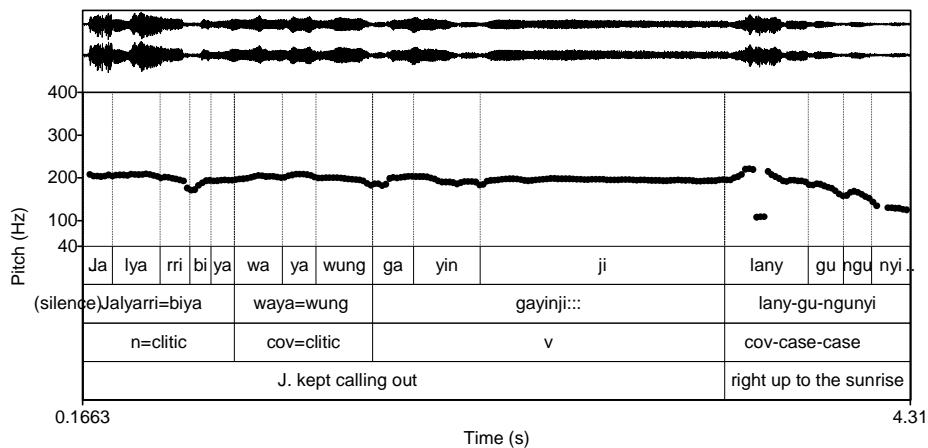
⁵⁵ Bishop gives a good summary of the construction in various Australian languages (2003, pp 89-95).

(96)

Jalyarri=biyang waya=wung ga-yinji... lany-gu-ngunyi
<subsection>=now call=RESTR 3sg-go:IMPF sunrise-DAT-ABL

'J. kept calling out, right up to the sunrise.'

[DB:ES97_01_03]

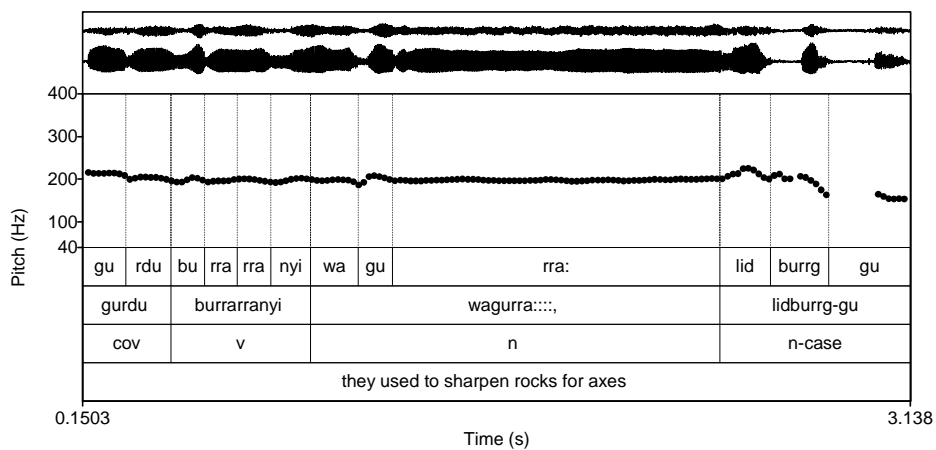


(97)

gurdu burr-arra-nyi wagurra :: lidburr-gu
grind 3pl:3sg-put-IMPF rock axe-DAT

'They used to sharpen rocks for axes.'

[IP:ES08_04_02]



(98)
 thanthiya buru-mayan ga-gba guruwuny
 DEM return 3SG-be.PST bottle.tree

jard gan-unga-ny
 upright 3SG:3SG-leave-PST

'Then he was going back, (and) he left a boab tree standing there.'
 [DD:ES99_08_01]

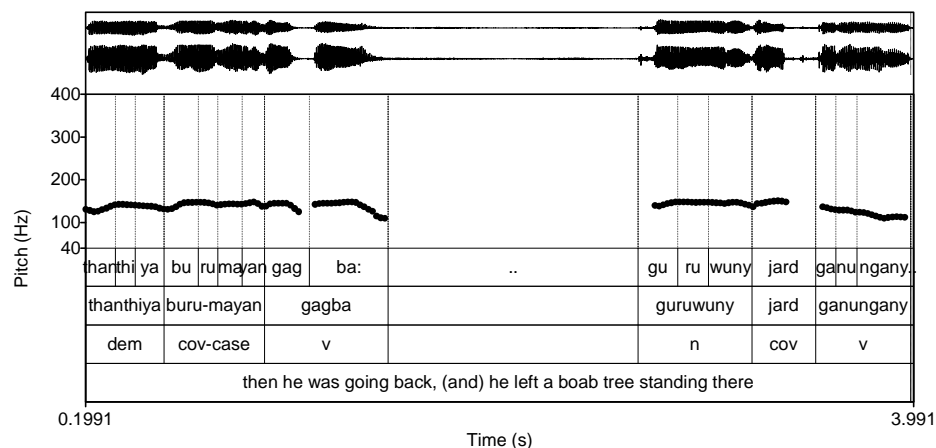


Figure 5-8 Iconic lengthening in the first IU followed, in the top pane, by an IU consisting of a non-finite predicate; in the middle pane by an NP; and in the bottom pane by a verbal clause.

5.2.3.2 First IUs

In this section, the correlates of the three contours found in the first IUs of complex constructions are measured, as well as those of the lengthened constructions. Contours are impressionistically characterized in the following manner.

Falling contours have a peak on the initial syllables: the pitch then decreases until the final syllable. The fall-rise contours are similar to the falling contours, with the distinction that there is a sharp rise in pitch in the last syllable of the IU. In level contours, after the initial peak, the pitch is maintained at the same level throughout the IU; there is often an audible drop in pitch in the last portion of the final syllables. The first IUs in the lengthened constructions also have a level contour. Whether it has the same correlates as the level contours just described is one of the questions this section seeks to answer.

The following describes the correspondence of the contours with the types of IUs found in second position in the complex sentences of the dataset (not including the lengthened constructions). The frequency of occurrences of each of the contours is shown in Table 5-2. Falling contours are the most frequent, accounting for 79% of all tokens in the dataset; fall-rise contours account for 18%, and level contours for 4%. IUs with falling contours are well spread out across the four subtypes of IUs in second position; preceding verbal clauses in 27% of the tokens, afterthoughts in 30%, reactivated topics in 21%, and non-finite predicates in 22% of all tokens. The fall-rise contours are found mostly before verbal clauses, with 64% occurrences. They are also found before reactivated topics in 14% of tokens; and before non-finite predicates in 21%. The level contours precede only verbal clauses and afterthoughts, with 67% and 33% of the tokens, respectively. The correlations do not reach statistical significance. Thus, the falling contours are most frequent, evenly spread between all the subtypes of second IUs. Fall-rise contours are the second most common, found mainly before verbal clauses; they also occur in the other subtypes, but not before afterthoughts. Level contours are the least frequent, and only occur before verbal clauses and afterthoughts.

Contours of 1st IU/count		Type of 2 nd IU				
		verbal	afterthought	reactivated topic	non-finite clause	Total
fall	Count	17	19	13	14	63
	%	27.0%	30.2%	20.6%	22.2%	100.0%
fall-rise	Count	9	0	2	3	14
	%	64.3%	.0%	14.3%	21.4%	100.0%
level	Count	2	1	0	0	3
	%	66.7%	33.3%	.0%	.0%	100%

Table 5-2 The rates of occurrence of the falling, fall-rise and level contours of the first IUs in the prosodic sentences in this dataset, correlated with the types of second IUs.

The measurements of the mean F0, excursion size, final velocity and duration are presented here. To capture the patterns, only the syllables at the left and right boundaries, including the first and last three syllables of the IUs, are retained. The results are illustrated in the form of graphs, shown below.

Considering mean F0 first, the differences in values between all the tested contours are statistically significant ($F(3, 622) = 8.507, p = .000$), the results are shown in Figure 5-9.

The falling contours have patterns of declining values reminiscent of those of the simple declaratives. In fall-rise contours, the observed patterns of F0 values are confirmed by the measurements; they decline until the penultimate syllables to rise only in the final syllable. Interestingly, the overall values are higher throughout than in the falling contours. The level contours start with lower pitch, around 160Hz, and rise slightly in the second syllables to decline only slightly until the last syllable.

The first IUs of lengthened constructions can not be equated with the level contours as the values in the syllables at their left boundary, around 200Hz, are extremely close to those in IUs with a falling contour. The values in the lengthened contours do not fluctuate much throughout the IU, indeed, there is no apparent declination.

Although the results for the level contours need to be treated with caution given the low rates of occurrence, the general statement can be made that their average pitch is certainly lower overall than that of the other contours. The first IUs of the lengthened constructions start like a falling contour (the most common contour) and the level is maintained until the final syllables. This matches with descriptions of this type of construction as ‘high sustained contours’ (Bishop, 2003) in other Australian languages.

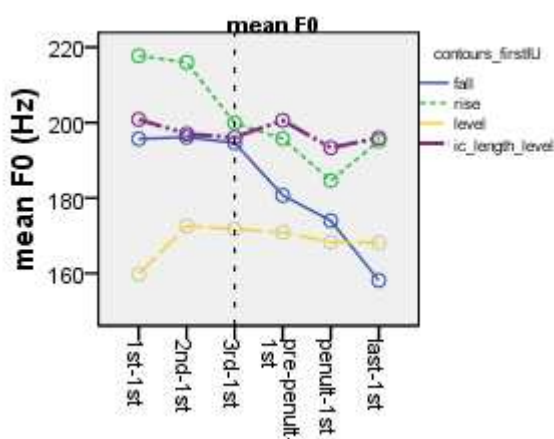


Figure 5-9 The values of mean F0 in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.

The variations in the excursion size are also significant ($F(3, 630) = 3.251, p = .021$) (Figure 5-10). The first syllables in all three contours have fairly similar excursion sizes, between 3 and 4st. The falling and fall-rise contours differ mainly in the syllables at their left boundaries, the fall-rise contours having much larger pitch excursions than the falling contours, while the syllables at the right have surprisingly similar values. The excursion size in IUs with level contours declines throughout, with much reduced expansion spans in the later portion of the IU, as expected.

The values of the syllables at the left boundary of lengthened IUs are close to those of the falling contours, with even wider excursions in their initial syllables. At the right edge, the lengthened IUs have only slightly smaller pitch excursions than the fall-rise and falling contours. The smallest pitch excursions are found in the level contours, which decrease over the first part of the IU to end with an average lower than 1st.

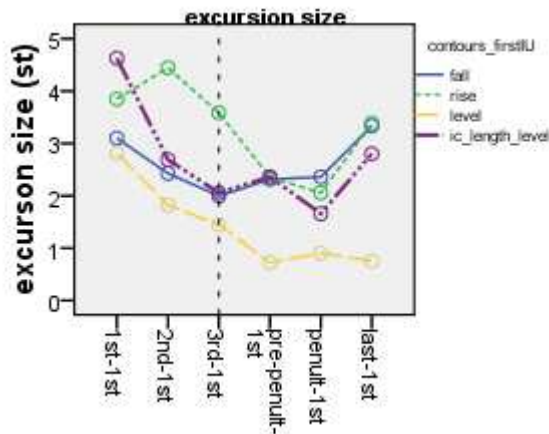


Figure 5-10 The values of excursion size in the first IUS with a fall, fall-rise, level and lengthened contours. Only the first and last three syllables of the IUs are shown.

As to the final velocity values, their difference is not statistically significant, but some trends are nevertheless found in the data (Figure 5-11). The first syllables of falling and level contours have values above zero which then fall to the vicinity of -10st/s. This is the expected pattern of an IU starting with a topic with an initial [high] target. In the fall-rise contour, the first syllable is near -7st/s and the second syllable around -17st/s, indicating a focused element in initial position. In lengthened contours, both syllables are around -10st/s. At the right edge, the penultimate syllables all have fairly similar values, between -5 and -10 st/s. They are then

differentiated in the final syllables, in accordance with the observed pitch patterns. In the fall-rise contours, the velocity averages 12st/s, indicating a [rise] target.

In the final syllables of falling contours, velocity values rise slightly from -10st/s to -5st/s, suggesting a [low] target. They do the opposite in the level contours, with a very slight drop from the penultimate to the final syllables. Interestingly, the pattern in the lengthened IUs and in the level contours is almost identical. These are interpreted as [mid] underlying targets.

The differences in the contours at the left edge are more likely to be associated with the element that is placed there, either a topic or the beginning of a focused constituent. The perceived differences in the contours are thus more likely to be encoded in the syllables at their right-edge. To confirm the significance of the targets at the right edge, a further test is done, this time filtering all but the IU final syllables. The differences then are significant ($F(2, 67) = 3.694, p = .030$). The last syllables for the falling contours have a final velocity value of -4.45st/s; that of fall-rise contours is 12.53st/s; for level contours, it is -9.06st/s; and for the lengthened contours it is -8.89st/s. The underlying targets suggested by the values of the penultimate and final syllables for falling contours are of a [low] as opposed to a [rise] in the fall-rise contours. As to the level and lengthened contours, the values could also suggest a [low] or a [fall], but the description of the target must be, based on perceptual impression, a [mid] target on the final syllable.

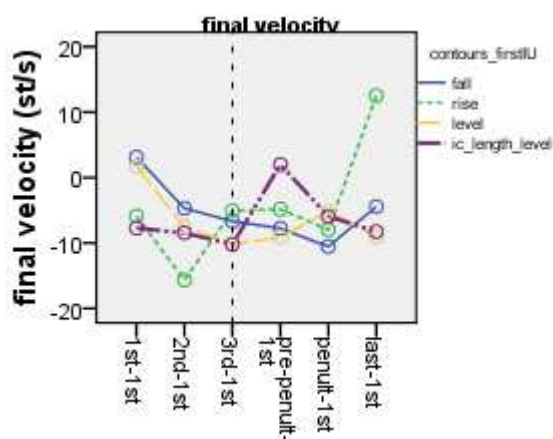


Figure 5-11 The values of final velocity in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.

The durations vary significantly ($F(3, 633) = 20.294, p = .000$), which is not unexpected with the extreme values of the lengthened IUs (Figure 5-12). When these are not taken into account, the differences in durations between falling, fall-rise and level contours are not significant. The values are indeed very close for the falling and fall-rise contours. Those in the level contours however have shorter durations at the left edge, and much longer ones at the right edge. The values of the last syllables for the falling, fall-rise and level contours are 186.55ms, 184.81ms, and 243.01ms, respectively. The long duration of the level contours may help in explaining the longer duration IU-finally than prosodic sentence-finally reported in the chapter on grouping (§3.3.3.1). Finally, the duration of lengthened syllables is on average 641.97ms, markedly longer than that of the previous syllables in the same IUs. Moreover, that average is only indicative, as the speakers may markedly lengthen the vowel, the longest duration in this dataset being 3713.93ms.

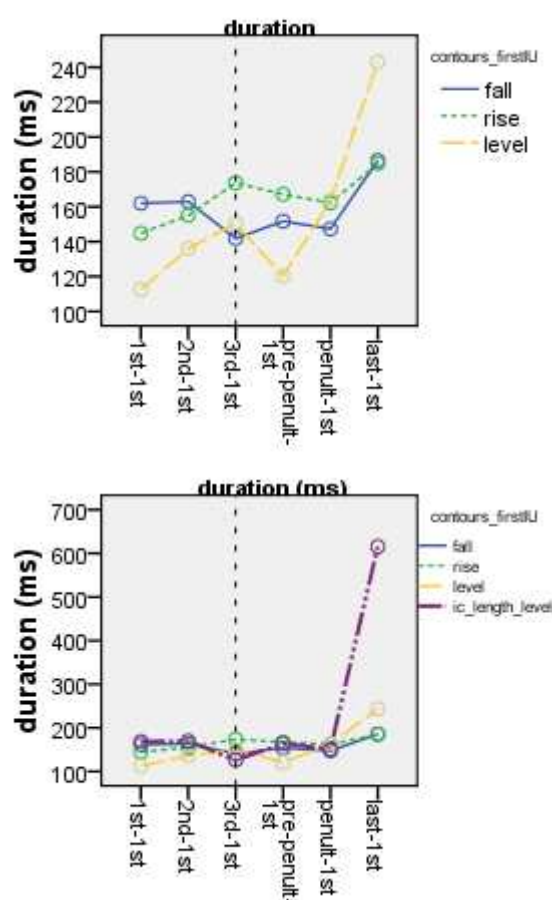


Figure 5-12 The values of duration in the first IUs with a fall, fall-rise, level, and lengthened contours. Only the first and last three syllables of the IUs are shown.

5.2.3.2.1 Summary: First IUs

The results of the measurements confirm that three contours are distinguished in the first IUs of complex sentences in Jaminjung: fall, fall-rise or level. The lengthened constructions also present specific characteristics.

In complex sentences, the most common contour in the first IU is the fall contour. Although the shape of the contour in the first IU cannot predict the type of IU in second position, a few general observations can be made as to their distribution in the complex sentences in this dataset. Most of the fall-rise contours precede IUs that are verbal clauses. There are also a few occurrences of fall-rise contours before reactivated topics or non-finite predicates, but none before afterthoughts. As to the level contours, they occur most frequently before verbal clauses, but a few cases are found preceding afterthoughts.

The falling contours have mean pitch values that, unsurprisingly, decline through the IU, ending in a final [low] underlying target. The movement in fall-rise contours is apparent in the last syllables of the IU, rarely earlier. The fall-rise contours have slightly higher mean F0 throughout and greater pitch excursions at their left boundary. They have similar patterns of final velocity as the falling contours, except evidently in their final syllables which have a [rise] target. It is difficult to assign a specific meaning to the contour, because its occurrence appears to be inconsistent. Rises are often considered as grammatical in the literature. In the AM framework, for example, their meaning “is to be interpreted with respect to a succeeding phrase “ (Pierrehumbert and Hirschberg 1990:305), or more generally as indicating that there is ‘more to come’. In the framework of discourse analysis, Selting (2000) has shown that rising intonation (or non-low in Dutch) is associated with turn-keeping phenomena. In Jaminjung, the interpretation of the rise as indicating ‘continuation’ is confirmed by the fact that the contour does not occur before afterthoughts. However, the rise which often occurs on the Kriol particle *na* cannot easily be interpreted as a ‘continuation rise’ as the particle indicates temporal succession relative to the preceding clause. This is

supported by the fact that the contour does not occur with the equivalent Jaminjung clitic =*biyang*. Obviously, more work needs to be done on this subject. At this stage, it can be said that the ‘continuation rise’ is a strategy available to Jaminjung speakers to indicate that their utterances are not yet complete.

The level contours are characterized by lower pitch values throughout and lesser pitch excursions. Their final syllables probably have [mid] underlying pitch targets and are marked by longer durations than those in both falling and fall-rise contours. As the contour is only attested in narrative, it may function as indicating a succession of events. This remains to be studied in further detail.

As to the IUs with iconic lengthening, they have steady pitch averages, [mid] underlying targets and, obviously, markedly longer final syllables. Concerning the first IUs, the comparison shows that the lengthened IUs are not a lengthened version of the level contours. At the left boundary, their average pitch is as high as that of the falling contours. The pitch is then maintained at the same level until the end of the IU. This finding confirms the impression of a ‘high’ level contour, as this ‘iconically lengthened’ construction has been referred to in other North-Australian languages such as Bininj Gun-wok (Fletcher and Bishop 2005:338).

5.2.3.3 Second IUs: Verbal clauses

In this section, the sentences with 2 verbal clauses will be examined. They will constitute, as well as the simple declaratives, a point of reference against which the other subtypes of IUs in second position will be tested in the subsequent section. Prosodic sentences composed of 2 verbal clauses can be schematically represented as:

1 st IU (falling)	without pause	2 nd IU (falling)
1 st IU (falling)	pause	2 nd IU (falling)
1 st IU (rising)	pause	2 nd IU (falling)
1 st IU (level)	pause	2 nd IU (falling)

These overall patterns provide the basis for the subgroups analysed in this section.

As verbal clauses may or may not be preceded by a pause, a distinction is maintained between the subtypes during this analysis. To avoid the risk of skewing the results because there are so few tokens in this sub-category, the tokens with a level contour are excluded from the analysis. Only the peripheral syllables, first, second, penultimate and last, are retained. In order to avoid cross-interferences, only the sentences with a falling contour in the first IUs are retained in the graphs illustrating the results for each of the correlates. IUs are separated by a vertical line.

The measures show significant differences in the mean F0 of the syllables when IUs are separated by a pause ($F(1, 126) = 55.855, p=.000$) (Figure 5-13). Interestingly, the average mean F0 of the entire sentence is lower when verbal clauses are separated by a pause than when they are contiguous. It would not be surprising if these lower values were observed in the second IUs only, considering that a pause might trigger an overall lowering of the pitch, yet the lower values of mean F0 are observed from the start of the sentences. The average mean F0 of the first IUs in complex sentences with a pause is less than 170Hz; it is mostly above 200Hz in the contiguous IUs. Importantly, there are no significant differences in the reset at the left boundary of the second IU, with an average value of 12Hz. The reset is obviously less apparent when the first IU ends in a rise.

The patterns are very similar in the second IUs, with a peak in the initial syllables and a steady decline in values until the final syllable.

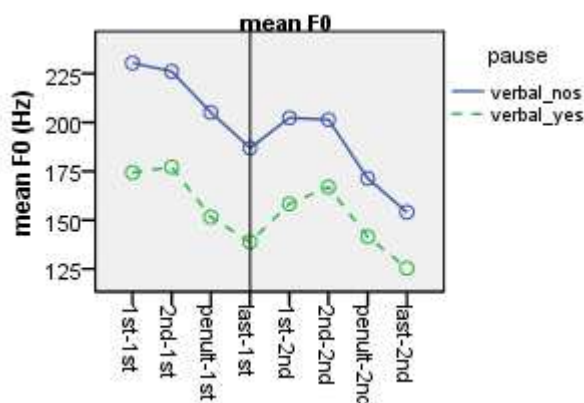


Figure 5-13 The mean F0 values in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.

The differences in excursion size are not significant between contiguous and IUs separated by a pause, except for that of the final syllables of IUs followed by a pause which is much wider than their counterparts, with respective values of 7.88st and 2.96st (Figure 5-14). The falling pitch movement in the final syllable in IUs that are followed by a pause is obviously much more pronounced than in those that are contiguous. The first and final syllables in the second IUs have wider excursions than the syllables in mid positions in the IUs, highlighting the sharper pitch movements at the IU periphery. The range of the excursions remains somewhat limited, the lowest value being 1.97st and highest 3.52st.

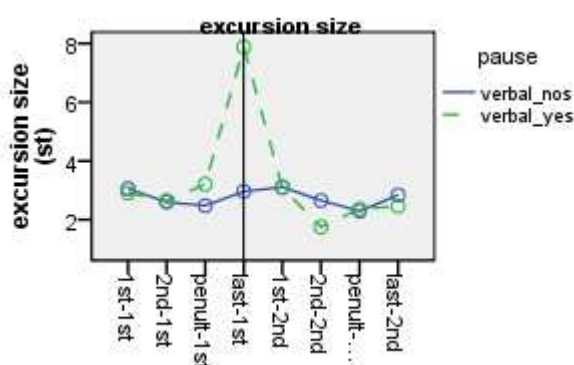


Figure 5-14 The values of excursion size in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.

The final velocity values do not show statistically significant differences depending on whether a pause is present or not (Figure 5-15). The patterns in the second IUs are not dissimilar to these observed in prosodic sentences formed of a single IU reported on in the previous section. The syllables at the left edge bear the marking of their information structure categories, those of the first IU with velocities associated with initial syllables of topics, and those of the second IUs associated with the marking of focus. This is plausible, as the topic is very likely to be omitted in the second IU.

The values of the syllables at the right edge have been discussed already for the different types of contours in the first IUs. In the second IUs, the penultimate syllables have the lowest final velocity values, averaging -17st/s, while those for the final syllables are around -8.0st/s. These measures suggest a final [low] target.

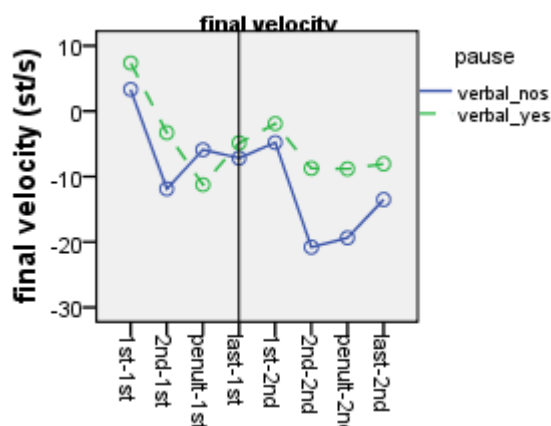


Figure 5-15 The values of final velocity in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.

The presence of a pause does not indicate statistically different values in durations, but, as with the excursion size, there is a great difference in the final syllables of the first IUs (Figure 5-16). Those followed by a pause have much longer durations, 231.35ms to 149.88ms, showing that IUs followed by a pause are evidently much more subject to final syllable lengthening. Another interesting difference is in the length of the first syllable in the second IU, which is much more salient in the contiguous IUs than in those that are separated by a pause. This may be explained by the fact that many of these contiguous IUs start with a monosyllabic coverb, often focused, which are bound to be longer than other bisyllabic words.

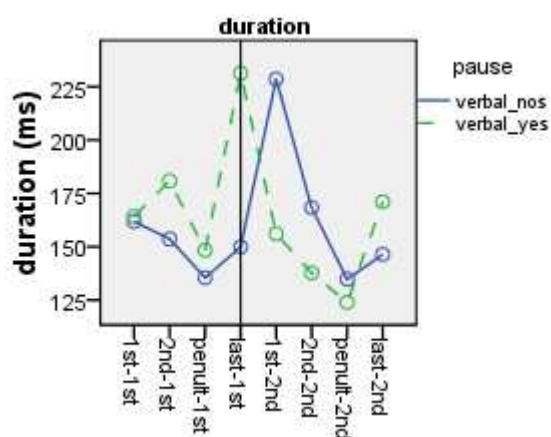


Figure 5-16 The values of duration in sentences consisting of 2 verbal clauses, distinguishing those that are separated by a pause and those that are contiguous.

5.2.3.3.1 Summary: 2 verbal clauses

To sum up, the measures in this section have shown that the presence of a pause influences the encoding of the IUs in syntactically complex sentences. Overall, the presence of a pause lowers the mean pitch of the entire sentence. However, the value of the pitch reset of the second IU is not influenced by the presence of a pause. Other tendencies are observed, notably that the final syllables in first IUs followed by a pause are longer and have a steeper falling contour than in those of contiguous IUs, which reflects their more independent status.

The second clauses/IUs are set off by a pitch reset on their first syllables. There is a pitch peak in the first two syllables after which the average pitch declines over the course of the IU⁵⁶. The velocity patterns are similar to those found in declarative sentences formed of a single IU reported on earlier. The values for the first syllables in this dataset suggest a [fall] pitch target, followed by a succession of [mid] targets. The left boundary is related to the marking of information structural categories, notably topic and focus. In chapter 4 on information structure, it was shown that a [fall] underlying target is associated with focus. It is not surprising to find such a target at the left boundary in the second IUs, given the propensity for omitting given topics in Jaminjung. Further, while the favoured strategy for focus marking is to move the focused element in IU-initial position, the focused element can also occur elsewhere in the IU, in which case the prominence is also displaced. When this occurs, the starting point of the falling contour is the first syllable of the element in focus. IUs introduced by particles are a case in point. Particles (§3.2.3.1.5) differ from clitics in that they are free forms and are rarely marked for prominence. In example (99) (Figure 5-17), the second IU is introduced by the particle *majani*, ‘maybe’. The focused element is the inflected verb, *burruriga* ‘they cooked it’. The falling movement starts at the beginning of the focused

⁵⁶ This does not imply that other contours are excluded; indeed speakers sometimes end prosodic sentences with a rising pitch contour. I would rather surmise that the falling contour is the preferred alternative.

inflected verb and not from the first syllable of the particle *majani*. The contour is representative of this type of construction with the main prominence falling on the first syllable of the second word in the IU. This accounts for some of the variations in the declarative contours found in the Jaminjung dataset.

(99)
 wirib=marlang=biyang yarrajgu ga-ngga...
 dog=given=now afraid 3sg-go.PRS

majani burru-rriga
 maybe 3pl:3sg-cook/burn.PST

'The dog is frightened, maybe they stung him.'
 [CP:ES96_18_02]

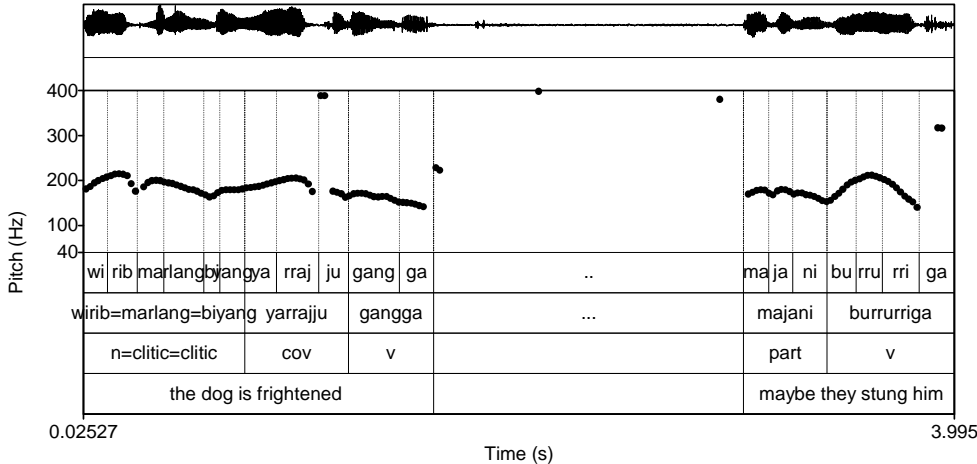


Figure 5-17 A sentence in which the second IU starts with the particle *majani*. The falling contour effectively starts at the beginning of the focused word *burruriga*.

Finally, at the right edge, the final syllable has a [low] target. The last syllable often receives creaky phonation, in which the speaker lets his/her voice drop to its lower reaches; there is also notable lengthening. Both of these are expected marking devices of major boundaries (see ch. 3).

5.2.3.4 Comparison of second IUs, verbal, non-verbal and NPs

In this section, the measurements of the correlates in second IUs in the four subtypes, verbal clauses, non-verbal clauses (coverbs as non-finite predicates), and NPs (afterthoughts or reactivated topics) are compared, with the aim of establishing whether they receive distinctive prosodic encodings.

First tests show that the contour of the first IU has no significant effect on the measurements of the second IUs; neither do the differences in speakers. All of the tokens will thus be included in the analysis. Secondly, as the verbal clauses are longer than all the other subtypes, it is interesting to assess whether the superficial differences in the contours are simply the effect of their greater number of syllables. The number of syllables in the IUs is thus considered. A comparison is made with all IUs in each of the subtypes, and then parallel comparisons are made with 2- and 4-syllable long IUs. However, with only two tokens of 4-syllables long non-finite predicates, the values in this sub-category cannot be treated as representative. To avoid interferences from other factors, only verbal clauses separated by a pause are retained and complex sentences with a level contour in the first IU are filtered out.

5.2.3.4.1 Mean F0

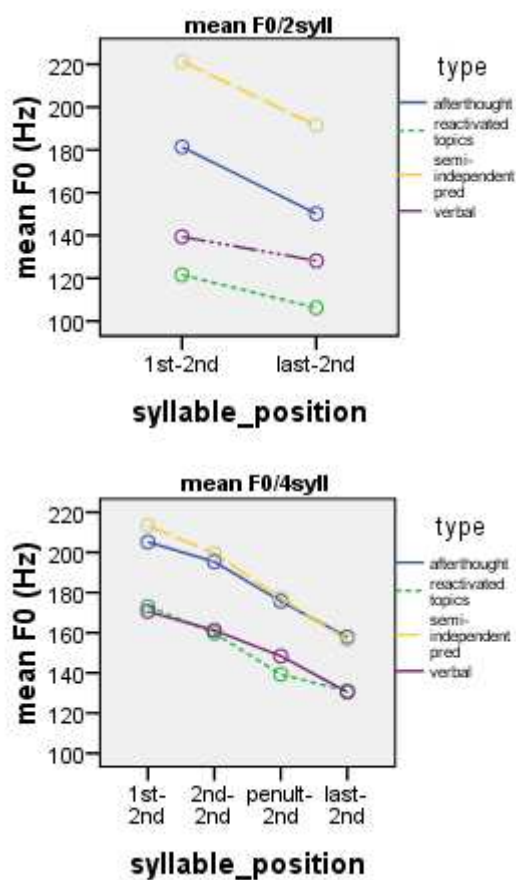
In all the various lengths of IUs tested, the mean F0 values in the four subtypes vary significantly (all: $F(3, 237) = 5.303$, $p = .002$, 2-syllable: $F(3, 33) = 16.521$, $p = .000$, 4-syllable: $F(3, 7) = 4.319$, $p = .008$). The values also differ significantly according to the position of the syllables in the IU, at least in the longer four-syllable IUs ($F(3, 70) = 4.115$, $p = .011$) and all IUs ($F(3, 237) = 20.195$, $p = .000$).

The results of the measurements are shown in Figure 5-18 and Table 5-3 below. When all IUs are considered, the non-finite predicates differ from all other subtypes. Verbal clauses and reactivated topics form a subgroup that differs from afterthoughts and non-finite clauses, a grouping that is much more easily observable in four-syllable IUs. The register of reactivated topics and verbal clauses is much lower throughout. Non-finite clauses are further demarcated by a higher F0 average in their first syllables. Reactivated topics, on the other hand, have lower pitch averages. The average pitch decreases over all the syllables in all subtypes.

In 2-syllable IUs, the values match those of the final 2 syllables in longer IUs, rather than those of their initial syllables, resulting in fairly low values. This pattern holds for the verbal clauses, afterthoughts and reactivated topics. In bisyllabic non-finite clauses, however, the syllables

have values that match those of the first and second syllables in the longer IUs, and are thus higher.

The F0 measurements are also useful in determining the pitch resets setting off the second IUs. Pitch reset occurs in all subtypes, as shown by the difference in pitch between the last syllables of the first IUs and the first syllables of the second IUs. The pitch reset in sentences with verbal clauses is 13.49Hz, in afterthoughts, 37.01Hz. It is 36Hz in reactivated topics and 30.18Hz non-finite clauses. The resets are thus much less marked in verbal clauses than in all other subtypes.



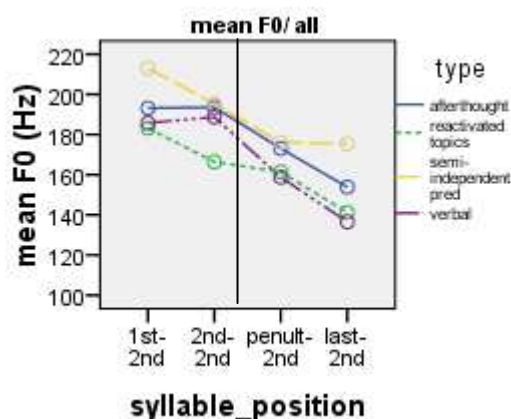


Figure 5-18 The mean F0 values in 2-, and 4-syllables second IUs, as well as the average of the values for the syllables at the left and right periphery in all IUS.

Syll position Nb syll in IU	Mean F0 (Hz)											
	verbal			afterthought			reactivated topics			non-finite clause		
	all	4syll	2syll	all	4syll	2syll	all	4syll	2syll	all	4syll	2syll
last syll 1st IU	172. 28			153. 97			141. 31			182. 82		
1st- second	185. 77	170. 54	139. 38	193. 15	205. 14	181. 36	183. 18	173. 10	121. 58	213. 00	213. 33	221. 33
second- second	188. 72	161. 22		193. 51	195. 34		166. 41	159. 73		195. 32	199. 45	
penult- second	158. 71	148. 35		173. 05	175. 67		161. 41	139. 29		176. 13	177. 29	
last- second	136. 58	130. 48	128. 16	153. 82	157. 87	150. 17	141. 28	131. 00	106. 18	175. 34	156. 34	191. 39

Table 5-3 The mean F0 values in 2- and 4-syllables second IUs, as well as the average of the values for the syllables at the left and right periphery in all IUs. The values of the final syllables of the previous IUs are also shown, as they are useful to calculate pitch reset.

5.2.3.4.2 Pitch excursion

The differences in pitch excursion size between the subtypes are not significant, except for the factor syllable position when all IUs are considered (all: $F(3, 240) = 5.423$, $p = .001$) (Figure 5-19). That is to say, the subtypes of IUs cannot be differentiated by a variation in pitch excursion. In two-syllable IUs, all subtypes end with excursions between 3 and 4st. except afterthoughts which have values close to 7st. The excursion size is greater in the final syllable than in the first syllable, indicating a steeper

movement at the right edge. Non-finite clauses are the exception, with a wider excursion in their first syllables. Patterns are harder to discern in four-syllable IUs and will not be discussed. When all IUs are used for the comparison, the patterns in excursion size become very easily apparent. All subtypes have similar pitch excursions in their initial syllables, between 2.8 and 3.8st. The size of the excursion then decreases slightly in the second syllables. At the right periphery, all subtypes have a wider excursion in the final syllable than in the penultimate syllable, except the afterthoughts which have larger excursions in the penultimate syllables. Non-finite predicates have a markedly wider excursion in their final syllables. The wider spans at the boundaries reflect the special status of these syllables in the encoding schemes of both the grouping function (boundary marking) and the emphasis marking function (first syllables locus of focus encodings).

The measurements suggest a steeper falling movement in the final syllables of non-finite predicates and afterthoughts; in the latter case, the fall is initiated in the penultimate syllable.

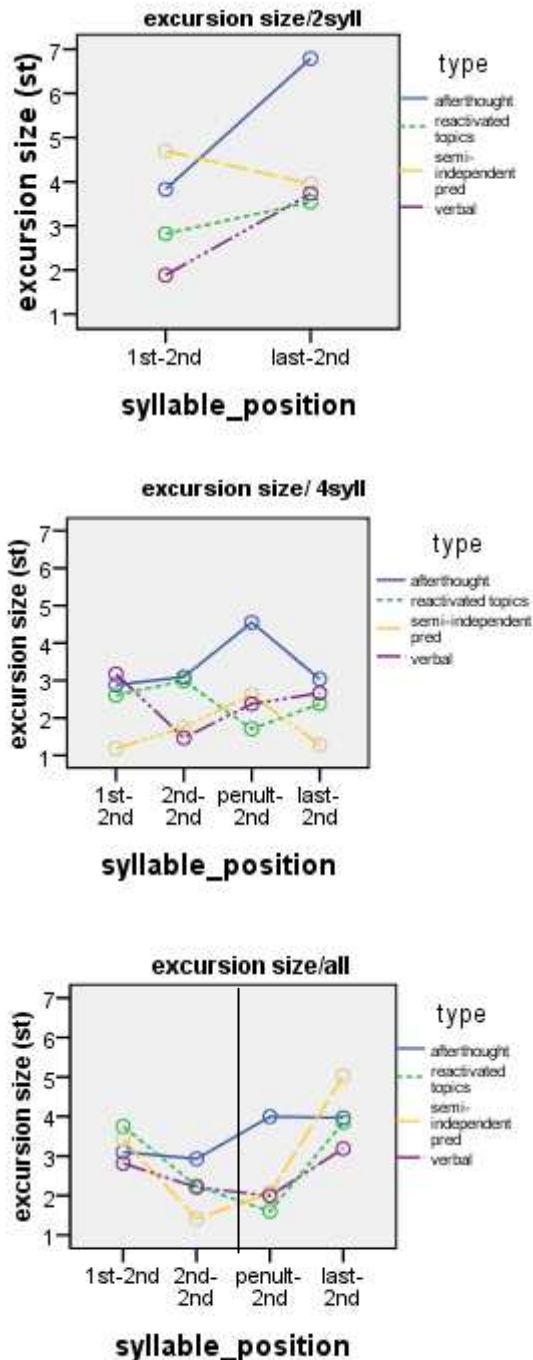


Figure 5-19 A comparison of the excursion size of syllables at the boundaries of in the second IUs, shown for IUs of various length.

5.2.3.4.3 Final velocity

Even if the differences in final velocity measurements between the subtypes of IUs do not reach statistical significance, it is interesting to note that non-finite clauses are the furthest away from the other types of IUs. The differences in the values according to syllable position are significant for 4-

syllable IUs and with all IUs (4-syllable: $F(3, 67) = 2.786$, $p = .005$; all: $F(3, 224) = 3.950$, $p = .009$) (Figure 5-20 and Table 5-4).

The lack of data for the final syllables of 2-syllable verbal clauses, probably due to creaky phonation, makes the interpretation of shorter IUs difficult. For this correlate, the four-syllable IUs seem to be most telling. At the left edge, the reactivated topics have values around 10st/s in the first syllable followed by a dip in the second syllables, which is consistent with the encoding of topics with an initial [high] target of reported earlier (§4.5.6). Afterthoughts and verbal clauses have very similar values, with values close to zero and falling around -10st/s in the second syllables, again consistent with their possibly, but not necessarily, harbouring a focused first syllable receiving a [fall] target. These subtypes could also have topics in initial position, which have higher final velocity values, and thus impact on average values. The first syllables of non-finite predicates have very low values, around -15st/s, the [fall] target associated with a focused element.

At the right edge, the pattern in all the IUs is that of a general rise in the final syllables. In all cases, the suggested underlying target is a [low]. Non-finite predicates have lower negative values. In four-syllable IUs, the same low values are also observed in afterthoughts. Thus, while these values do not necessarily suggest a different underlying target, it is suggested that in these IUs the movement is sharper than in the other subtypes.

For this correlate as with the mean F0, the values of the bisyllabic IUs seem to mimic those in the penultimate and final syllables of longer IUs, the evidence being somewhat less compelling than with the mean pitch. The patterns of underlying targets in the second IUs suggested by the measurements are thus:

	1 st syll	subsequent syll		penult	final syll
verbal clauses	[fall]	[mid]	[mid]	[mid]	[low]
non-verbal clauses (SIP)	[fall]	[mid]	[mid]	[mid]	[low]
NPs afterthoughts	[fall]	[mid]	[mid]	[mid]	[low]
NPs reactivated topics	[high]	[mid]	[mid]	[mid]	[low]

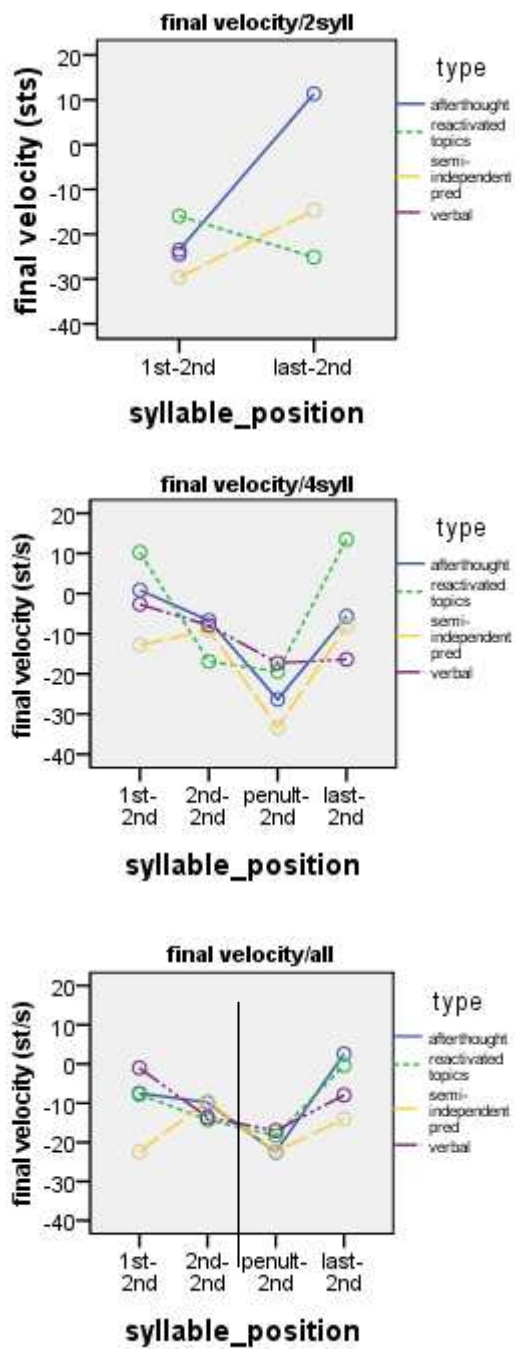


Figure 5-20 A comparison of the final velocity values in the second IUs, shown for IUs of various length.

Syll position Nb syll in IU	Final velocity (st/s)											
	verbal			afterthought			reactivated topics			Non-finite clause		
	4syll	all	2syll	4syll	all	2syll	4syll	all	2syll	4syll	all	2syll
1st-second	-2.67	-1.98	-24.44	0.83	-8.56	-23.51	10.22	-7.58	-15.94	-12.84	-22.41	-29.62
second-second	-7.91	-14.22		-6.70	-8.78		-16.92	-14.27		-9.06	-9.52	
penult-second	-17.31	-15.40		-26.36	-21.00		-19.41	-18.30		-33.28	-22.33	
last-second	-16.38	-9.90		-5.65	2.80	11.35	13.47	-0.35	-25.14	-8.11	-14.04	-14.63

Table 5-4 The final velocity values of in the second IUs, shown for IUs of various length.

5.2.3.4.4 Duration

The differences in duration are not statistically significant, neither for the subtypes nor according to syllable positions. The results are shown in Figure 5-21, Table 5-5 and Table 5-6.

At the left boundary, the first syllable is longer than the second in non-finite predicates and verbal clauses, while it is shorter in afterthoughts and reactivated topics.

As to the right boundary, the patterns are surprisingly similar, except for the afterthoughts seem to be longer overall. Final lengthening is apparent in all subtypes. It is calculated as a ratio of the duration of the final to the penultimate syllable; the results are shown in Table 5-6. Final syllable lengthening is not apparent in two-syllable IUs, except in verbal clauses. Otherwise, final syllables are lengthened in all subtypes. The bottom right graph in Figure 5-21 shows the average length of the pause preceding the second IUs, and shows clearly that the pause is much shorter before reactivated topics, thus signaling its closer syntactic integration with the previous clause.

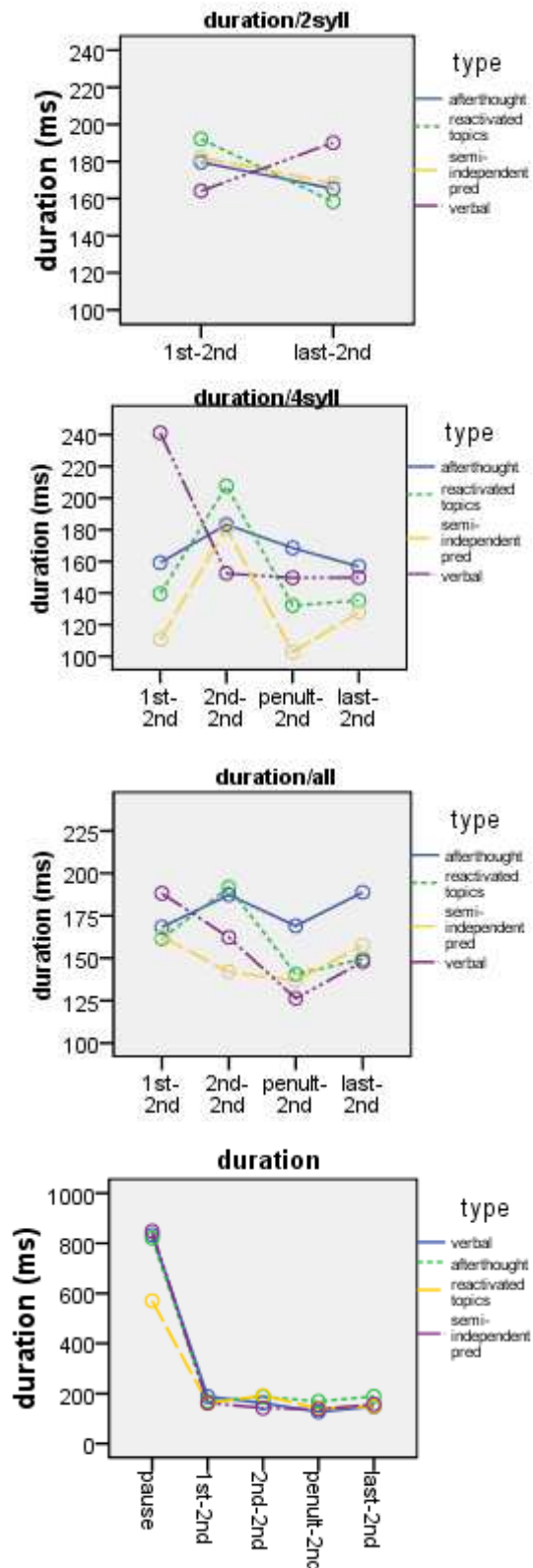


Figure 5-21 The duration of the syllables in the second IUs, shown according to IU length. The bottom right graph shows the overall durations, including that of the preceding pause.

Syll position Nb syll in IU	Duration (ms)											
	verbal			afterthought			reactivated topics			non-finite clause		
	4syll	all	2syll	4syll	all	2syll	4syll	all	2syll	4syll	all	2syll
1st-second	241.14	188.22	164.14	159.14	168.31	179.63	139.54	161.53	192.14	110.65	162.64	182.16
second-second	152.40	162.23		183.21	187.29		207.58	191.79		181.00	141.93	
penult-second	149.62	126.31		168.47	169.01		132.14	140.55		102.66	136.94	
last-second	149.82	147.93	190.16	156.77	188.84	165.36	135.32	149.50	158.29	127.81	157.55	168.27

Table 5-5 The duration of the syllables in the second IUs, shown according to IU length.

Ratio durations final/penultimate syll	4syll	all	2syll
Verbal clause	1.01	1.17	1.16
Afterthought	.93	1.12	.92
Reactivated topic	1.02	1.06	.82
Non-finite predicate	1.51	1.16	.96

Table 5-6 The duration of the syllables in the second IUs, shown according to IU length.

Before discussing these results and to complete the portrait of complex sentences, a further test is conducted, this time comparing the second IUs of lengthened constructions to those of non-lengthened sentences.

5.2.3.4.5 Comparing second IUs that follow lengthened first IUs to other second IUs in complex sentences

Of particular interest is a comparison between the second IUs in lengthened and non-lengthened sentences. It has been proposed to treat the sequence of a lengthened IU and the subsequent IU as a single construction, defined by a specific contour: “the final syllable of the main clause is lengthened, often considerably, with level intonation, and the delimiting predicate has a sharply falling boundary intonation” (Schultze-Berndt 2000:24). The alternative is to consider the meaning of the IUs compositionally, the second IU simply receiving the encoding of ‘secondary prosodic IUs’ according to their semantic relationship with the main clause.

The dataset for this analysis contains the prosodic sentences with iconic lengthening and the complex declarative sentences used in the previous sections, omitting the reactivated topics.

The measures of mean F0 indicate significant differences between the subgroups ($F(5, 206) = 5.110, p = .000$). Post hoc tests indicate that verbal clauses after non-lengthened IUs differ from all other subtypes, but that these do not vary between each other. They are marked by lower averages throughout (Figure 5-22).

When they are verbal clauses, or NP afterthoughts, the second IUs of lengthened constructions have slightly higher overall mean F0 values. However, when they are non-finite predicates, they are lower, at least at the left boundary. The non-finite predicates also seem to differ in having a sharper fall at the left boundary, and a less marked fall in the right boundary, a pattern that is the reverse of that of the other two subtypes.

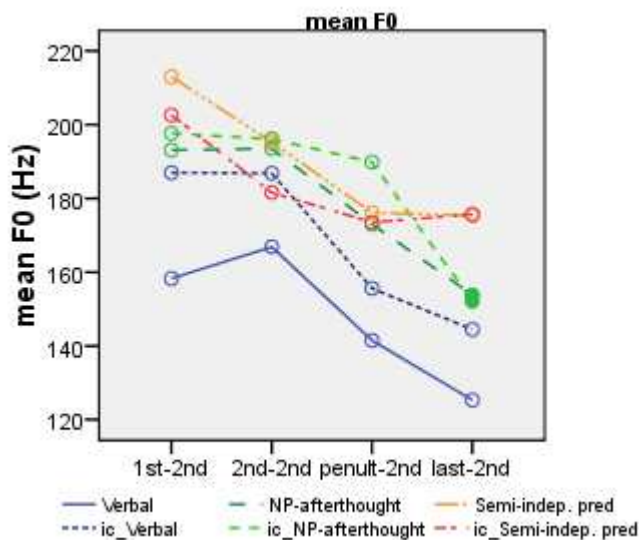


Figure 5-22 A comparison of the mean F0 values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.

The excursion sizes do not vary significantly. The non-finite predicates have a very different pattern of values in the last two syllables, with smaller excursions in the penultimate syllables and notably wider excursions in the final syllables.

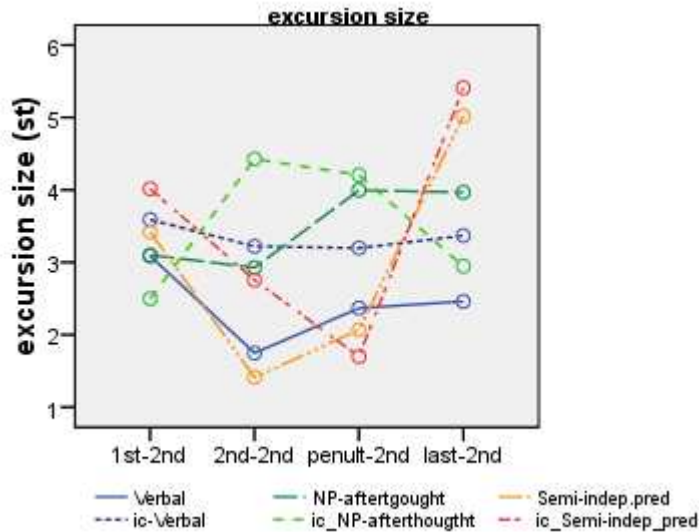


Figure 5-23 A comparison of the excursion size values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.

The measurements of final velocity do not indicate a significant difference between the subgroups tested. However, the patterning suggested by the mean F0 values is also apparent in the final velocity. Figure 5-24 illustrates these patterns. Verbal clauses have slightly higher values than other types of IUs in the first syllables. The values for the other subgroups vary from 0 to -10st/s. At the right boundary, the IUs have a similar pattern: a lower value in the penultimate syllables, followed by a rise in the final syllable. The suggested underlying target is a [low].

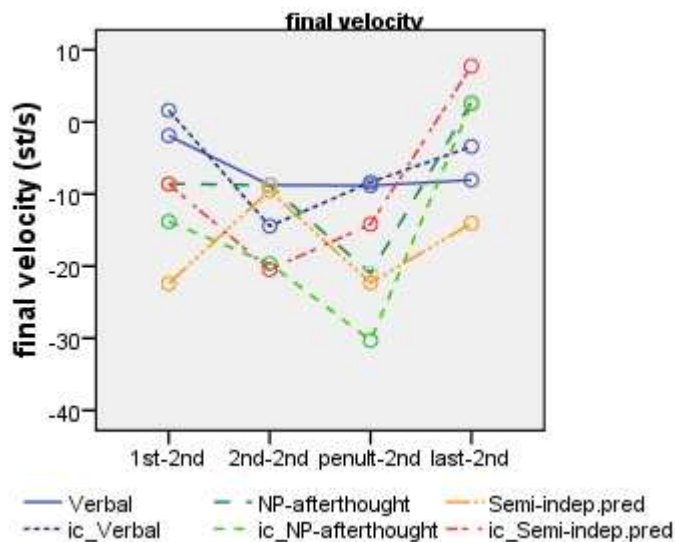


Figure 5-24 A comparison of the final velocity values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.

The durations do not differ significantly either, the most striking pattern being the longer values of the final syllables.

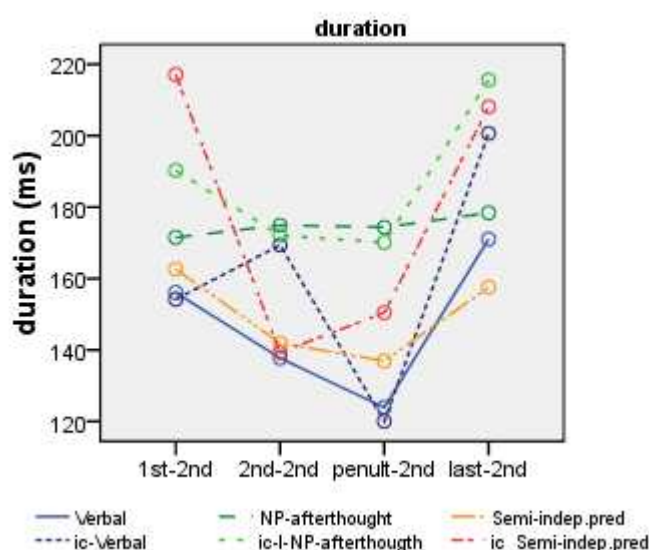


Figure 5-25 A comparison of the duration values of the various subtypes of IUs in second position, according to whether they follow a lengthened IU or not.

This comparison of the correlates in second IUs shows that there are no significant differences in the correlates of non-finite predicates and NPs/afterthoughts, whether they are uttered after a first IU that is lengthened or not. However, verbal clauses following non-lengthened first IUs differ from all other types of IUs in second positions by having a lower average F0 overall. The implications of these results are discussed in §5.2.4 below.

5.2.4 Discussion: complex declarative sentences and iconic lengthening

Three main hypotheses were posited. The first one is that structural differences at the syntactic level should translate into differences in the prosodic encoding. Secondly, the informational value of the IUs determines their encoding, encoding ‘new’ and ‘given’ information distinctly. Thirdly, the patterning of the sequences of IUs in complex sentences is ‘compositional’, thus each IU receives a specific encoding according to its function. These hypotheses were tested by establishing the specific correlates associated with the IUs in complex sentences and then comparing the results in the various subgroups posited using statistical analysis.

In sentences composed of two IUs, the first IU is usually the main clause, with a falling contour resembling the ones described earlier for the simple declaratives in §5.2.2. First IUs can also have fall-rise and level contours. Interestingly, it is found that IUs with iconic lengthening are not simply elongated level contours. They are rather similar to falling contours at the start, with a steady level pitch from the lengthened syllables onward, resulting in a contour better described as a ‘high sustained level contour’.

As to the second IUs, when it is an independent or a subordinate verbal clause, its pattern is a repeat of the previous IU, the ‘default declarative’. Indeed, the sequence of underlying targets in the second verbal clauses is similar to those of the first IUs, with the syllables at the left boundary displaying encodings associated with focus marking. The final pitch movement at the right boundary is realized with the [low] target. Interestingly, it is found that pauses influence the average pitch of the whole sentence: the average mean F0 of IUs separated by a pause is lower. After the analysis of verbal clauses, a comparison was undertaken with other types of IUs as second IUs, including non-verbal clauses, namely non-finite clauses (non-finite predicates) usually consisting of a lone coverb; and NPs, functioning either as afterthoughts or reactivated topics, with the aim of establishing the prosodic correlates of the IUs. The statistical analysis confirms that the tested subtypes of IUs are distinguished by pitch registers. When uttering non-finite clauses and afterthoughts, speakers make use of a higher pitch register, while verbal clauses and reactivated topics are uttered in a lower register. I suggest that higher registers mark secondary prosodic units, positioned at the right edge of a prosodic sentence, which functionally bring in new information related to the main clause. Non-finite clauses and afterthoughts both contribute new information directly related either to an argument or to the predicate of the main clause. The most marked is the non-finite clause which not only has high overall pitch, but is also marked with values of final velocity and excursion size that suggest that even if it does not have a different underlying target on its final syllable, the [low] target is realized more sharply. Verbal clauses evidently contribute new information as well, but are new at the discourse level, rather than at the sentence level. Reactivated topics do not contribute new information,

instead; the information they carry is being made accessible, reactivated from a pool of ‘semi-active’ information that is “present in a person’s peripheral consciousness” (Chafe 1987:25). De Vries (2009:236), in a study of Dutch dislocations, comes to similar conclusions, which he expresses in terms of ‘backgrounding’, where the right-dislocated topic is “deaccented, whereas an afterthought, like many parentheticals, receives an additional intonation contour containing a pitch accent”.

Structurally, the distinction involves one construction being ‘dependent’ on another (the main clause), as in the case of reactivated topics, non-finite predicates and afterthoughts, or being ‘independent’ in the sense that it can function on its own, as in the case of verbal clauses. A dependent clause cannot function on its own; it needs semantic and/or pragmatic information from another clause in order to be understood. Verbal clauses (in second position) are relatively independent of the main clause, more semantically ‘self-contained’ and are therefore encoded with the default declarative contour.

The original hypothesis that the syntactic integration of the IUs is encoded in their prosody is clearly evidenced in the much shorter pauses that precede reactivated topics. As to the contours of the IUs, independent and subordinate verbal clauses have the ‘default’ declarative encoding. The measurements also suggest a secondary prosodic contour, which rather reflects semantic-pragmatic (information structural) considerations, namely that new information is conveyed with more salient prosodic encodings than old information. It seems that speakers intentionally create a system of prosodic units to emphasize relationships between semantically related sections of the discourse and to highlight some elements they judge important.

	Structural integration dependent ↔ independent	Information status new ↔ given
verbal clauses	+	+
semi-independent predicates	+	+
afterthoughts	+	+
reactivated topics	+	+

Table 5-7 The parameters ruling the prosodic encodings of second IUs in complex sentences.

The following examples summarize and illustrate my findings for the IUs in complex sentences.

Example (100) is a complex sentence consisting of two verbal clauses. The second verbal clause has an encoding similar to the first, simply continuing its declination. The example is from the same Emu-Brolga narrative as above. The participants have both been mentioned before and are omitted as sentence topics, being part of the common ground. Both IUs have the default declarative contour. The first syllables have higher final velocity values suggesting a [high] target, followed by [mid] pitch targets in subsequent syllables. The final syllables have a [low] target and are lengthened. In the second IU, which consist of the lone inflected verb *bunyingayija* ‘the two argued’, the prominence is heard on the first syllable, but the highest pitch is realized on the second syllable. As is very common, the sentence final syllable ends with creaky phonation.

(100)
 jarlaj gani-yu=nu... buny-ngayija
 jealous 3sg:3sg-do.PST=3sg.OBL 3du-see-REFL:PST

'She got jealous of her ... the two argued'
 [DM:MH96_19_01]

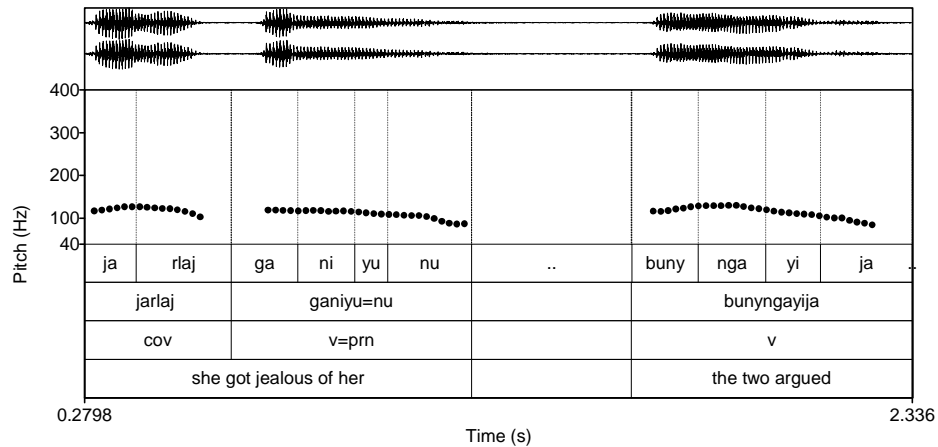


Figure 5-26 A sentence showing two independent verbal clauses.

Example (101), like (100), is part of a mythological narrative. The participants, the Emu and the Brolga⁵⁷, have both been mentioned before and are part of the common ground. The second IU consists of the NP *gudarrg=marlang* which refers to the brolga, with the clitic *=marlang*, the information structural marker of givenness (§4.5.4.2). It displays the characteristics of reactivated topics: an overall gently falling contour, an average pitch using a lower register, no salient focus marking on the first syllable, and no extensive lengthening of the final syllable.

⁵⁷ Australian birds.

(101)

Lagarniny=biyang ga-yu... gudarrg=marlang
 placename=now 3sg-be.PRS broлга=GIVEN

'At L. she is now, the Broлга is.'

[DM:MH96_19_01]

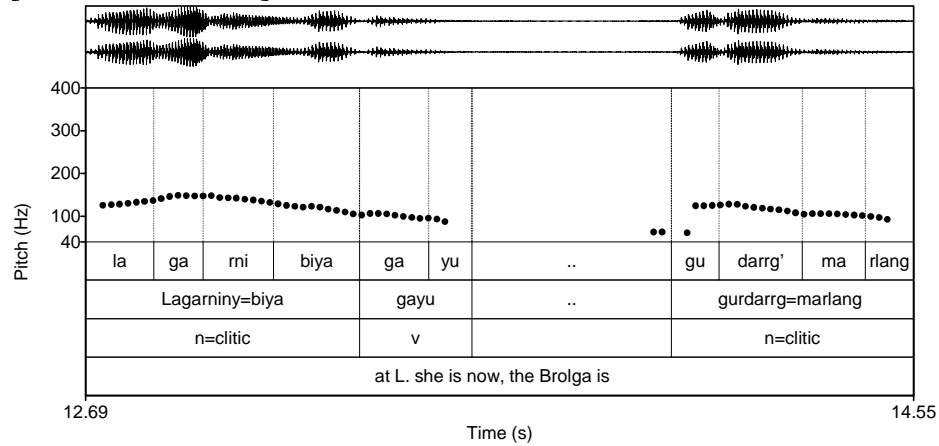


Figure 5-27 A prosodic sentence in which the second IU is NP, functioning as a reactivated topic.

Afterthoughts and non-finite predicates are grouped together as secondary prosodic IUs, with very similar encoding patterns. They both make use of an elevated pitch register. The non-finite predicates distinguish themselves in their low final velocity values in the first and final syllables, which suggest that they receive [fall] targets. This special encoding, especially of the first syllables, is congruent with the findings of the encoding of focus described earlier (see ch 4). The perceptual impression in both cases is of a sharp falling contour.

Example (102) is from a retelling of the 'Frog Story'. It shows an NP functioning as an afterthought, *wirib-ni-mij*, 'with the dog'. A circumstantial predicate complement such as this is not likely to be topical. The sentence topic in this case is *jarlig* 'the boy' in the first IU. All the participants in the story have been introduced. Here, the afterthought clarifies the description given by the speaker in the first clause, it has overall falling contour, with high average pitch. The first syllable is more prominent, with the associated pitch peak.

(102) (also (26))

jarlig janju-ni=marlang ganurru-ngawu
 child DEM-ERG=GIVEN 3sg:3pl-see.PST

warrb-bina... wirib-ni-mij
 be.together-ALL dog-ERG-COMIT

'That child saw them sitting together, together with the dog.'

[CP:ES96_18_02]

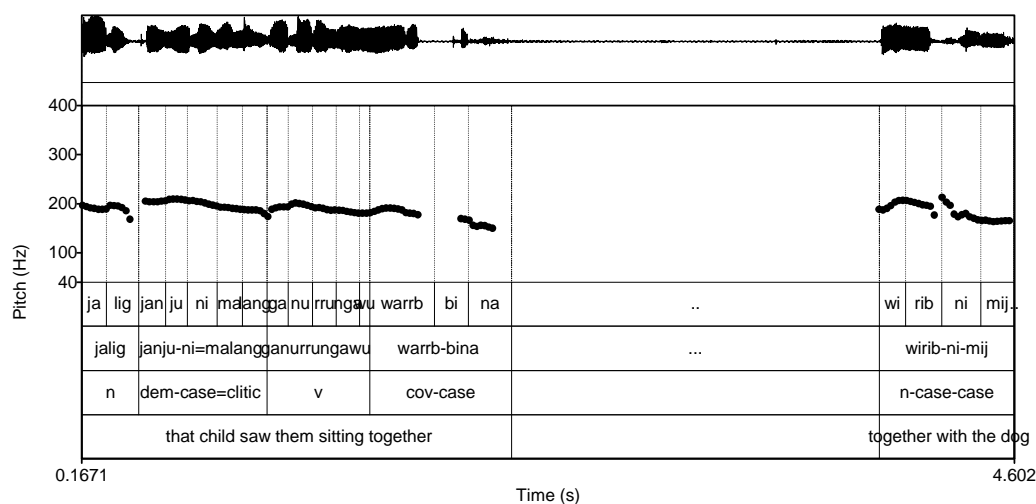


Figure 5-28 A prosodic sentence in which the second IU is an NP functioning as an afterthought.

Example (103) shows a non-finite predicate, from an elicitation session on verb forms. The speaker is illustrating her explanations with a sort of 'role playing' about people coming to ask her for food. The second IU, consisting of the coverb and its case marker *thawaya-wu* 'for eating' has a steep falling contour, higher average pitch, and a [fall] target on its first syllables, followed by a [fall] target at the right boundary. The final syllable is lengthened.

(103)

ngarrgina mangarra wurlug... thawaya-wu
 1SG:POSS plant_food alone eating-DAT

'That's the food for me alone / that food is only mine, to eat.'

[IP:ES96_08_02]

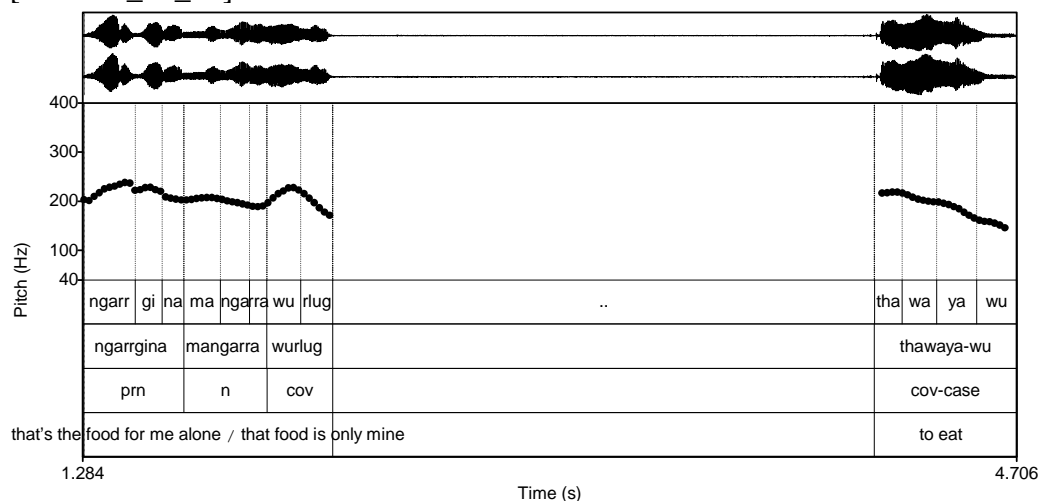


Figure 5-29 A prosodic sentence showing a non-verbal clause in the second IU, functioning as a non-finite predicate.

The final test was aimed at describing the constructions that mark the duration of an event in their first IUs through the lengthening of the final vowel. It is very common for these IUs to be followed by an IU consisting of a non-verbal clause (coverb as non-finite predicate), an NP (afterthought), and at times a verbal clause (usually an inflected verb only).

The second IUs have correlates that are very similar to those of the non-verbal clauses and afterthoughts already described. Indeed, the comparison between all the second IUs, whether they follow a lengthened IU or not, shows no significant difference in their encoding.

I suggest that the second IUs in Jaminjung complex sentences (including the lengthened constructions) can take one of two contours. The first option is the 'default' declarative contour; the second is the 'secondary prosodic IU contour'. Prosodically, they are distinguished mainly by a difference in pitch register, the 'secondary prosodic contour' being uttered in a much higher pitch. The declarative contour is used mostly with independent or subordinate verbal clauses, while the secondary prosodic contour is associated with coverbs/non-finite predicates and NPs/afterthoughts. The verbal clauses that sometimes follow an iconically

lengthened first IU, usually consisting of a lone inflected verb, are also encoded in this way.

The specific function of IUs encoded with the secondary prosodic contour is to contribute new information related to one of the elements in the main clause, either delimiting the event of the main predicate in time and space, or specifying one of the referents. Their being called ‘secondary’ highlights their dependence on the main clause, both morpho-syntactically and prosodically; such contours cannot occur on their own. They also have a demarcative function as they are obligatorily positioned at the right edge of a prosodic sentence. The higher pitch and the steep contour⁵⁸ which is the consequence of the secondary prosodic contour being shorter, have a stylistically marked effect in Jaminjung. They undoubtedly contribute to the vividness of the narration or description.

Verbal clauses evidently contribute new information as well, but it is new at the discourse level, rather than at the sentence level. Reactivated topics do not contribute new information, instead, the information they carry is being made accessible, reactivated from a pool of ‘semi-active’ information.

I further argue that the meaning of the encoding schemes of the IUs in complex sentences is to be treated as compositional⁵⁹, in that each IU contributes its meaning to the sentence as a whole. This is supported by the comparison between the lengthened constructions and other complex sentences reported above. First IUs receive different encodings to mark their specific function. The marking of durative aspect in Jaminjung is encoded in a specific prosodic scheme displaying iconic lengthening which differs from that of declarative statements. As to the second IUs, the results of the comparison confirm that the secondary prosodic contours are found both

⁵⁸ And probably the greater intensity as well, although this correlate can not be taken into account in this analysis.

⁵⁹ The use of ‘compositional meaning’ here is not to be equated with the same term used by Pierrehumbert and Hirshberg (1990), who state that the meaning of an intonational contour can be compositionally derived from its components, namely its pitch accents, phrase accents and boundary tones.

after iconically lengthened first IUs and non-lengthened declarative statements.

5.3 Interrogatives

The goal of this section is to describe the encoding of interrogatives in Jaminjung. Three subsets of interrogatives are tested, question word interrogatives, yes-no interrogatives and question-tag interrogatives. In the first type, a speaker requests information, in the second he/she inquires about the truth of a proposition (Cruttenden 1997:161), and in the third, he/she checks that the proposition is really in the common ground with the hearer.

The contours of polar and information questions are examined through a comparison of their correlates. Tag questions are treated separately.

5.3.1 Hypotheses

Dixon (2002:327) states that polar questions in Australian languages are often signaled by rising intonation. He also states, regarding information questions, that the interpretation of some pronouns as either interrogative or indefinite rests mainly on intonation, without further description of the patterns involved. In Jaminjung, based on informal observations, the prosodic distinction between polar and information questions is not evident. I will state as a starting hypothesis that they receive the same encoding, and that this encoding is similar to that of declarative sentences. Even if both rising, fall-rise and rising contours are attested in both types of interrogatives, polar questions are probably more likely to conform to the generalization of a rising intonation. Following this, tags issued after statements i.e. *ngih*, *init*, glossed as ‘isn’t it’, which act as a check on the common ground, and *gurra*, ‘right’, which seeks agreement with a previous assertion in the discourse, should be considered as polar interrogative forms and thus receive similar encodings.

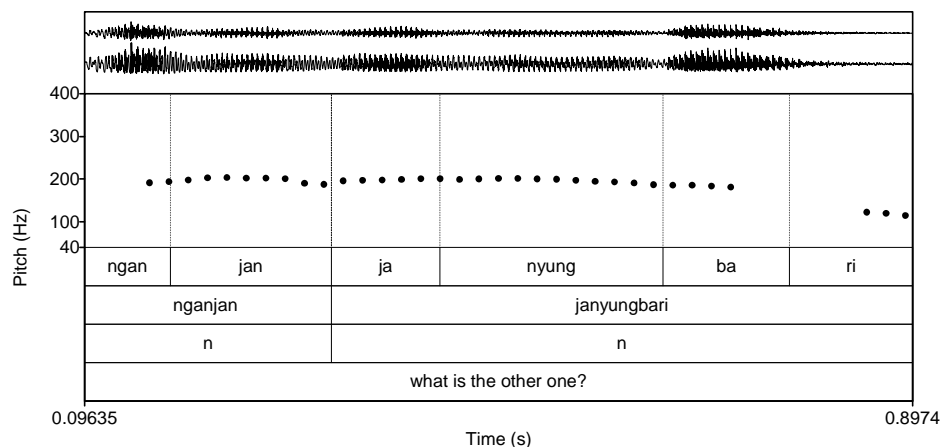
5.3.2 Dataset

The analysis of information questions is based on 58 tokens spoken by 7 different speakers. These questions usually contain one of the following nominal interrogatives: *nanggayin* ‘who/someone’, *nganthan* ‘what/something’ and its variants *nganthanug* ‘why’ (reason, objective), *nganthan-ngunya* ‘why’ (reason/cause, lit/ ‘what from’, see example (104) below), *ngajang* ‘how many’, *warnang* ‘where’, and *nyangulang* ‘when’. Jaminjung also has an interrogative coverb, *warndug* ‘how, do what’. The clitic =*warra*, glossed as ‘DOUBT’, is also associated with interrogatives; it conveys ignorance about the intended referent (‘I don’t know wh’).

The question word usually receives the encoding associated with a focused argument (§4.3). Information questions mostly have a falling contour (more on contours hereafter), but a rising contour can also occur, where the rising movement starts with the focused question word and continues until the end of the IU, as well as a contour called the fall-rise, similar to that observed in declaratives, where the rising movement occurs in the final syllables of the IU. Examples are shown in Figure 5-30 below. An example of a falling contour is given in example (104) where the clause is non-verbal. A rising contour is shown in example (105) and a fall-rise contour in example (106).

(104)
 nganjan janyungbari
 what other-QUAL

‘What is the other one?’
 [BH:CS07_72_01]

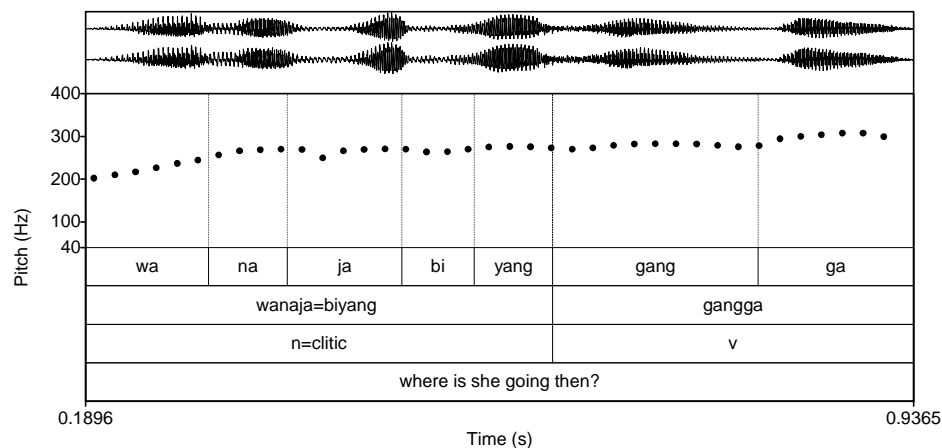


(105)

wanaja=biyang ga-ngga
where:DIR=now 3sg-go.PRS

'Where is she going then?'

[IP:ES96_08_03]



(106)

nganjan-nyunga yinju-nud
what-ORIG PROX-COLL2

'Where is this lot coming from?'

[JM:CS07_73_01]

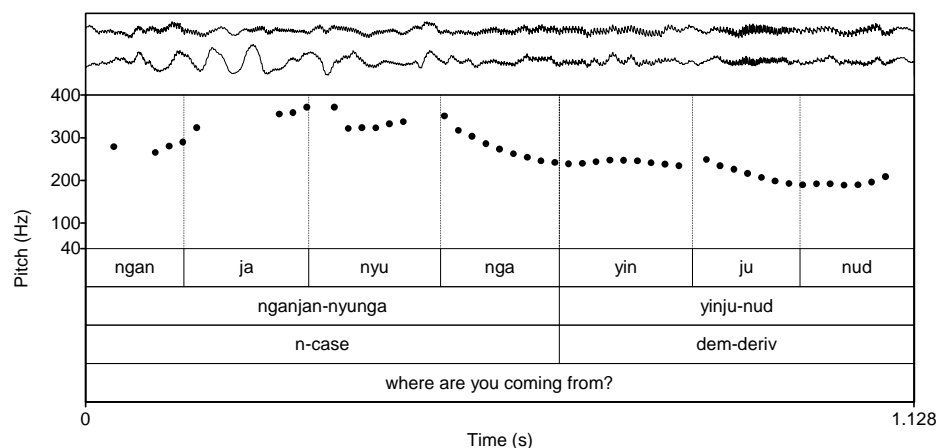


Figure 5-30 Information questions, displaying a falling, a rising, and a fall-rise contour.

The dataset contains 22 polar questions. Polar questions have the same structure as declarative sentences in Jaminjung. The clitic =*ja*, glossed as 'QU' may be used as a polar interrogative marker, but it is not frequent. Given the relative similarity of declaratives and polar questions, the

following properties were used to identify polar questions, originally developed for English and French (Safarova et al. 2005:3): they must be turn-final, and be followed by a reply from the addressee that contextually entails ‘yes’, ‘no’, or ‘I don’t know’.

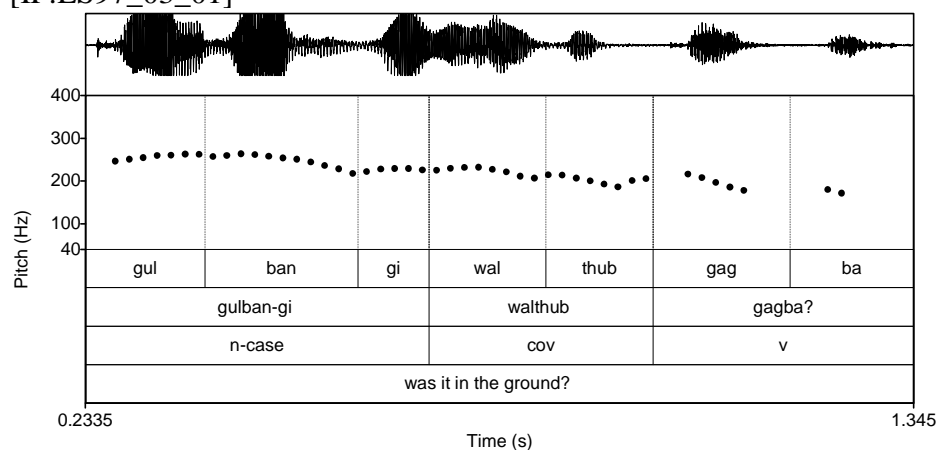
Figure 5-31 shows an example of a polar question with a falling contour in (107) and one with a rising contour in (108). The interpretation of the IU as a polar question rests on contextual cues.

(107)

gulban-gi walhub ga-gba
ground-LOC inside/enclosed 3sg-be.PST

'Was it in the ground?

[IP:ES97_03_01]



(108)

digirrij ga-jga-ny... yawayi
die 3sg-go.PST yes

'Did he die? Yes.'

[BH:CS07_72_01]

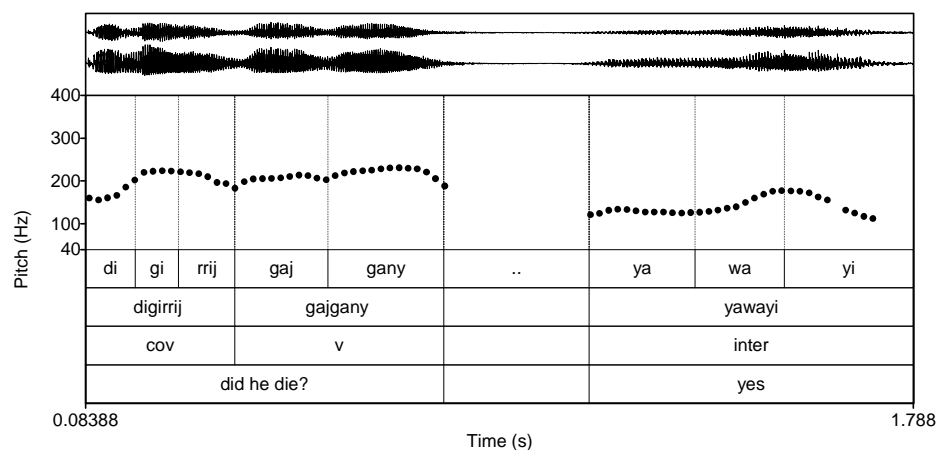
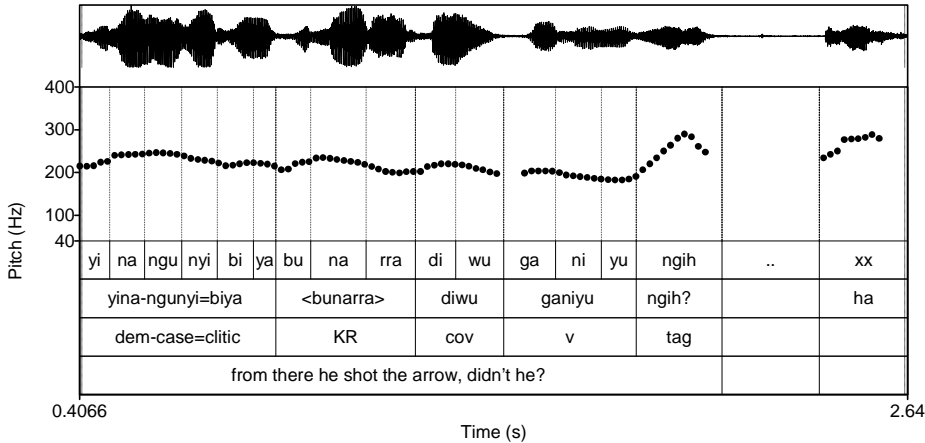


Figure 5-31 Polar questions with a falling and a rising contour.

Finally, the dataset contains 16 tag questions, which consist of a main clause followed by the interjection *ngih*, or its Kriol equivalent *yintit*, both glossed as ‘TAG’ and loosely translated as ‘isn’t it?’ The other construction analysed is formed with the clitic =*gurra*, restricted to post-verbal positions, and glossed as ‘EMPH’, with a meaning similar to the English tag ‘right?’ An example of each of the tag constructions (108-109) is given in Figure 5-32.

(109)
yina-ngunyi=biyang <bunarra> diwu
DIST-ABL=now bow.and.arrow fly

gani-yu ngih
3sg:3sg-say/do.PST TAG
'From there he shot the arrow, didn't he?
[IP:ES97_03_05]



(110)

wirib=mang gugu-g ga-yu gurra
 dog=SUBORD water-LOC 3sg-be.PRS TAG

'The dog is in the water, right?'

[IP:ES97_03_01]

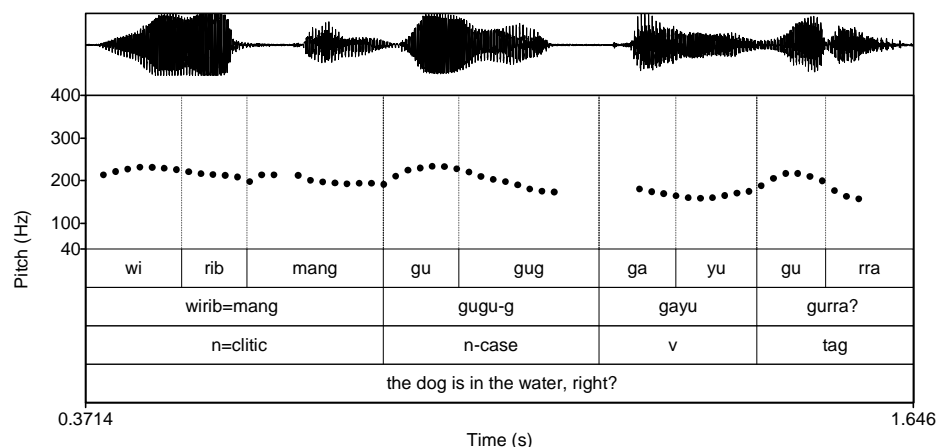


Figure 5-32 Tag questions, the top pane an instance of *ngih* 'isn't it', and the bottom pane *gurra* 'right?'.

5.3.3 Information and polar questions

The overall contour of the IUs is labeled. It is 'falling' when the overall impression is that of a fall throughout the IU; 'rising' when the overall impression is that of a rise from the question word onwards; or 'fall-rise', when a slight rise is heard in the very last portion of the question. The first and second syllables at the left periphery and the penultimate and final syllables at the right periphery are retained for this analysis. A comparative analysis of the correlates of information and polar questions will establish whether they receive the same encoding or different ones.

The questions are organized into subgroups according to their types and their overall contours. A comparison between all of the subgroups is made first, in order to discover whether measures for each of the correlates vary significantly between the polar and the information questions, and between the subgroups established on a perceptual basis.

The subgroups are:

1. Information questions with

- falling
- rising

- fall-rise contours.

2. Polar questions with

- falling
- rising
- fall-rise contours.

The rates of occurrence of each of the contours are counted first (Table 5-8). 55% of tokens in information questions and 59% of polar questions have falling contours. The second most frequent is the fall-rise contour in 33% of information questions and 32% of polar questions. The rising contour is the least frequent, accounting for 12% of information and 9% of polar questions. Importantly, with regard to the starting hypothesis, the distribution of occurrences of the three perceived contours is virtually the same in polar and information questions. The polar questions are thus not more likely to have a rising intonation.

Overall contour	information		polar	
	n	%	n	%
fall	32	55	13	59
rise	7	12	2	9
fall-rise	19	33	7	32
total	58		22	

Table 5-8 The rates of occurrences of the three types of contours in information and polar questions.

A first test is done to ensure that the three contours vary in a significant manner. Given that the main differences in contours are found mainly at their right boundaries, only the two final syllables are retained and an ANOVA test is performed. The contours are found to vary in their mean F0 and their final velocity ($F(2, 155) = 16.767, p = .000$), and ($F(2, 157) = 4.198, p = .017$), but not in their excursion size or duration. The averages are shown in Figure 5-33.

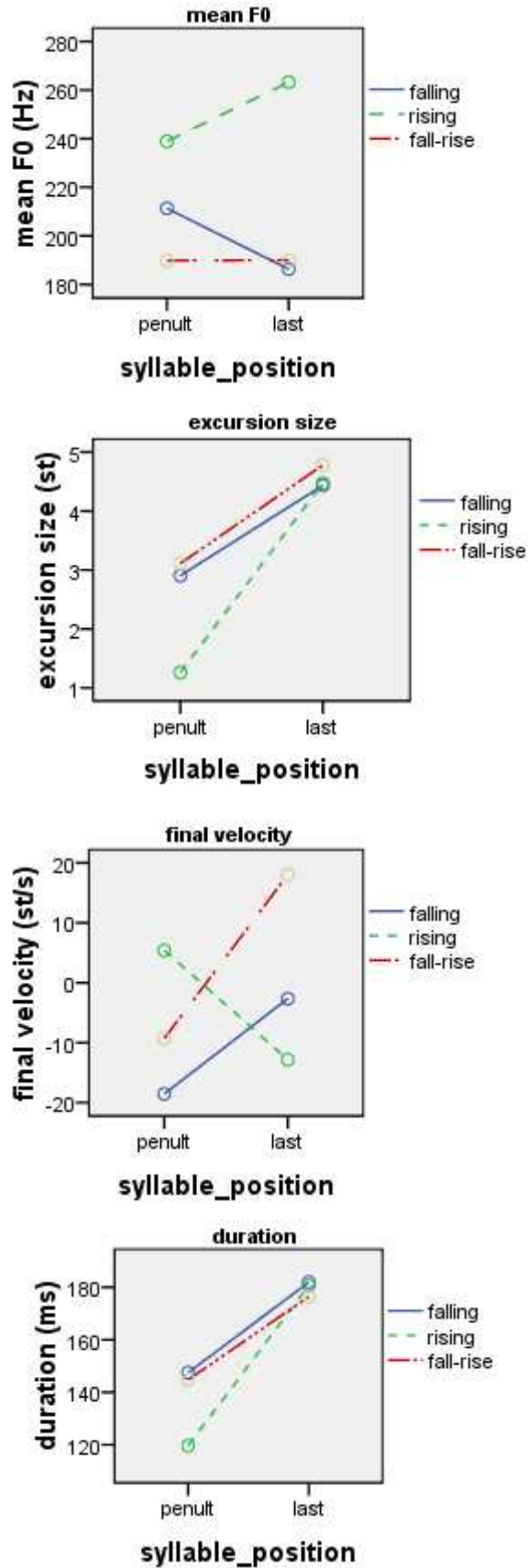


Figure 5-33 The average values for the correlates of mean F0, pitch excursions, final velocity and duration in the penultimate and final syllables of the three observed contours.

5.3.3.1 The falling contour

This is the most frequent of the contours found in this dataset, both in information and polar questions. The first, second, penultimate and final syllables are retained for this analysis.

ANOVA tests show that the variations according to the position of the syllables in the IU are significant for all the correlates. As to the differences between the polar and information question, the tests show that the measurements of average mean F0 and excursion size do not differ significantly between the polar questions and the information questions. The final velocity and duration do differ significantly between the two types of questions ($F(1, 176) = 5.206, p = .024$ and $F(1, 180) = 14.818, p = .000$). The results are shown graphically in Figure 5-34 below.

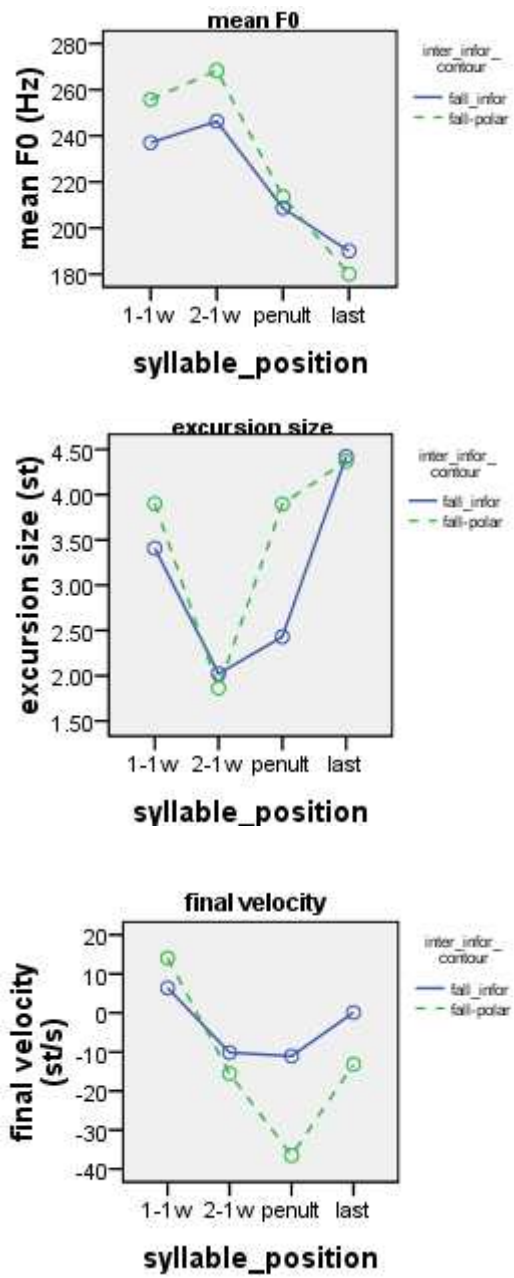
The mean F0 values are indeed very close: polar questions (dotted line) having slightly higher values at the left boundary and lower values on the final syllables than information questions. The values for the excursion size are very similar, except in the penultimate syllables where the polar questions average 3.89st and the information questions have much lower values with 2.43st.

The final velocity values point to similar underlying targets on the initial syllables, but vary significantly at the right edge where the penultimate and final syllables of polar questions have values of -36.53st/s and -13.19st/s, respectively, suggesting that a [fall] target occurs on the penultimate syllables, and that the falling movement is continued into the last syllable. For information questions the values in the same syllables are -11.01st/s and .11st/s, a pattern reminiscent of that found in declarative sentences, suggesting a [low] target.

The differences in the values for duration are also significant. Polar questions have shorter durations than information questions in all syllables but the first.

This suggests that, although the polar and the information questions with falling contours have superficially very similar contours, they vary in their encodings. The polar questions have steeper falls, evidenced by the higher values in the excursion size of their penultimate syllables, the lower

values in their final velocity which may signal a different target, and their shorter durations.



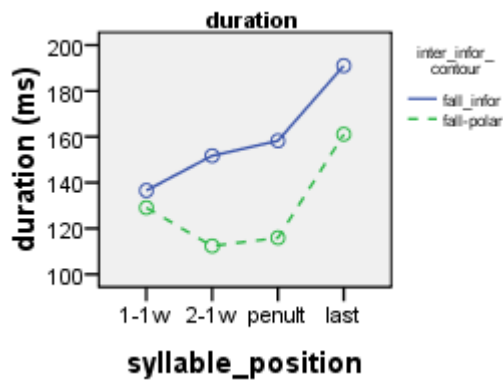


Figure 5-34 A comparison of the correlates of polar and information questions with a falling contour.

5.3.3.2 The rising contour

The measurements for the rising contours, where the rise is heard at least from the onset of the final word, do vary significantly according to the positions of the syllables in the IU, but not between the polar and the information questions. The results are shown in the graphs in Figure 5-35 below.

The mean F0 in the final syllables is markedly higher than in the first syllables, particularly in the polar questions. The excursion size is also much greater in the final syllables.

The information questions have high final velocity values in their first syllables, again reminiscent of the values in the declarative contours, suggesting a [high] target on the first syllable of the first word, in this case most probably the focused question word. At the right boundary, the final velocity values suggest that the rise is realized in the penultimate syllables, with a turn towards zero in the final syllable that would suggest a [rise] underlying target. Again the average values are markedly higher in polar questions.

The durations indicate some degree of lengthening over the last syllables, polar questions being less lengthened than information questions. Overall, the values of mean F0 confirm the perceived rise. The final velocity measurements indicate that the penultimate syllable is the locus of realization of the [rise] target. Polar questions have slightly steeper rises

than information questions, as shown by the wider excursion size and shorter durations.

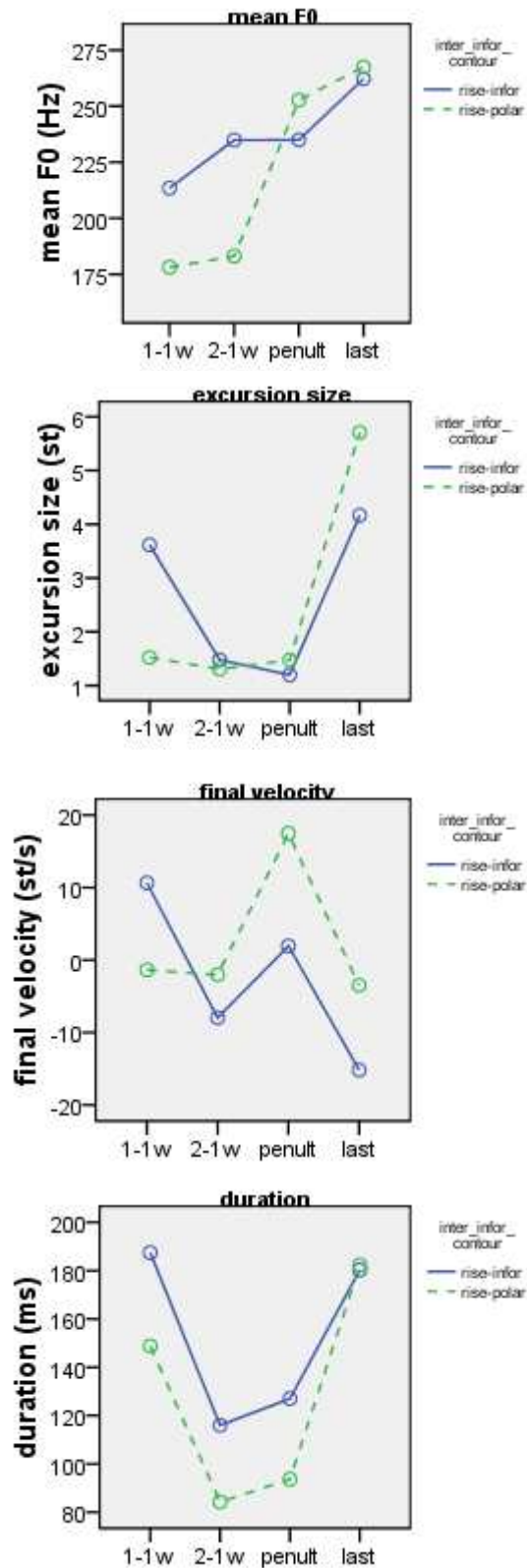


Figure 5-35 The average of mean F0, excursion size, final velocity, and duration in the first, second, penultimate and final syllables of interrogatives with a rising contour, distinguishing between information and polar questions.

5.5.3.3 *The fall-rise contours*

The fall-rise contour is characterized by a falling movement until the penultimate syllable. In the last syllable of the IU, a small but perceptible rise is heard and is also usually evident in the pitch track.

The ANOVA tests show that the differences in the measurement in the correlates vary significantly according to syllable position, but that the polar and information questions do not differ between each other (Figure 5-36).

The mean F0 of information questions does indeed rise in the last syllables, but not that of polar questions, which decreases slightly. The excursion size is important in the syllables at the right boundary, mostly in the final syllables. The perceived rise is confirmed by the final velocity measurements, with values in the penultimate syllables of -7.79st/s in information questions and of -13st/s in polar questions, and in the final syllables of 21.87st/s and 10.1st/s, respectively. The underlying target, realized in the final syllables, is probably a [rise] because of the high positive values of the velocity.

Thus, the contour is marked mostly by a [rise] target in the final syllable, and does not differ between polar and information questions.

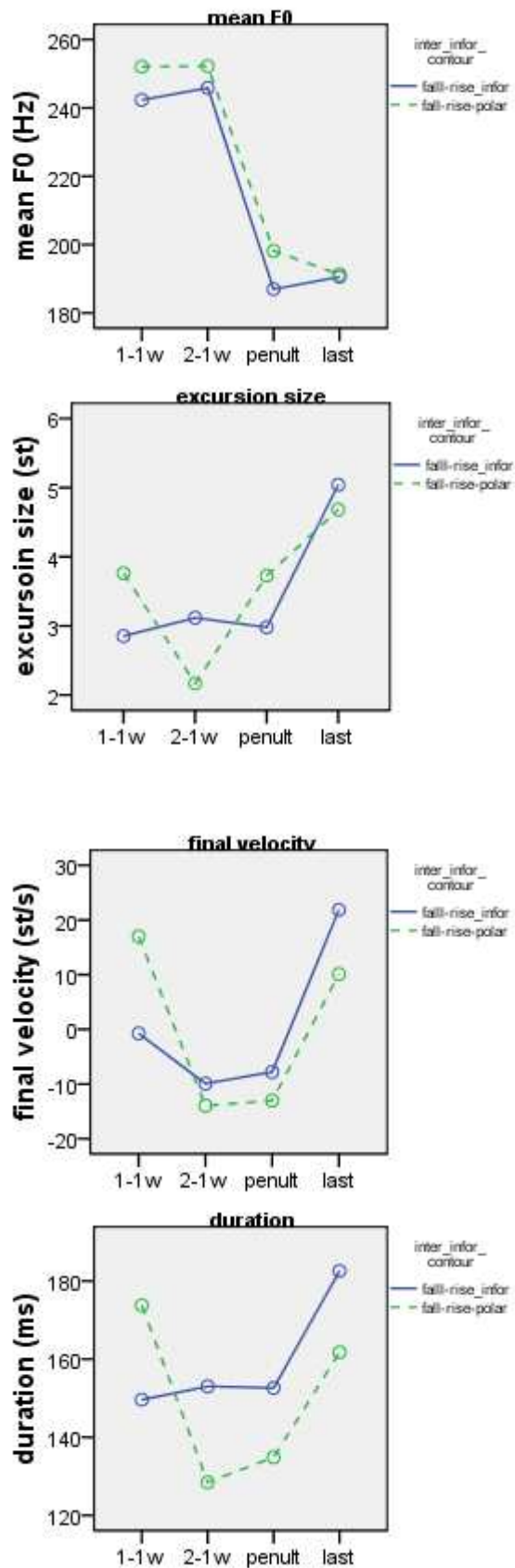


Figure 5-36 The average of mean F0, excursion size, final velocity, and duration in the first, second, penultimate and final syllables of interrogatives with a fall-rise contour, distinguishing between information and polar questions.

5.3.4 Tag questions

A bipartite construction, tag questions consist of a main clause, which exhibits the encoding of the default declarative, followed by the tag. This is illustrated in Figure 5-37 which shows the average pitch of the syllables in the sentences comprised in this dataset. The first IUs (left side of the graph) indeed have a very similar pattern whatever the following tag. The remainder of this section will thus concentrate on a comparison of the tags. Perceptually, *ngih* and *init* display a rising intonation, and *gurra* a falling intonation.

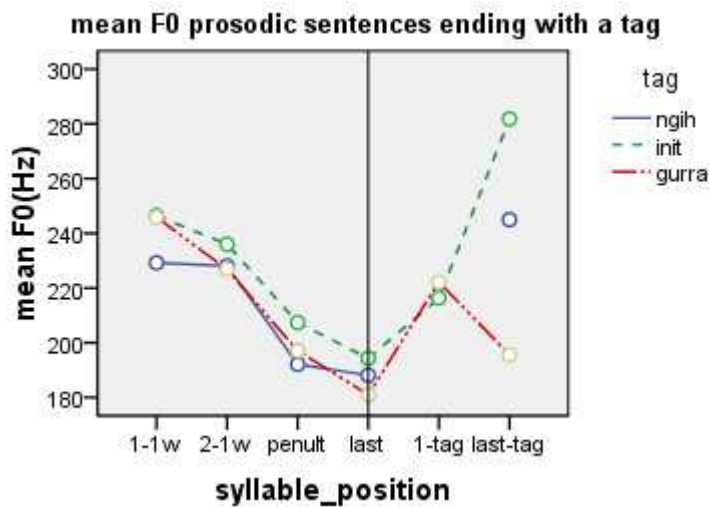


Figure 5-37 The mean F0 values in sentences ending with a tag. The first IUS have very similar default declarative patterns.

The difference in average pitch of the tags is statistically significant ($F(1, 29) = 3.732, p=.039$), *gurra* being different from the other two, as expected. This is evidenced in the right section of the graph 5.37 above.

Figure 5-38 shows the results for the correlates of excursion size, final velocity and duration. The difference in excursion size is significant ($F(1, 29) = 4.856, p=.017$). The final syllables in *gurra* have a value of 3.6st, those of *init* 7.43st, and those of *ngih* 8.08st. The rising contours in *ngih* and *init* are much steeper than the falling one in *gurra*.

The final velocity measures do not vary significantly. Nevertheless, some patterns are suggested. The syllables in *gurra* have values of -28.99st/s and 11.65st/s, suggesting a [low] target. In the tag *init*, the values are 25.81st/s and -4.61st/s and in *ngih* the value is -6.75st/s. The very close

values in the last syllable of *init* and in *ngih* continue to highlight their similarity and it is probable that the rising movement in *ngih* is realized early in the syllable, a movement that is not captured in these measurements.

Finally, the measurements of the duration are significantly different ($F(1, 29) = 15.602, p=.000$), understandably, as the monosyllabic *ngih* is markedly longer than both bi-syllabic tags.

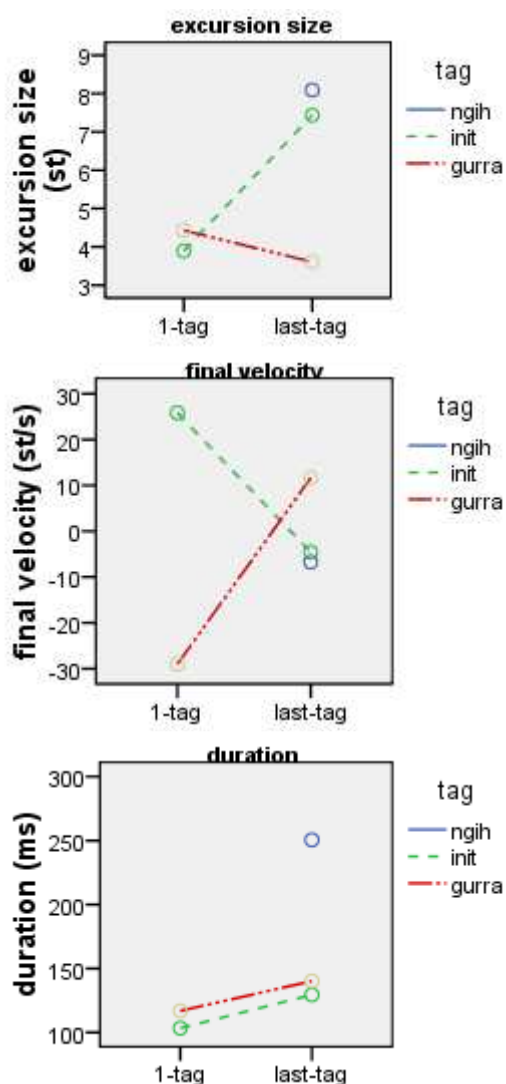


Figure 5-38 The average excursion size, final velocity, and duration in the syllables of the tags.

5.3.5 Summary: interrogatives

The results partially disprove our starting hypothesis, stating that polar questions are more likely to end with a rising intonation. Three contours are found to occur in interrogatives, none of which can be

attributed solely to the encoding of either the polar or the information questions. The most frequent contour is the falling contour, followed by the fall-rise contour, and finally by the rising contour. Although the differences between the interrogative types are not significant for most correlates, in all three contours, polar questions seem to have more marked patterns.

The only statistically significant differences in the measurements of the information and polar questions concern the falling contours for the correlates of final velocity and duration, differences that are not apparent in the syllables at the left periphery, but in those of the right periphery, quite understandably. Information questions have final velocity values that suggest a [low] target in the final syllables, which are also fairly long. The final velocity values of the last syllables in polar questions suggest that they have a [fall] target. Overall, polar questions have a steeper falling contour, which is also evidenced by the excursion sizes, which are wider. Rising contours are characterised by a rising movement that starts early in the IU, at the latest in the first syllable of the last word. The penultimate syllables in the rising contours have a [rise] underlying target. In fall-rise contours, the last syllables are the locus of the pitch movement, with a [rise] target.

In tag questions, the similar encoding of the two tags, *ngih* and *init*, is confirmed, they are marked by a [rise] target realized on the first syllable of *init* and very early on in *ngih*, and a [fall] target on the first syllable in *gurra*. The prosodic correlates of *gurra* are in line with a pragmatic interpretation of its meaning as a weaker signal of questioning/uncertainty.

5.4 Imperatives

5.4.1 Hypothesis and dataset

When listening to Jaminjung speakers, it is fairly clear that imperatives are uttered with greater intensity, a correlate which cannot easily be measured within the parameters of this study. With regards to the correlates of mean pitch, excursion size, final velocity and duration, imperatives are expected to be similar to the default declaratives.

The basic criterion used in the selection of the 23 IUs analysed in this section is the presence of an inflected verb in the imperative form. This verb may occur in any position in the IU. Example (110) is shown in Figure 5-39, where the speaker is warning a child not to walk on her fishing line.

(111)

ba-ngawu thanthu=gayi mununggu
IMP-see DEM=ALSO string

'Watch that fishing line too!'

109 [IP:ES97_01_03]

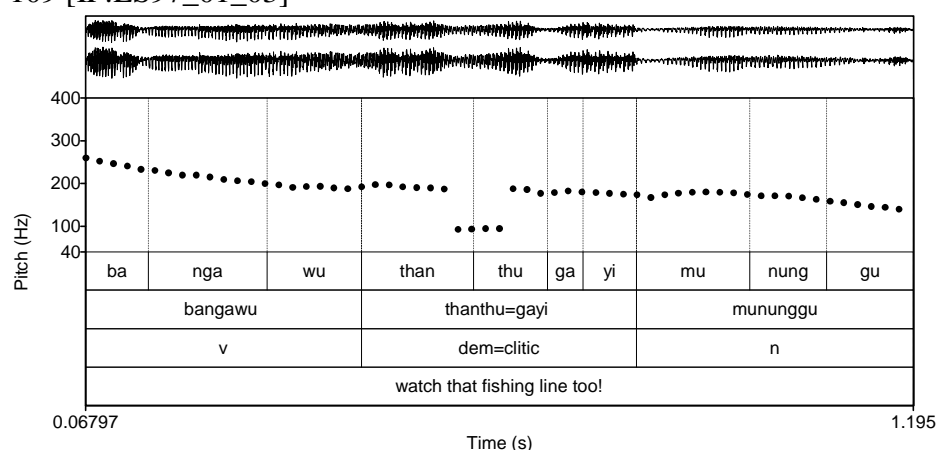


Figure 5-39 An imperative sentence with a verb in the imperative form.

5.4.2 The correlates of imperative sentences

All IUs containing imperative verb forms in this dataset have an overall falling contour, with the expected prominence at its left boundary and a reset on each prosodic word. The values in the syllables of the last word seem to drop quite sharply. This overall contour is illustrated in a boxplot in Figure 5-40.

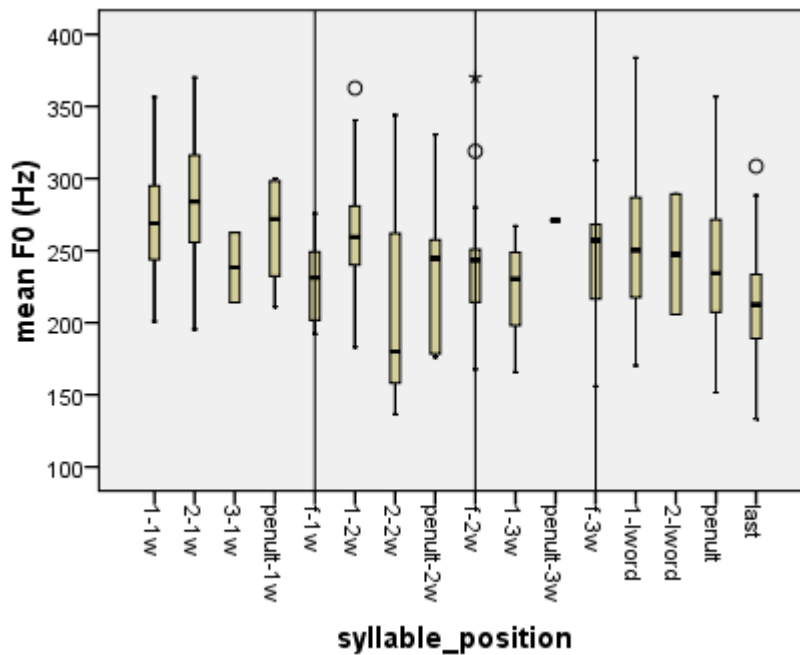


Figure 5-40 The mean F0 values of the syllables in imperative sentences. Words are separated by vertical lines.

The values for each of the correlates of mean F0, excursion size, final velocity and duration are shown in graphs in Figure 5-41 below. The values are shown for the first and second syllables in the IUs, and then for the penultimate and last syllables in the final word.

The measures of the correlates of mean F0 ($F(3, 75) = 10.355$, $p = .000$), final velocity F0 ($F(3, 76) = 3.332$, $p = .024$), and duration F0 ($F(3, 75) = 5.227$, $p = .003$) vary significantly according to the position of the syllable in the IUs.

The mean pitch values in the first and second syllables display the pattern associated with the pitch peak at the left boundary. In the final syllables, mean F0 values decline quite sharply from the penultimate to the last syllable, with averages of 242.42Hz to 213.88Hz. The steep decline is also conveyed by the wider pitch excursions in the syllables at both edges. The patterns of final velocity in the penultimate and final syllables are very similar to those of declarative sentences, with values of -22.89st/s and -8.14st/s, with the same interpretation as a [low] target.

The final syllables are markedly longer than those at the left boundary, by almost 40ms, with respective values of 129.62ms, 126.84ms, 161.89ms and 179.29ms.

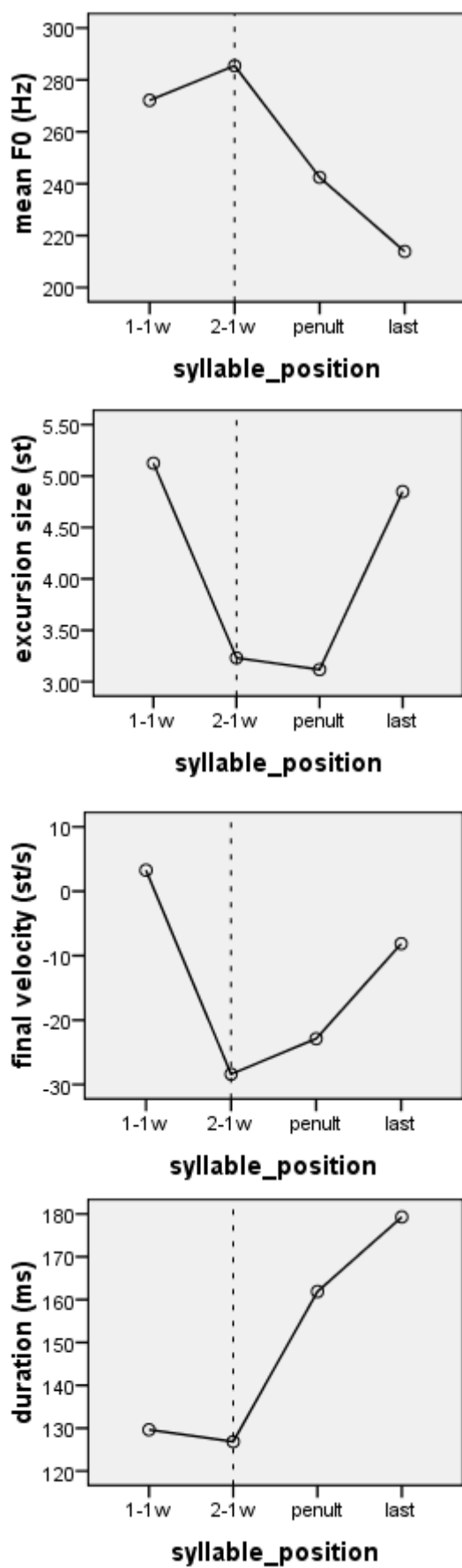


Figure 5-41 The average of mean F0, excursion size, final velocity, and duration in the 1st, second, penultimate and final syllables of imperative sentences.

5.4.3 Summary: imperatives

The measurements indicate a steep falling contour at the right edge in imperatives, evidenced by the wide excursion sizes. The final velocity values in the penultimate and final syllables suggest a [low] underlying target.

Perceptually, imperatives are uttered with more force than declaratives, but this will have to be confirmed in later research, when recording conditions make it possible to test intensity measurements.

5.5 Comparison between declaratives, interrogatives and imperatives

To conclude this section, a comparison is made between the contours in declaratives (sentences formed of one IU), interrogatives and imperatives, particularly as they seem to share a similar falling contour.

The dataset comprises the IUs analysed for the simple declaratives which all had falling contours. In interrogatives, the falling contours of the information and polar questions are distinguished, as their encoding was found to differ. The subtypes of interrogatives are not distinguished for rising and fall-rise contours, as no significant difference between them was found (§5.3.3). Imperatives only have falling contours. Only the first and second syllables and the penultimate and final syllables in the IUs are retained. As the patterns for each of the subtypes have already been described, the comparison will seek to establish if, and how, they vary.

The mean pitch varies significantly between the different contours/sentence types ($F(5, 584) = 39.691, p = .000$). Declaratives differ from all other types and so do the imperatives, while the interrogatives form a subgroup of their own. The averages are shown in Figure 5-42. Of special interest are the falling contours, which show that sentence types are distinguished by a different register: imperatives have the highest overall values, declaratives the lowest, and the interrogatives lie somewhat between these two extremes.

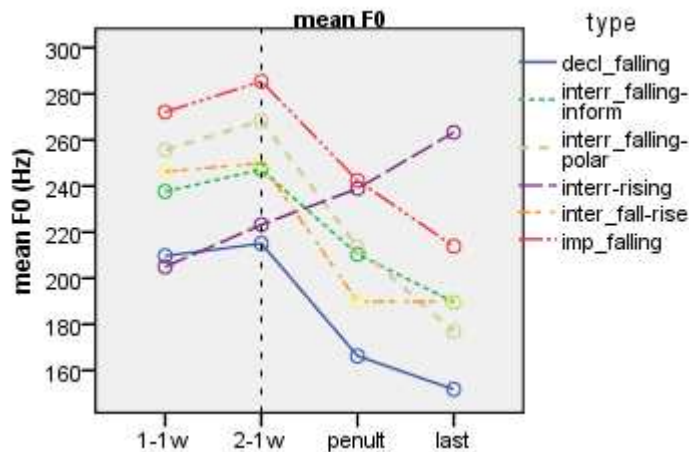


Figure 5-42 A comparison of the mean F0 values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences, distinguished according to their contours.

Although the differences in excursion size do not reach statistical significance, it is interesting to note the wide excursions in the first syllables of the imperatives. The interrogatives with a rising contour seem to have a less salient pitch excursion, at least in the non-peripheral syllables.

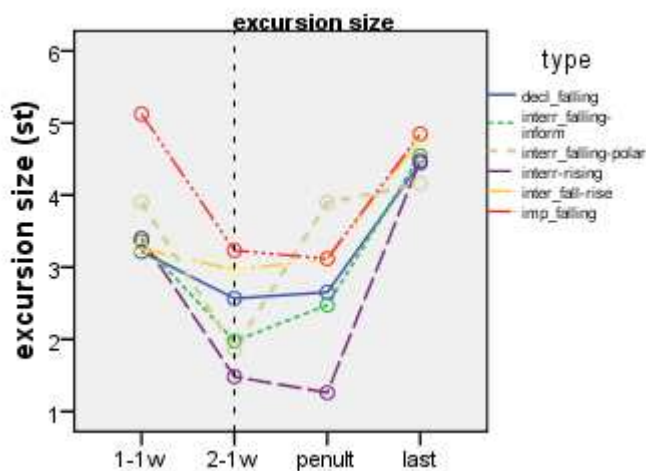


Figure 5-43 A comparison of the excursion size values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.

The measures of final velocity vary significantly between the sentences types ($F(5, 577) = 3.999, p = .001$) (Figure 5-44). All the falling contours have very similar patterns. The most strikingly alike are those of declaratives and information questions, while polar questions are distinguished by their lower values in the penultimate syllables. The pattern for the imperatives is also very similar, at the right boundary, but differs at

the left boundary, with much lower values in the second syllables.

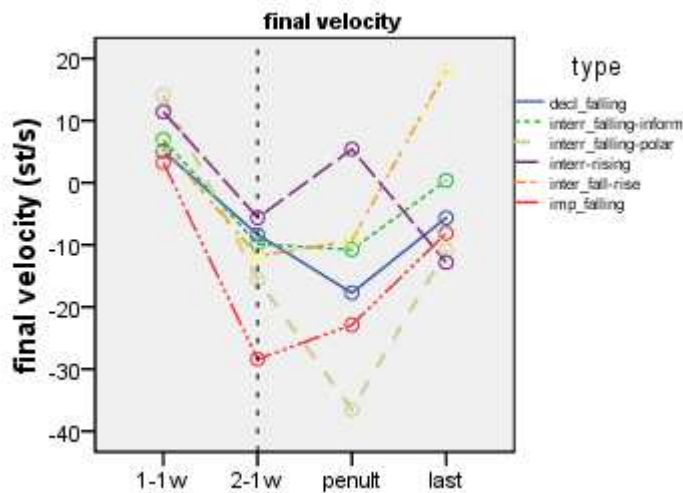


Figure 5-44 A comparison of the final velocity values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.

The different sentence types/contours do vary significantly in their syllable durations ($F(5, 599) = 2.707, p = .020$). The significant differences are between the durations of the polar and information questions with a falling contour, and the interrogatives with a fall-rise. The values are indeed surprisingly close in the final syllables for all the subtypes.

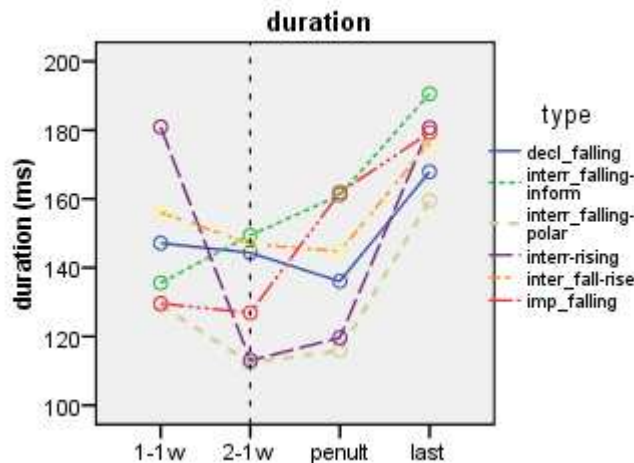


Figure 5-45 A comparison of the duration values of the syllables at the left and right boundaries of declarative, interrogative, and imperative sentences.

5.5.1 Conclusion and discussion: comparison of sentence types

In the PENTA model of intonation, it is assumed that each communicative function has to have at least one dominant encoding

characteristic for it to be identified by the hearer. One basic assumption on which this chapter is based is that sentence types are differentiated by prosodic means in Jaminjung. The comparison between sentence types yields an important result: although declaratives, interrogatives and imperatives are all predominantly uttered with a falling contour, they are clearly differentiated by pitch register – declaratives use lower reaches, imperatives higher reaches, and interrogatives somewhere in between. There are some other differences in the sentence types investigated. The fall-rise and rising contours are found in interrogatives of both types, as well as in the first IUs of complex declaratives, but are not attested in imperatives. These contours can be interpreted as continuation rises, but they are not being used systematically in this function, and the lack of other clear motivations for their use would warrant further research.

Imperatives and the polar questions with a falling contour have steeper falling contours, as shown by their wider excursions and more extreme velocity values, making them the more marked structures.

The findings reported here for Jaminjung, particularly for the contours of the interrogatives, is not unlike those found in other languages of Australia. King (1994:144) reports very similar patterns in a study of Dyirbal, a language of Northern Queensland, where information and polar questions exhibit the same falling contour, but where 27% of the tokens analysed end in a ‘high final boundary tone’. The patterns of interrogatives described by Bishop and Fletcher (2005:338) for Bininj Gun-wok also resembles those described here, including the realization of the falling pattern on the penultimate syllable.

The lack of a rising pattern in interrogatives may be associated with socio- or ethnolinguistic considerations. One of these considerations concerning Aboriginal culture is the importance of ‘indirectness’ in social interactions. Indeed, Eades⁶⁰ (1988:105ff.) points out that although Aboriginal social life is very public, “the Aboriginal way of interacting indirectly preserves a considerable degree of personal privacy”. She further

⁶⁰ Her observations are based on her work in Southeast Queensland.

states that Aboriginal people use different strategies than do English speakers in seeking information.

Questions are used to find out about another's kin, movements, and country, but not to seek substantial information, such as important personal details, a full account of an event, or the explanation of some event or situation. In these situations indirect strategies are used: the speaker contributes some of their own knowledge on a topic and then leaves a silence, to lead the person with the knowledge to impart information. Important aspects of substantial information-seeking are the two-way exchange of information, the positive, non-awkward use of silence, and often considerable time delays (frequently several days) between the initiation of substantial information-seeking and the imparting of information (Eades 1988:107).

The same patterns of language use are found amongst Jaminjung speakers. While they do occur, direct questions tend to be avoided, and repetitions and other strategies are used to seek confirmation and request further information. The common use of tag questions in Jaminjung, of which two are analysed in this chapter, is a case in point.

In a wider perspective, although the correlation between rising and falling contours, and questions and assertions, respectively, is amply recognized in the literature (see Ladd 1996 for a review), some authors, such as Nilsenova (2006), show that rising intonation does not always go hand in hand with question intonation and proposes that the primary linguistic interpretation of final rises is uncertainty (rather than questioning). Another interpretation is suggested by House (2004), based on a study of Swedish where final rises are also optional in interrogatives. He makes use of the categories proposed by Bell and Gustafson⁶¹ (1999) to categorise utterances according to whether the speakers try to retrieve information, or whether they are seeking feedback (socializing). He finds evidence that the final rise in the material analysed is a signal of socializing, while the final low is a

⁶¹ Developed for the annotation of the August spoken corpus.

signal of information seeking. He goes on to link this interpretation with the biological codes proposed by Gussenhoven (2002) for the universal meanings of intonation. Gussenhoven proposes three codes: a frequency code which implies that a raised F0 is a marker of submissiveness or non-assertiveness and is therefore associated with question intonation; an effort code which implies that articulation effort is increased to highlight important focal information producing a higher F0; and a production code which associates high pitch with phrase beginnings (new topics) and low pitch with phrase endings. House goes on to link the final rises in ‘socializing’ questions and the falls in ‘information seeking’ questions with the frequency code.

This section has established clear differences in the encoding of declarative, interrogative and imperative sentences in Jaminjung. Whether these differences are culturally motivated, or whether they are artifacts of more universal prosodic phenomena is beyond the scope of this research and will have to remain a subject for further research.

To complete this inventory, two more subtypes of clauses will be analysed, vocatives and quoted speech.

5.6 Vocatives

Vocatives are informally defined as nominal forms used as address or in addressing people. Two basic functions are thus given to vocatives, they are used either as *calls* or as *addresses*. Calls are designed to catch the hearer’s attention while addresses simply aim to maintain or emphasize the contact between the speaker and the addressee (Zwicky 1974). Vocatives have posed a challenge to prosodists because they are usually associated with specific, or ‘stylised’ intonational patterns that are not easy to account for in compositional models of intonational meaning. Nespor and Vogel (1986) state that in languages that make use of them, they are set off from the rest of the sentence by some special intonation and that they obligatorily form independent intonational units. For English, in AM theory, their pattern is described as a H* !H- L% standing for a downstepped high

phrase tone ending in a low tone. Further, it is suggested that these contours may change according to their position in the sentence.

This section will present an analysis of vocatives in Jaminjung.

5.6.1 Dataset

The vocatives analysed in this dataset are addresses rather than calls, used to emphasize the contact with the listener, often uttered in a context of language elicitation when the speaker addresses the linguist, using their subsection name, usually a three-syllable word as with *Nambijin*, in example (112) in Figure 5-46.

(112)
 nginju=biyang burr-angga=mindi Nambijin
 PROX=now 3PL-go=12du <subsection>

'Here now, they are going away as we can see, Nambijin.'
 [JM:CS07_62_01]

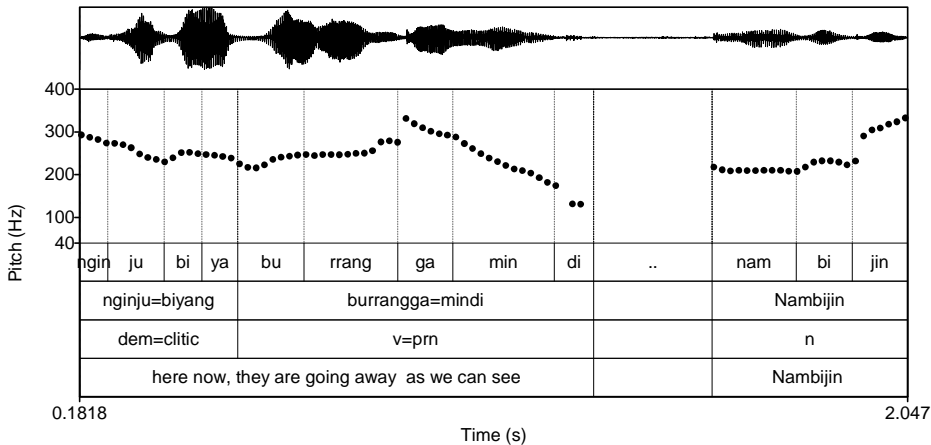


Figure 5-46 A vocative in sentence-final position, used as an address.

In Jaminjung, the vocatives form an intonation unit of their own and tend to occur either sentence medially or in sentence final position. In the prosodic sentences in this dataset (14 tokens), the position of the vocative in the prosodic sentence is labeled, as well as their number of syllables and relative position.

5.6.2 The correlates of vocatives

The pitch contour in vocatives is found to rise in all but one token in the dataset, a pattern that is not influenced by the position of the vocative in the prosodic sentence. The mean F0 of the first, second and final syllables are shown in the boxplot in Figure 5-47.

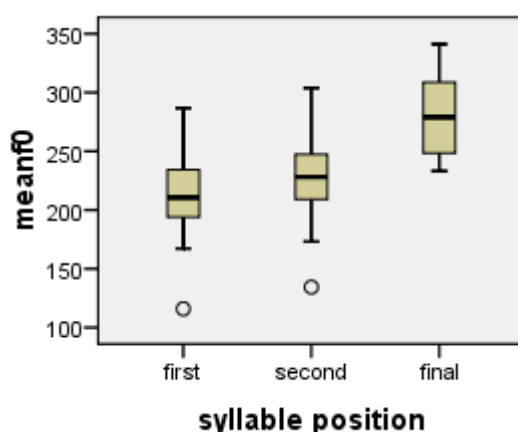


Figure 5-47 The mean pitch of the syllables in vocatives, showing a rise in values.

The measurements for all four correlates are shown in Table 5-9. The mean pitch has been mentioned above. The excursion size in the last syllable is greater in the final syllables than in the preceding ones, indicating a sharp rising movement. The measures of final velocity suggest that the rising movement is initiated right from the first syllable. The low values of the final syllables suggest that the underlying target is probably a [high] rather than a [rise]. The durational values follow the expected pattern of a longer syllable in initial position, a notably shorter mid syllable, and a final syllable marked by a longer duration.

Correlate	Syllable position	Mean	Std. Dev	N
mean F0	first	212.94	44.75	13
	second	231.39	45.47	14
	final	268.99	49.85	14
excursion_size	first	2.09	1.75	14
	second	2.09	0.81	14
	final	4.93	2.60	14
final_velocity	first	6.00	17.22	14
	second	-7.18	22.92	14
	final	-2.30	18.43	12
duration	first	185.39	54.69	14
	second	130.39	32.18	14
	final	157.15	41.30	14

Table 5-9 Values of mean pitch, excursion size, final velocity and duration in the syllables of the vocatives.

5.6.2.1 Comparing the vocatives to the surrounding IUs

The specificity of the patterns for vocatives can only be understood by comparing them with the surrounding IUs. The IUs adjacent to the vocatives are retained for this comparison, either immediately preceding or following it. Given the heterogeneous nature of the IUs before and after the vocatives (uncontrolled for focus, position in the sentence, etc), the average measurements of their syllables presented here is to be treated with caution. The syllables at the left and right boundaries of the units are retained, that is, the first and second, and the penultimate and the last syllables. The patterns are shown in the graphs below. The vocatives are shown in a continuous line, while the preceding and following IUs are shown in dotted lines.

The IUs adjacent to the vocative display the expected patterns of mean F0 (Figure 5-48) associated with declarative contours - a slight increase at the beginning of the IU, and a decrease throughout, ending with a slightly higher value on the final syllable, indicating continuation. The vocatives however have a different pattern. There is very little indication that they are set off by a pitch reset and they have a very steep increase in pitch throughout.

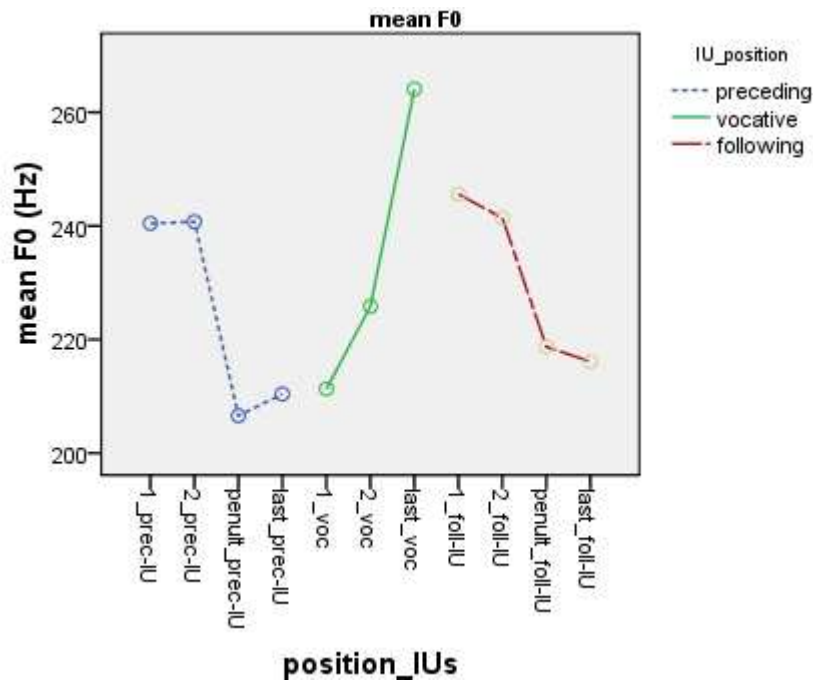


Figure 5-48 A comparison of the mean F0 values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).

The excursion size varies significantly between the syllables in the IUs ($F(18, 109) = 3.065, p = .004$). The pattern of excursion sizes in the adjacent IUs is not unexpected, the syllables at the left and right edge having wider excursions. The vocatives are distinguished by a very salient excursion on their final syllables, with an average of 5.1 st.

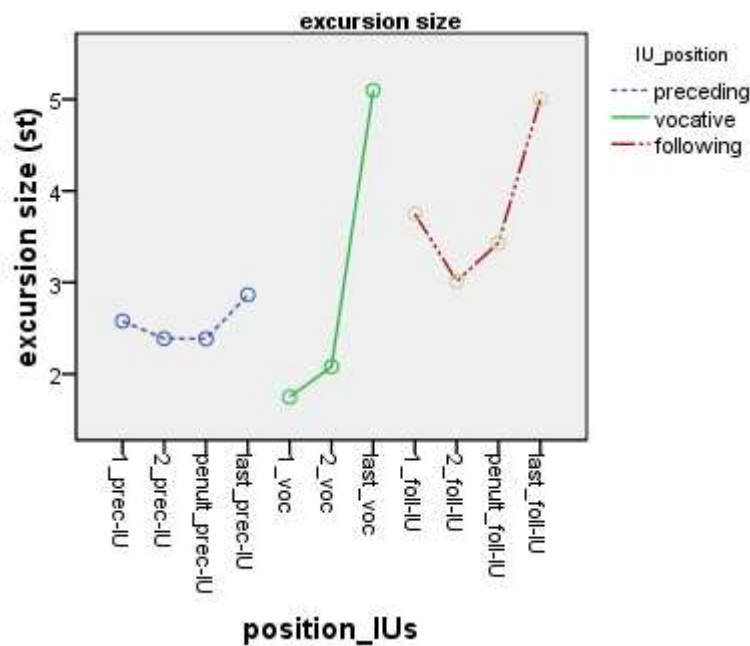


Figure 5-49 A comparison of the excursion size values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).

As to the final velocity, the differences in the measurements are not significant. The low positive values in the first syllables and a movement towards zero in the subsequent syllables, combined with the mean F0 measurements, suggest not a [rise] underlying target, but a [high] target in the vocatives.

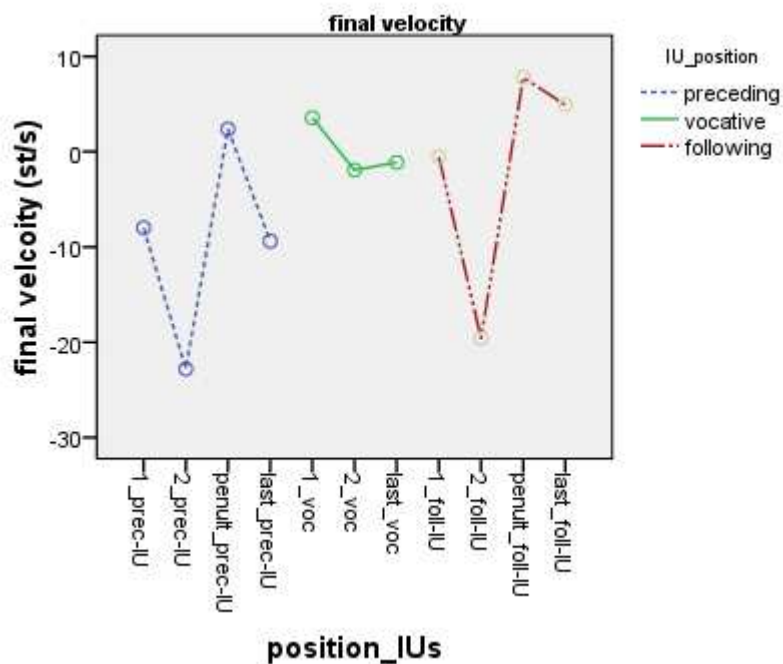


Figure 5-50 A comparison of the final velocity values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).

The differences in syllable duration are significant ($F(8, 109) = 2.444, p = .019$). Vocatives appear to conform to the pattern of final syllable lengthening that is expected in an IU (even in a single-word one), but this last feature seems to be fairly attenuated, particularly when compared to that of the preceding IU, as shown by the values of the ratios of the final to the penultimate syllable: 1.8 in preceding IUs to 1.17 in vocatives.

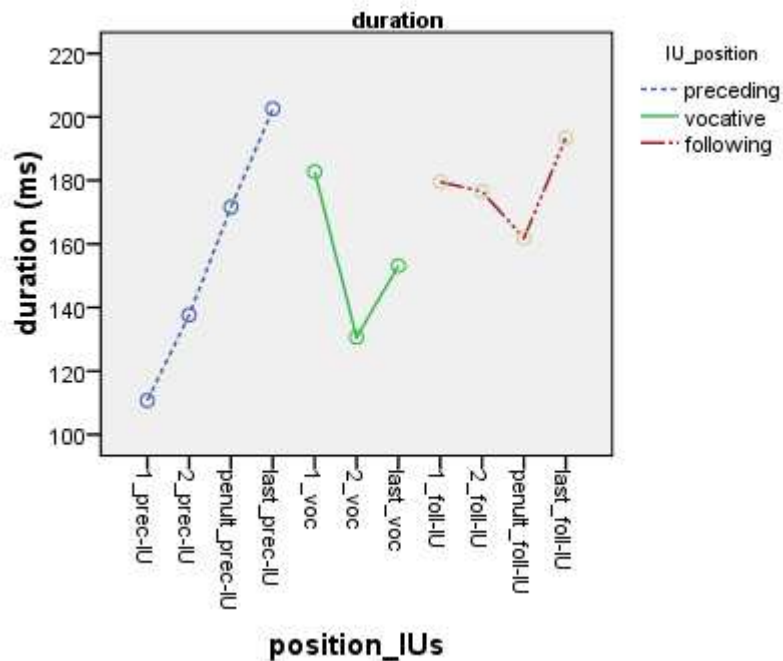


Figure 5-51 A comparison of the duration values in the syllables of the IUs preceding and/or following vocatives (shown in a continuous line).

In short, vocatives have an easily identifiable rising contour. They are marked by a rising pitch throughout and by a large excursion size on their last syllable. The values of the final velocity indicate [high] underlying targets. The position of the vocatives in the prosodic sentence does not influence its prosodic encoding.

5.7 Reported or ‘quotation’ speech

Repeating what another person has said, or paraphrasing in their own words are two strategies available to speakers in any language to report on what someone else has said. Marking quotations prosodically is very common in Jaminjung, and is easily recognisable in all the speakers’ utterances, giving texture to Jaminjung narratives and other forms of speech. This section will establish the parameters used to encode quotations. Evans (2009) proposes a tri-partite categorization of reported speech based on deictic elements: “(a) direct speech, in which the perspective of the original speaker is completely maintained, (b) indirect speech, in which the perspective is entirely assimilated to that of the reporting speaker (c)

biperspectival speech, in which the twin perspectives of original and reporting speaker are simultaneously coded at every relevant points “. Aikhenvald (2004:132) reports that some Australian languages make do without indirect speech constructions. In Jaminjung, reporting speech usually involves a direct quotation, and is thus associated with the first of Evans’ categories.

5.7.1 Dataset

The constructions analysed in this section are all quotations introduced by an IU containing a quotative verb form (usually the SAY/DO verb *-yu(nggu)*) and an IU that consists of the quotation itself, sometimes followed by other IUs, similar to English utterances such as ‘the old man said “there is one by the river”, and the kids went to look that way’. There are a few tokens where the verb is omitted, the speaker simply stating who is talking in the subsequent quoted speech, a strategy often used in the course of a longer narrative, i.e. ‘Paul now: “this is a big fish, this one”’. The IU containing the quotative verb can also be placed after the quotation. Blythe (2009) refers to these constructions, in any of the variants described here, as *quotative expressions*, a term which I will adopt in this section. The working hypothesis, based on perceptual impressions, is that these IUs are marked by both a higher pitch register and wider pitch excursions.

The dataset consists of 24 IUs. A typical example (113) of reported speech in the form of direct quotation is given in Figure 5-52. The quotation follows the verb, starting with the Kriol *maitbi* ‘maybe’ and ends with the verb *ganinyba* ‘it bit you’. The speaker continues with her narrative in the next prosodic sentence. Both prosodic sentences are shown in the figure to show the context in which the quotation occurs.

(113)

en burru-yu=rrgu <maitbi> jalarriny-ni
and 3pl:3sg-say/do.PST=1sg:OBL maybe centipede-ERG

ganiny-ba burr-ijga-ny birdij burr-ina
3sg:2sg-bite.PST 3pl-go-PST find 3pl:3sg-chop.PST

<thet> jalarriny
that centipede

'And they said to me, "maybe a centipede bit you". They went (and) found the centipede.'

[IP:ES97_03_02]

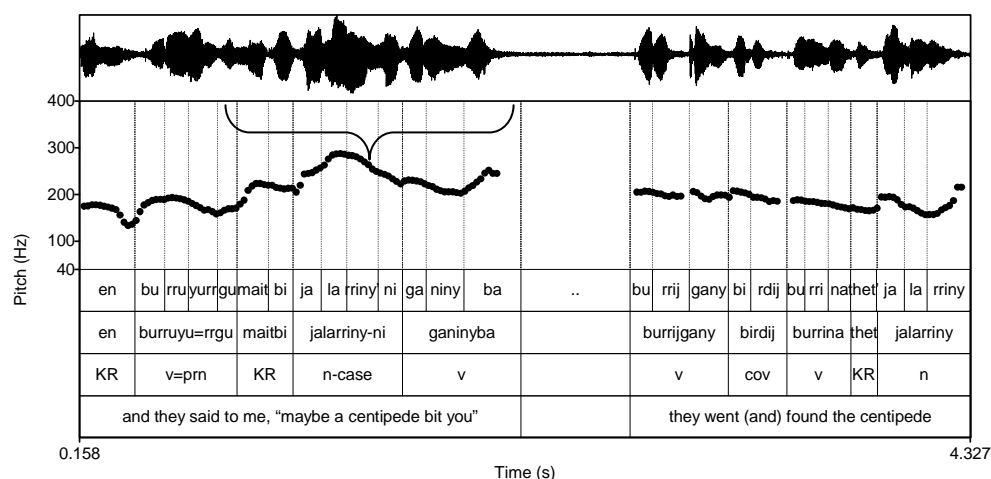


Figure 5-52 A typical example of reported speech in the form of direct quotation. It is shown by a bracket, starting with *maitbe* and ending with *ganinyba*.

The syllables in the prosodic sentences are labeled according to their positions within the IU; the position of the IU in the prosodic sentence is also coded. A further code is added to locate the quotative verb, indicating whether it is found in the IU preceding or following the quotation. The syllables containing the minimum and maximum pitch in the prosodic sentence are also coded.

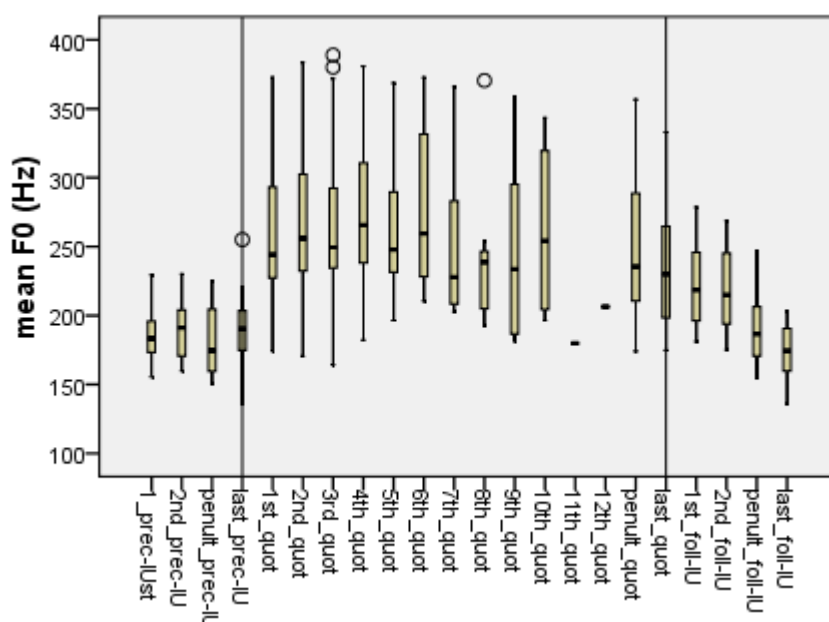
5.7.2 The correlates of the quotative expressions

The general profile of the utterances in the dataset is illustrated with boxplots in Figure 4-53. The top pane shows the average mean pitch and the bottom pane the pitch excursions of the syllables. Only the boundary syllables are shown for the preceding and following IUs to avoid overcrowding the graphs. The IUs are separated by vertical lines.

The higher register of the quoted speech is very apparent in the mean pitch averages. The first IU (left portion of the graph), corresponding to the IU containing the quotative verb, is distinguished by its very low overall pitch values, which is unusual in the first IU in a complex sentence (§5.2.3.2).

The pitch minima and maxima in the prosodic sentence give an indication of the pitch expansion⁶² of the quoted material. In this dataset, the pitch minima are more likely to occur at the right edge of the prosodic sentence, in 41.7% of cases, but they can also occur either in the first IU or within the quotation. The pitch maxima, on the other hand, are always found in the quoted material. The counts are shown in Table 5-10.

Finally, the pitch excursion values indicate that the first syllables of the quotative expressions are uttered with particularly wide excursions.



⁶² The highest and lowest pitch reaches within the IU.

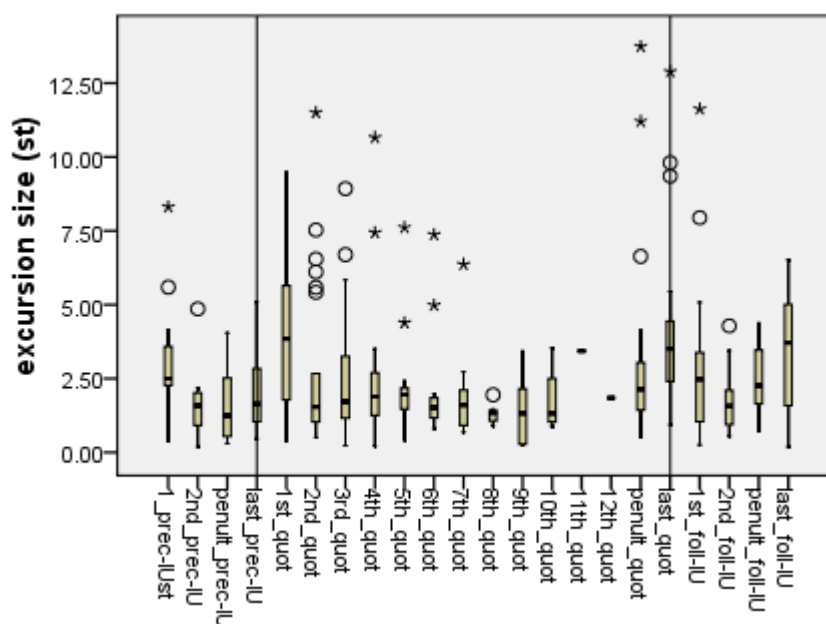


Figure 5-53 The values of mean F0 and excursion size in the IUs surrounding the quotative expressions. IUs are separated by a vertical line.

F0 min and max		Prec_IU	Prec_IU (quot-verb)	quot	Foll_IU (quot-verb)	Foll_IU	Total
min	Count	3	4	7	4	6	24
	%	12.5%	16.7%	29.2%	16.7%	25.0%	100.%
max	Count	0	0	24	0	0	24
	%	.0%	.0%	100.0%	.0%	.0%	100%

Table 5-10 The location of the pitch minima and maxima in the sentences containing quotative expressions.

Finally, the measurements of the correlates are compared between the quotative expressions and the syllables at the boundaries of the surrounding IUs.

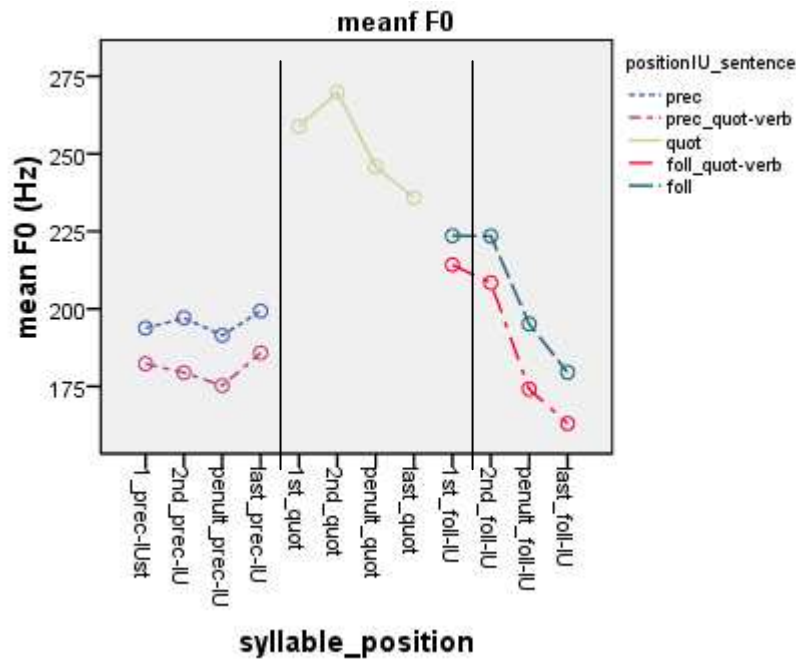
Differences in mean F0 according to syllable position are significant, particularly distinguishing those of the quotative expressions from all others ($F(9, 213) = 2.617, p = .007$). When the preceding and following IUs are coded according to whether they contain the quotative verb, the results indicate that the IUs with a quotative verb tend to have lower mean F0 than other IUs, both when they occur before and after the quotation. The patterns are shown in Figure 5-54.

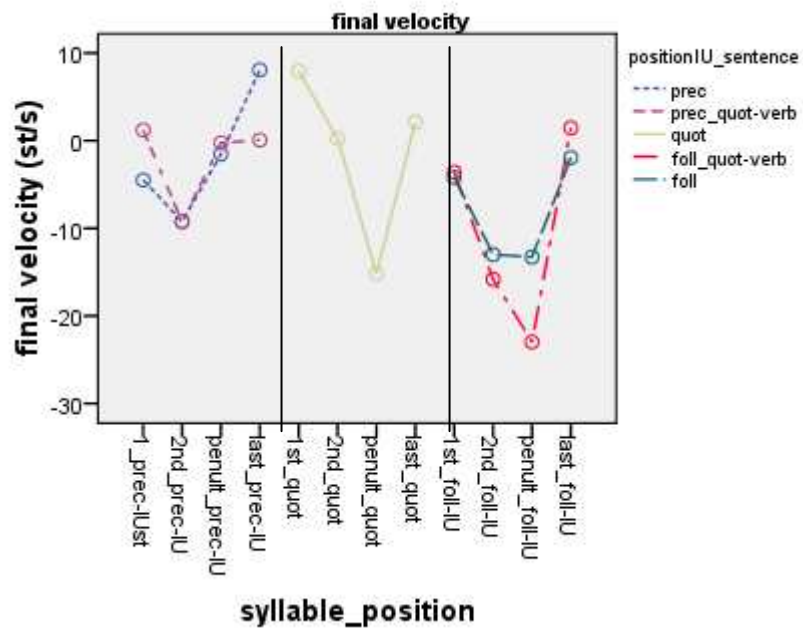
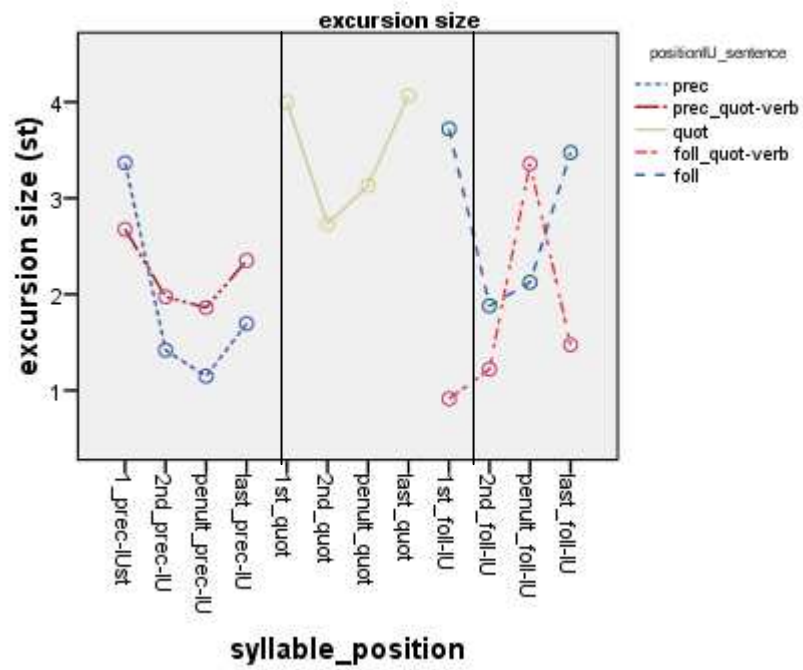
The differences in excursion size are not significant in the syllables of IUs before and after the quotative expressions, but again that of quotations differs from all of these. As shown in Figure 5-54., IUs with the

quotative verb seem to have less salient variations in excursion sizes from syllable to syllable, suggesting a flatter contour.

The final velocity values do not differ significantly. The IUs have values that seem to follow those in other declarative constructions.

The values for duration show significant differences between both the types of IUs and the positions of the syllables within them ($F(2, 212) = 4.121, p = .018$ and ($F(9, 212) = 3.182, p = .001$)). The durations of the syllables in quotative expressions follow the same patterns as in the other IUs, except in their final syllable which is much longer, as shown by the ratios of the durations of the final to the penultimate syllables, which are 1.23 to .89 (without verb) and 1.06 (with verb) in the preceding IUs, or 1.05 (without verb) and 1.06 (with verb) in the following IUs.





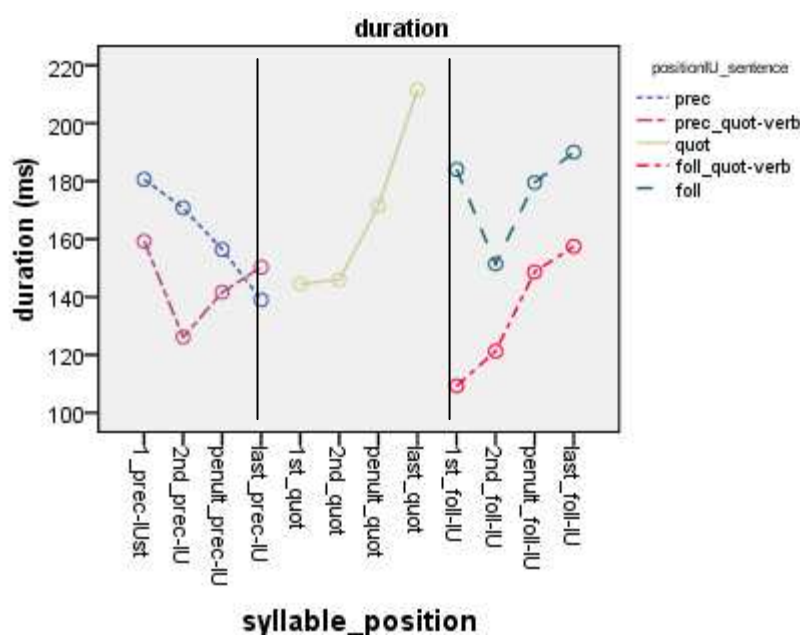


Figure 5-54 The average values of mean F0, excursion size, final velocity and duration in the syllables of the IUs surrounding the quotation expressions.

To conclude, the quotative expressions analysed in this dataset have a raised pitch register, a wider pitch excursion, and a considerably lengthened final syllable. Despite the importance of the pitch movements at their left boundaries, the pattern of underlying pitch targets is not different from the default declarative pattern. The initial prominence is more salient in direct speech, a characteristic also noted for Warlpiri (Chapman 2007). As to the surrounding IUs, those containing the quotative verbs tend to be uttered with a lower register, whether they occur before or after the quoted material.

There is no doubt that the use of prosody in quotation speech has much wider scope than that of a single utterance. When listening to Jaminjung speakers telling a story or even in conversations, it is apparent that they mark the speech of various participants by changing the intensity, pitch and even rhythm of their voice. Some researchers have sought to describe these patterns, for instance, the extensive use of prosodic means to track referents in the context of reported prior interaction by speakers of

Murriny-Patha⁶³ is studied in detail by Blythe (2009:250ff.). Establishing how prosodic patterns span wider stretches of speech, be they narratives or conversations, remains a subject for further research in Jaminjung.

5.8 Conclusion: sentence types

In this chapter, the encodings of various constructions have been described. They can be summarized as follows:

The default declarative contour is characterized by an overall decline in pitch. IUs are set off by a pitch reset; they have a prominence on the first syllable, which is associated with the encoding of focus. The first syllable may have a [high] or a [fall] depending of the information structure of the IU. It is followed by a sequence of [mid] targets, and demarcated by a final [low] target.

In complex declaratives, when they consist of a succession of verbal clauses, the most common pattern is for a repetition of the default declarative contour. IUs are usually bounded by pauses, but they may adjoin each other without any break.

The first IUs in complex sentences can also have a fall-rise or a rising contour. In the fall-rise contour, the pitch falls, as in a default declarative contour, until the final syllable in which there is a clear rise. These syllables receive a [rise] pitch target. In the rising contour, the rise start at the beginning of the IU-final word, the same [rise] target is realized, sometimes from the penultimate syllable.

First IUs can also have a level contour which is marked by lower pitch values throughout and lesser pitch excursions. Their final syllables have [mid] underlying pitch targets and are marked by longer durations than those in both falling and fall-rise contours.

IUs that are iconically lengthened are not longer level contours. At their left edge, they have pitch patterns that are like those of the declarative contour.

⁶³ A non-Pama-Nyungan Australian language of Northern Territory of Australia, probably related to the Daly group.

Pitch is then maintained at that level, thus giving the impression of a ‘stylized high contour’. They have [mid] targets on the lengthened segment, including the final syllable.

Another contour is described for the second IUs (or rather sentence-final, as they could well occur in sentences with more than 2 IUs), termed the secondary prosodic contour. It usually coincides with non-finite predicates and afterthoughts (but may also apply in some verbal clauses) which functionally bring in new information related to the main clause, either delimiting the event of the main predicate in time and space, or specifying one of the referents. It often occurs after an iconically lengthened first IU. It is characterised by a high register and a steeper pitch decline. It may also have [fall] underlying targets at the right edge.

Other NPs may form an IU of their own in second position, such as reactivated topics. They are set off by a pitch reset, have a gently falling contour with low mean pitch values throughout, and diminished pitch excursions.

As to interrogatives, the falling contour is most common, accounting for nearly 60% of the tokens in the dataset. The fall-rise, marked by a rise on the last syllable, is also frequent, found in more than 30% of the tokens, while the rising contours account for the remaining tokens. The fall-rise and rising contours are fairly evenly distributed between polar and information questions. The two types of interrogatives are distinguished in that most polar questions have more marked values than information questions. The three contours, when they are uttered in interrogatives, use a slightly higher pitch register than declaratives with the same contours.

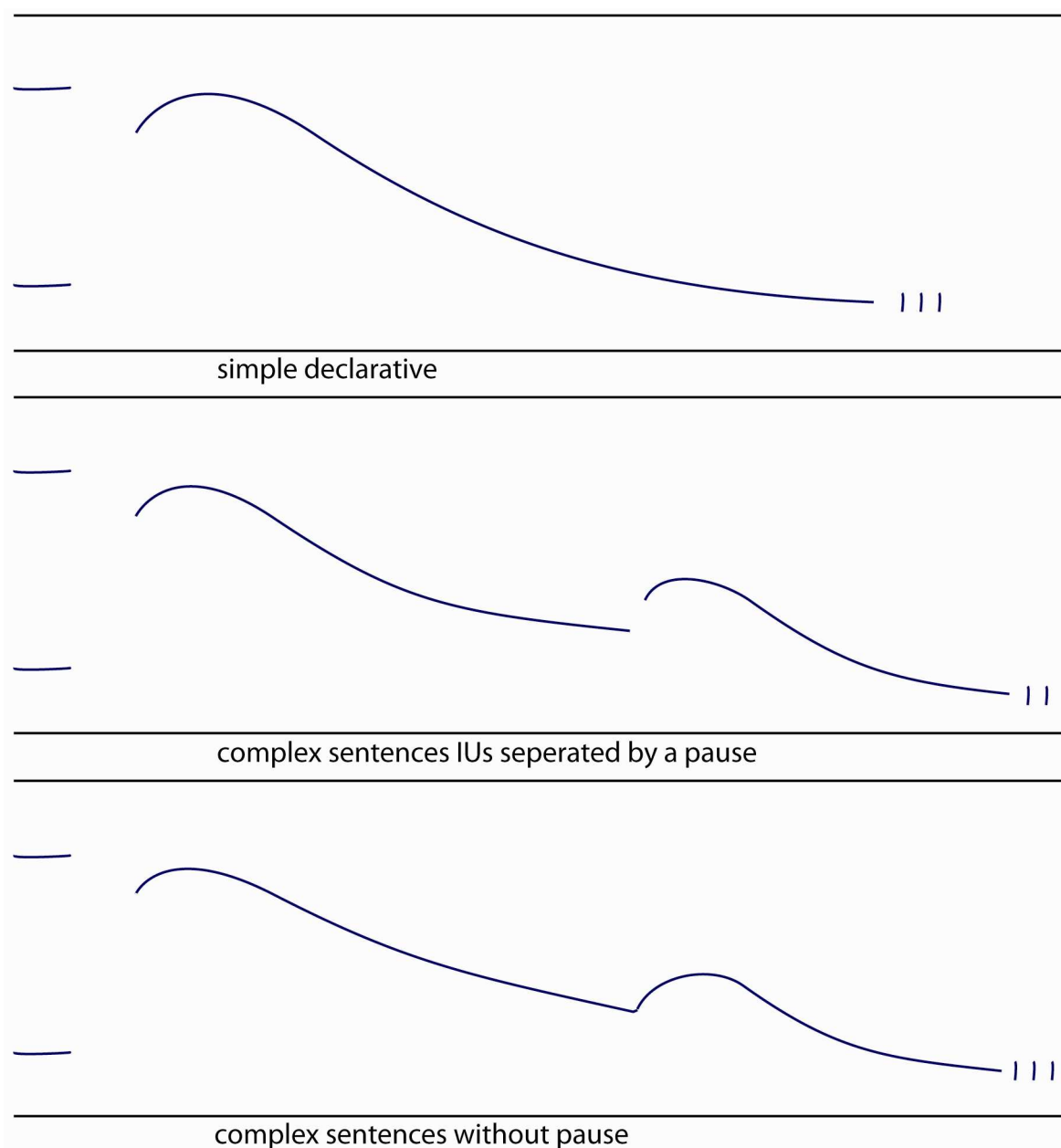
Tag questions are common; *ngih* and *init*, meaning ‘isn’t it’, have a rising contour, and *gurra*, meaning ‘right?’, has a falling contour.

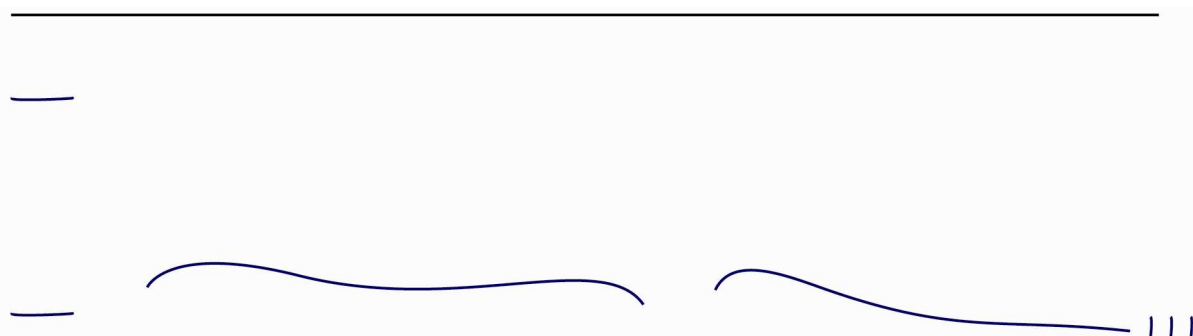
Imperatives are also uttered with the default declarative contour, but with the highest pitch register. They have wider excursions on all their syllables, and the prominence at their left boundary is more salient than that of the other types.

Finally, vocatives are usually uttered with a steep rising contour, evidenced by the large excursions, most notably on their last syllable. The final velocity values indicate a [high] underlying target. Quotative

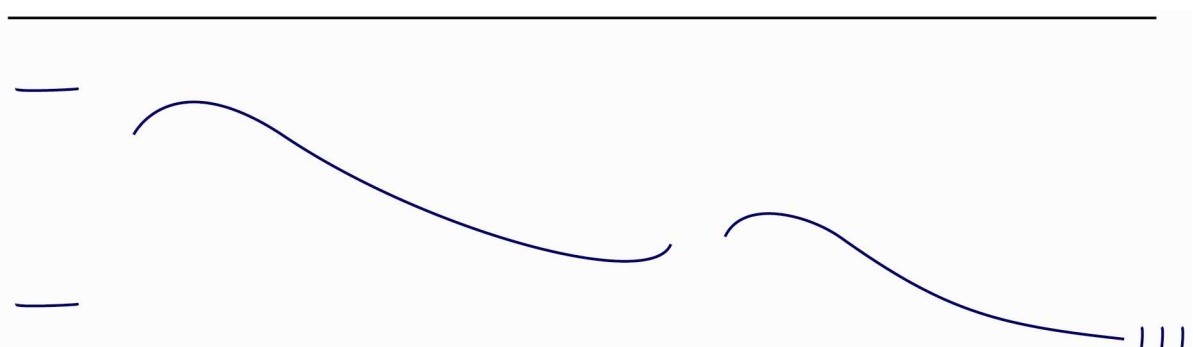
expressions for their part tend to have a default declarative pattern, but are uttered with a raised pitch register, wider pitch excursions, and a considerably lengthened final syllable. The surrounding IUs, particularly those that contain a quotative verb, are uttered with a low register, whether they occur before or after the quoted material.

A schematic representation of these patterns is provided hereafter. Markers on the left indicate the ‘normal’ pitch range. The three short vertical lines at the right edge indicate creakiness.

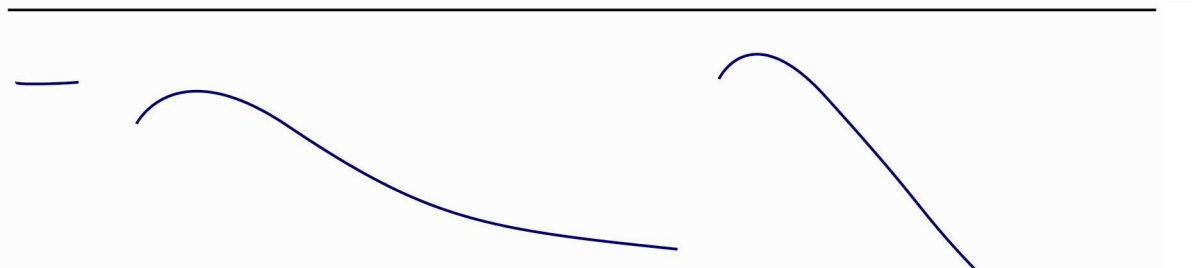




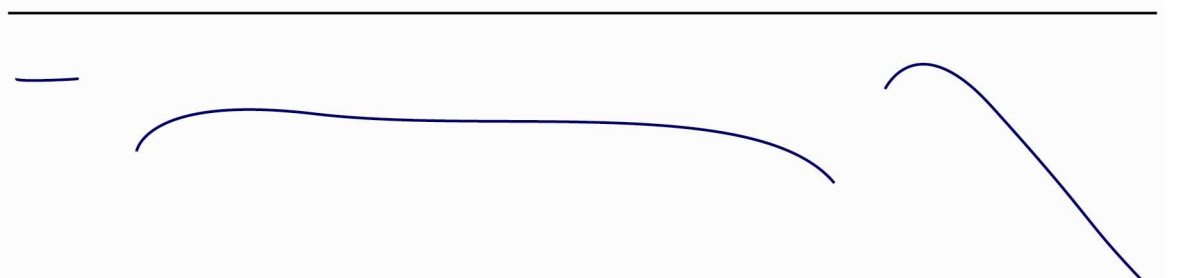
complex sentences first IU level



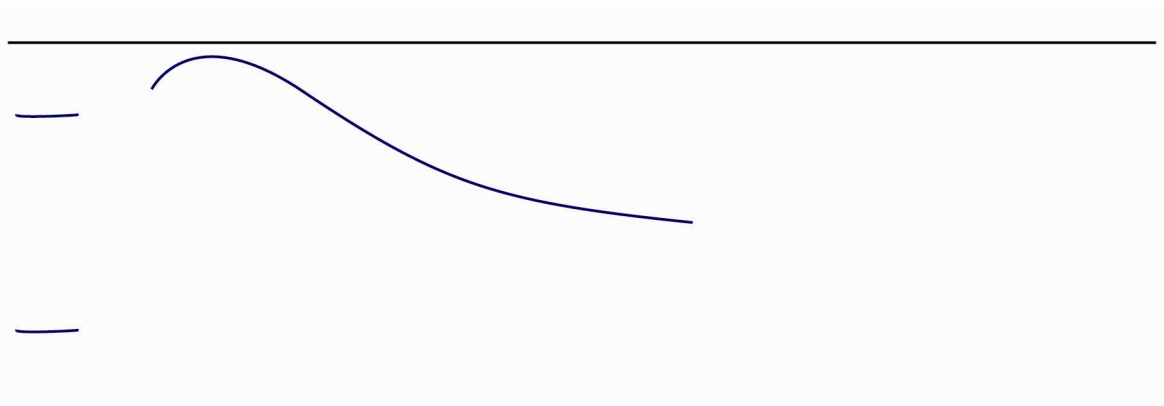
complex sentences first IU fall-rise



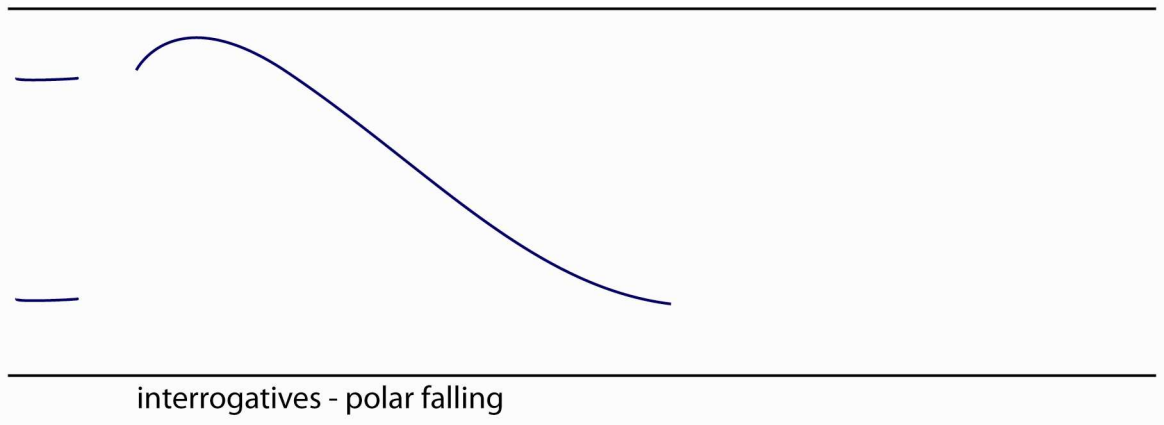
complex sentences first IU falling - second IU secondary prosodic IU



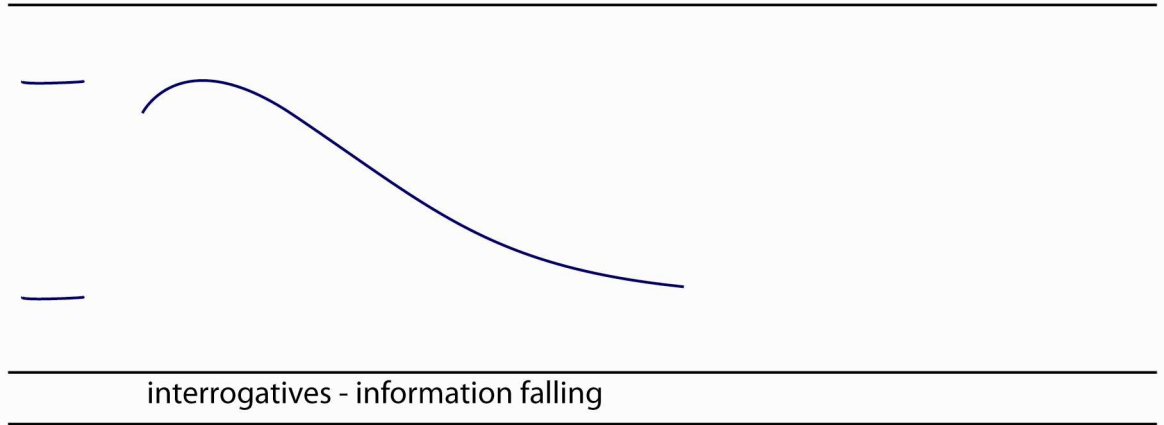
complex sentences first IU lengthening - second IU secondary prosodic IU



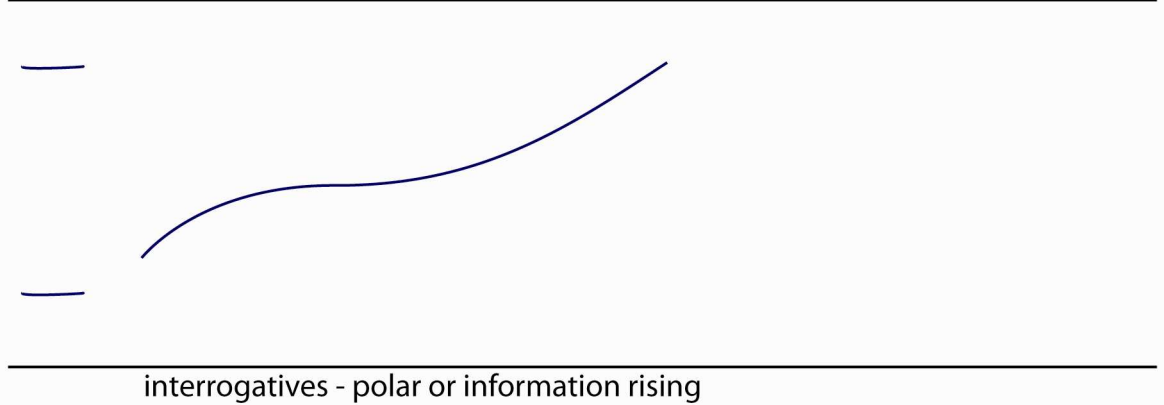
imperatives - falling with higher register



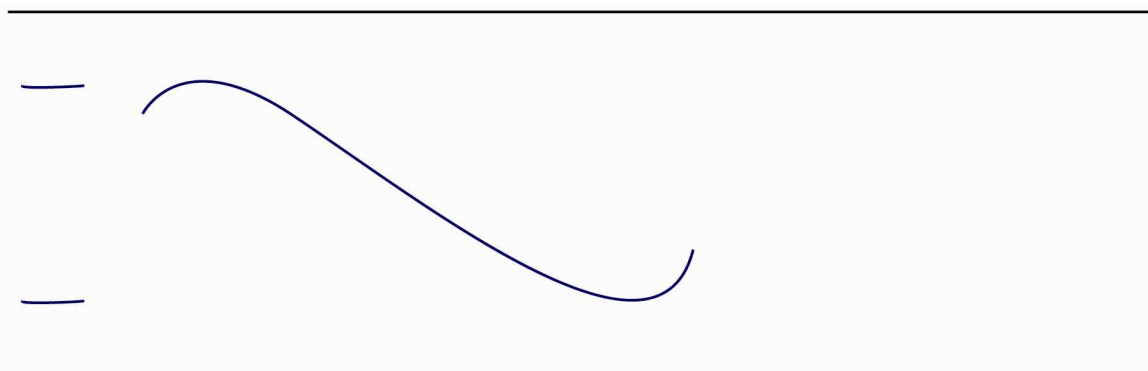
interrogatives - polar falling



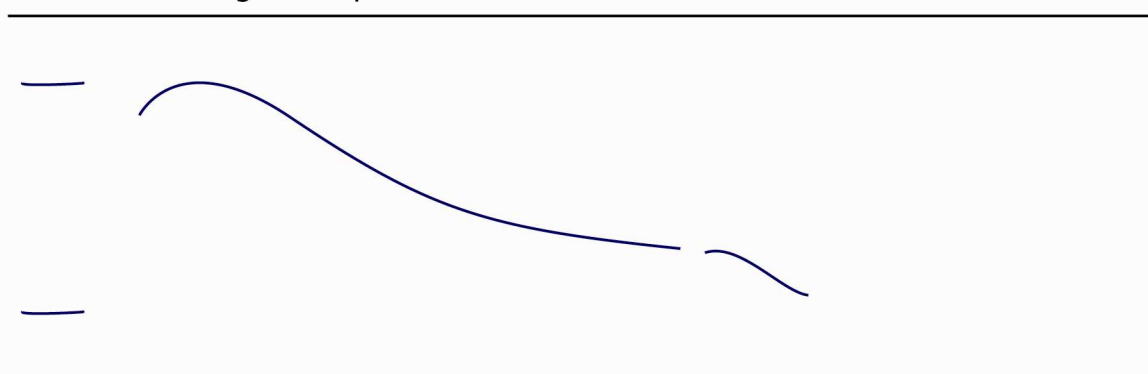
interrogatives - information falling



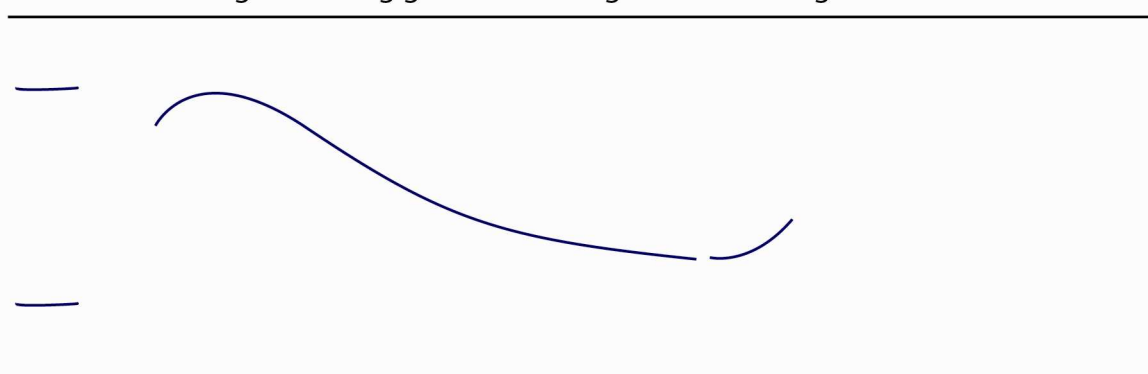
interrogatives - polar or information rising



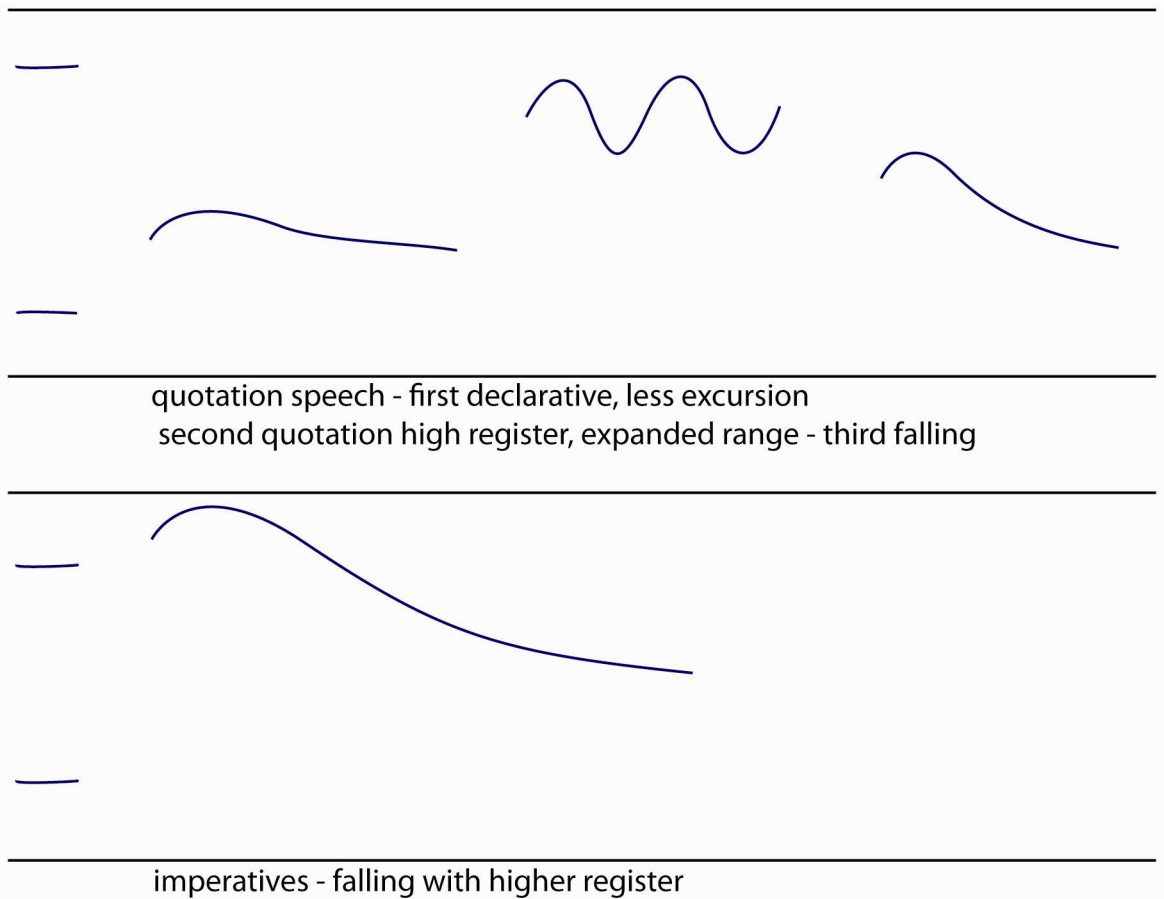
interrogatives - polar or information fall-rise



interrogatives - tag gurra first falling - second falling



interrogatives - tag ngih and init - first falling - second rising



Chapter 6 Conclusion: the prosodic contours of Jaminjung

Prosody is one of the fundamental components of language. In the last decades, a renewed interest in intonation studies has provided important insights into its workings as well as its interactions with other linguistic components. Because some aspects of its realization are specific to every language, we stand to learn a lot by examining typologically diverse languages. In recent years, interest has grown in documenting and describing intonational phenomena in lesser-known or endangered languages. As a case-study, this examination of the intonational patterns of Jaminjung helps our understanding of the language itself, and that of prosodic patterns amongst Australian languages, and others that are shared by most of the languages studied so far.

In this chapter, I will first appraise the model used in this study; I will then summarize the main findings of the previous chapters and their implications for Australian linguistics and intonational typology. I will point to areas worth considering in further studies along the way.

The PENTA model (Parallel Encoding and Target Approximation) has been used to make this description of the intonational patterns of Jaminjung. It is the first time the model has been used to provide an overall description of the intonational patterns in one language; it is also the first time it has been used with unscripted speech, as made necessary in the Jaminjung context.

The model has two major advantages: firstly it accounts for the fact that intonation and its most obvious manifestation, the surface F0 contour, serve many functions *at once*; secondly, it provides tools for a quantitative evaluation of the prosodic parameters used to encode communicative functions, with no a priori decision as to what is or is not valid information in the signal. This avoids the risks of a description based on a ToBI-like

annotation system⁶⁴, particularly when only one transcriber/annotator (the main researcher) is available, and native speakers can not be expected to participate in the analysis.

PENTA assumes that communicative meanings are at the core of prosody, and that there are biophysical mechanisms that underlie speech production. In the model, observed F0 contours are the outcome of implementing pitch targets, which are synchronized with the syllable, through encoding schemes that specify the melodic primitives (including pitch range, articulatory strength and duration). In this manner, the PENTA model both describes and explains F0 patterns in utterances. Importantly, PENTA does not specify the encoding schemes for a specific language; these need to be discovered through empirical investigations in which potential contributors to surface F0 contours are systematically controlled (Xu 2005:246). This has informed the methodological decisions in this research: in a first step, communicative functions have been delineated, then clear examples of their instantiation in Jaminjung were selected, the parameters used in their encodings were then sought out. The PENTA model also makes it possible to apply quantitative methods usually reserved for larger corpora to relatively limited datasets and thus makes patterns more easily discernable and verifiable.

In a way, and as pointed by Frid⁶⁵ the encoding schemes that specify local pitch targets could roughly be compared to a ‘tonal grammar’ in intonational phonology, in the sense that they constitute the link between the communicative functions and the surface F0 contour. However, unlike in intonation phonology, other parameters are also considered, so that pitch range, articulatory strength and duration can also be specified in the encoding schemes. Furthermore, while descriptions based on models such as

⁶⁴ See Wightman (2002) and Dilley et al. (2006) for a discussion/criticism. In short, although reasonably high agreement amongst labelers has been found, systematic disagreements exist, indicating that some intonational events may be harder to identify, or possibly flaws in the methodology.

⁶⁵ From the Function- and production-based modeling of Swedish prosody project <http://www.ht.lu.se/o.o.i.s?id=23922&p=212>

the AM theory may perform reasonably well in describing surface F0 contours, they lack in explanatory power; with no obvious link between phonetic and semantics; indeed while the AM model's ultimate aim is to link the grammar (phonological) to the communicative functions of prosody, this is usually not an integral part of the model.

Some of the findings reported in this research, however, pose some challenges to the model. In the encoding of focus for instance, the length of the word in focus impacts on the encoding, suggesting that the domain of the encoding scheme could be the prosodic foot instead of the syllable. This in fact does not go against the basic principle of parallel encoding via articulatory execution that is at the core of the model as PENTA has no restriction on the temporal domain of an encoding scheme. Rather, the restriction on the syllable as a temporal domain is for local pitch targets and is based on empirical findings.

Further, the present results may be interpreted as helping to clarify the nature of targets under PENTA. They are phonetically distinguished through a combination of F0 characteristics, including slope, peak and valley alignment, global pitch range, and relative height. In this research, the measurement of final velocity was used for the first time with unscripted speech data, and has shown to be a fairly reliable indicator of the underlying pitch targets with the Jaminjung data, although it still needs to be used in conjunction with auditory perception and a visual inspection of the pitch track. Further experimentation and comparable analyses with other languages may continue to strengthen its validity, notably with taking the measurements at 30ms before syllable offset.

Importantly, some of the findings reported here could not have been captured if other models had been used, highlighting the descriptive strength of the PENTA model. This is most strikingly exemplified with the difference in register between the three main sentence types, declaratives, interrogatives, and imperatives, or the specific [fall] target and its alignment associated with focus. Another aspect is that of 'pitch scaling', understood here in the sense of variations in F0 excursions, which plays a part in the encoding of contrastive categories in Jaminjung. In the AM model, intonation consists of linearly concatenated local tones, and surface F0 comes

from interpolation between these local tones. The model only recognizes local pitch accents and edge tones as phonological units, and ‘pitch scaling’ is deemed not to have a phonological role.

This study of intonational phenomena in Jaminjung assumed three main functions of intonation: chunking the flow of speech into units of various lengths, marking information structure categories, notably topic, focus and contrast, and distinguishing sentence types.

In Chapter 3, I examined the prosodic constituents in Jaminjung. Four units are proposed. Prosodic words, phrasal constituents, IUs and prosodic sentences. This study is unusual in prosodic research in its examination of a unit larger than the IU. It has shown that particular contours coincide with the final boundary of prosodic sentences i.e. the [low] targets, and further that larger units indeed establish coherence between ‘primary’ and ‘secondary’ units (chapter 5).

More specifically, prosodic words, which consist of content words and their associated bound morphemes, are the domain of stress which falls on their first syllables, cued by longer duration and higher intensity. Phrasal constituents, which correspond to NPs, complex predicates, non-finite predicates, and particle groups, are set off by a small pitch reset, with the consequence that the declination line appears to be halted. They also display some final lowering at their right-edge, which is not associated with phonation events. The two larger units are the IUs, which correspond for the most part to clauses, but can also be dislocated elements, or interjections; and the prosodic sentences, which may consist of semantically connected independent clauses, or a main clause with a subordinate clause or dislocated element. IUs and prosodic sentences are both set off by substantial pitch resets, smaller in IUs than in sentences. At the right boundary, both larger units are lengthened; pauses after IUs are usually shorter than those after sentences. Both IUs and prosodic sentences have final lowering, their final syllables receiving a [low] target. The right-edge lowering is not, however, without exceptions as reported in chapter 5. Speakers sometimes use breathy phonation at the end of IUs, and creaky phonation at the end of prosodic sentences.

In short, the units are distinguished at their left boundaries by pitch resets which increase from unit to unit. This reset accounts for some of the observed pitch movements on their first syllables (except in prosodic words which are not reset), here the [high] target associated with the grouping function. Note, that this is not the sole pitch event that coincides with the left boundary, as a [fall] target is associated with the marking of focus as shown in chapter 4. The correlates observed at the right edge of the larger constituents, F0 lowering and syllable lengthening are cues usually associated with finality in many languages. Then again, as the measures display a gradience, it could be argued that prosody does not mark finality in a categorical manner. The strongest cue for the identification of the units is the affinity index, which clearly delimits the four units. The values of the IOI (inter onset interval) of the units indicate that stronger dependencies between immediately adjacent units correlate with a weaker boundary between them. Phenomena at the word and phrase level would warrant further investigation, particularly the encoding of clitics as a group distinguished from affixes, and between specific members of the group.

As current typological classifications of intonation are mainly based on word prosodies that is, whether words can contain lexical stress, pitch-accents, or lexical tones, these distinctions forming the criterial definition in the classification of the tonal structure of languages leading to classes such as intonation languages, tone languages and possibly also pitch-accent languages (Ladd:1996, 2008). In this sense, the findings for Jaminjung are in accordance with those of other Australian languages: there is no evidence of lexical tone or pitch accents, although ‘stress’ occurs on the first syllable, a feature which is interpreted in this context as serving as a marker of the ‘grouping’ function and which places Jaminjung amongst intonation languages.

In chapter 4, I investigated the encoding parameters of some major information structural categories. This study is the first to provide such an in-depth prosodic analysis of IS categories in an Australian language. Three focus domains and two major topic subtypes, and the notion of contrast were examined. Some attention was given to information structure-related clitics such as =*marlang* ‘GIVEN’ and =*gun* ‘CONTRAST’.

As for the realization of focus, a prominence is heard on the first syllables of the initial word of the construction in focus. In comments, it is often the coverb, thus signaling its greater semantic weight. The constituent in focus usually is in 1st position in an Intonation Unit (but following any topical constituent), which accounts for much of the word order variation observed in Jaminjung. A [fall] pitch target is associated with the first syllables of focus constructions. Its alignment may vary with the length of the focused element, thus longer words may have a [high] on the first syllables, the [fall] occurring only on their second syllables. Prominent syllables are also marked by wider pitch excursions and longer durations.

This research has paid special attention to the encoding of focus according to its scope. A comparative analysis of three focus domains was conducted, including argument, comment and sentence focus. Such analyses are rarely done in prosodic studies, most particularly concerning sentence focus which is rarely researched and has not often been considered in works on Australian languages. The results showed a gradation in the prosodic correlates according to the scope of focus, from the least marked sentence focus domain to the most marked argument focus domain with the comment focus domain in between. Most notably, the sentence focus domain seems to lack a clear prominence.

These results conform to analyses of other Australian languages in establishing the IU-initial, or at least pre-verbal, position (following any topical constituents) as the position of focused constituents (and the locus of prominence in when the focused constituent includes the verb), and in the use of an expanded pitch range in focus marking. It has not been possible, because of the spontaneous nature of the speech in the datasets used in this study, to determine whether the expanded pitch range is coupled with a post-focus compression, as noted in typologically diverse languages such as Mandarin and English. The assignment of a [fall] pitch target on the prominent syllable in the focused constituents in Jaminjung is also consistent with the patterns in other Non-Pama-Nyungan languages, but unlike the Central Australian language Warlpiri, for which no culminative marking is described. The specific marking of argument focus (or narrow focus as it is often referred to) often distinguishes between languages that

mark narrow focus through word order (without necessarily changing intonation pattern), such as Spanish and Italian, and those that mark it through intonation alone (i.e. without a focal word order), such as English or German (Ladd 1996). Jaminjung falls somewhat in between these major classes, in that the argument in focus is usually moved into the first position, the favoured position for focus marking by prosodic prominence, but the parameters used to encode its prominence are also more salient, placing Jaminjung amongst languages that use both mechanisms of focal marking.

Other information structure categories tested included topics and contrast. Topics do not receive a prominence and are marked by a [high] target on their initial syllables. The prosodic encoding of topics follows a scale of ‘givenness’, where more given topics are less marked than less given topics. The presence of the clitic =*marlang* ‘GIVEN’ correlates with lowered mean pitch of the entire topic constituent, wherever it occurs. Contrast in focused arguments and topics is encoded with a [fall] target on the initial syllable and with a wider pitch excursion on all the syllables of the contrastive element. This last feature marks contrast as an independent information structure category from focus and topic. Finally, contrastive elements which also have the clitic =*gun* ‘CONTRAST’, have less marked values in their prosodic correlates.

The results reported for the clitics =*marlang* and =*gun* lead me to suggest that the use of morphological markers influences the prosodic encoding parameters in Jaminjung, either in attenuating or accentuating them.

In chapter 5, sentence types were examined, including declaratives, interrogatives and imperatives; specific constructions such as the iconically lengthened constructions, vocatives and reported speech were also examined.

IUs displaying the ‘default’ declarative contour are set off by a pitch reset, followed by an overall decline in pitch; they usually, but not always, have a prominence on the first syllable (associated with the encoding of focus). The syllables at their left boundaries receive targets that are related to the information structure of the IU. They are followed by a sequence of [mid] targets, and demarcated by a final [low] target.

When complex declarative sentences consist of a succession of verbal clauses, the most common pattern is of a repetition of the default declarative contour, with a continuation of the line of declination throughout the sentence. Their IUs are usually bounded by pauses, but they may adjoin each other without any break. There are some variations to this pattern. For example, the first IUs may have a fall-rise or a rising contour. In the fall-rise contour, the pitch falls as in a default declarative contour, to rise sharply in the final syllable. These syllables receive a [rise] pitch target. In the rising contour, the rise starts at the beginning of the IU-final word, and the [rise] target is realized from the penultimate syllable. First IUs can also have a level contour, marked by lower pitch values and lesser pitch excursions throughout. The final syllables in level contours have [mid] underlying pitch targets and longer durations than those in both falling and fall-rise contours.

IUs that are iconically lengthened are not elongated level contours. At their left edge, they have pitch patterns that are like those of the declarative contour. Their pitch is then maintained at that level, thus giving the impression of a 'high level contour'. They have [mid] targets on the lengthened segment, including the final syllable.

A specific contour is described for some IUs in second position within a larger unit (prosodic sentence) consisting of two IUs. This contour, termed the secondary prosodic contour, is characterised by a higher pitch register and a steeper decline, possibly reflecting a [fall] underlying target in the final syllables. It usually coincides with non-finite predicates and afterthoughts (but may also occur in short verbal clauses) which functionally bring in new information related to the main clause, either delimiting the event of the main predicate in time and space, narrowing event type, or specifying one of the referents. This contour often occurs after an iconically lengthened first IU.

As to interrogatives, the falling contour is most common. The fall-rise, marked by a rise on the last syllable, is also frequent, while overall rising contours occur but are more rare. Polar and information questions can both have fall-rise or rising contours. The two types of interrogatives are distinguished in that the correlates in polar questions are more marked, with wider pitch excursions and a higher pitch register. Tag questions are

common, *ngih* and *init*, meaning ‘isn’t it’, have a rising contour, and *gurra*, meaning ‘right?’, has a falling contour.

Imperatives are also uttered with the default declarative contour, but with a high pitch register. They have wide excursion on all their syllables, and a salient prominence at their left boundary.

The comparison between sentence types yields an important result: although declaratives, interrogatives and imperatives are all predominantly uttered with a falling contour, they are clearly differentiated by pitch register – declaratives use lower reaches, imperatives higher reaches, and interrogatives somewhere in between. There are some other differences in the sentence types investigated. Imperatives and polar questions with a falling contour have steeper falling contours. The fall-rise and rising contours are found in interrogatives of both types, as well as in the first IUs of complex declaratives, but are not attested in imperatives. These contours could be interpreted as continuation rises, but they are not being used systematically in this function, and the lack of other clear motivations for their use would warrant further research.

Finally, vocatives are usually uttered with a steep rising contour. Quotative expressions, for their part, have patterns similar to that of the default declarative, but are uttered with a raised pitch register, wider pitch excursions, and a considerably lengthened final syllable. The surrounding IUs, particularly those that contain a quotative verb, are uttered with a low register, whether they occur before or after the quoted material.

These results could be corroborated and augmented with the inclusion of the correlate of intensity, which could not be measured in a satisfactory manner given the recording conditions of the materials used in this study. It is probably used in the encoding of stress and contrast, as well as in imperatives and reported speech.

Regarding the intonational events associated with sentence types, Jaminjung proves to be another language in which information requests and questions are not marked with rising intonation at the right-edge of the IU, a situation that has been reported in other Australian languages and in many African languages. The use of different registers to differentiate between

imperatives, declaratives and interrogatives has not been systematically investigated in other languages so far, to our knowledge.

This analysis of the intonation system of Jaminjung thus contributes to intonational typology by supplementing the existing knowledge with data from the vastly underrepresented Australian languages.

References

- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Alsina, A., J. Bresnan & P. Sells (eds). 1997. *Complex predicates*. Stanford: CSLI.
- Ameka, Felix. 1992. Interjections: *The universal yet neglected part of speech*. *Journal of Pragmatics*, 18: pp. 101-118.
- Amir, N. & V. Silver-Varod & S. Izreel, 2004. *Characteristics of Intonation Unit Boundaries in Spontaneous Spoken Hebrew – Perception and Acoustic Correlates*. In B. Bel & I. Marlien (eds.), *Speech Prosody 2004*, Nara, Japan, March 23-26, 2004: Proceedings, SProSIG (ISCA Special Interest Group on Speech Prosody). pp. 677-680.
- Anderson, A., M. Bader, E. Bard, E. Boyle, G. Doherty, M. Garrod et al. 1991. *The HCRC Map Task Corpus*. *Language and Speech*, 34, pp. 351-366.
- Aronoff, Mark & Janie Rees-Miller (eds). 2001. *The handbook of linguistics*. Blackwell Handbooks in Linguistics. Oxford: Blackwell.
- Auer, P. 1996. *On the Prosody and Syntax of Turn-continuations*. In E. Couper-Kuhlen & M. Selting (eds), *Prosody in conversation: Interactional studies*. Cambridge: Cambridge University Press. pp. 57-100.
- Austin, J. L. 1962. *How to Do Things With Words*. Oxford University Press: Oxford, England.
- Austin, Peter. K. 2001. *Word order in a free word order language, the case of Jiwarli*. In Jane Simpson & David Nash & Mary Laughren & Peter Austin & Barry Alpher (eds), *Forty years on: Ken Hale and Australian languages*. Canberra: Pacific Linguistics. pp. 205-323.
- Austin, Peter K. & Joan Bresnan. 1996. *Non-configurationality in Australian Aboriginal languages*. *Natural Language and Linguistic Theory* 14. pp.215-268.
- Averintseva-Klisch, Maria (to appear.). *German Right Dislocation and Afterthought in Discourse*. In Sassen, Claudia & Benz, Anton & Kühnlein, Peter (eds), *Constraints in Discourse*.
- Baumann, Stefan & Martine Grice & Susanne Steindamm. 2006. *Prosodic Marking of Focus Domains – Categorical or Gradient?* *Proceedings of Speech Prosody 2006*. Dresden, Germany. pp. 301-304.

- Beaver, D & E. Flemming & F. Jaeger & M. Wolters. 2004. *When semantics meets phonetics: Acoustical studies of second occurrence focus*. Cambridge: Cambridge University Press.
- Beck, David. 1999. *Words and prosodic phrasing in Lushootseed narrative*. In T. A. Hall & U. Kleinhenz (eds), *Studies on the Phonological Word*. Amsterdam: Benjamins. pp. 23-46.
- Beckman, Mary E. 1996. *The Parsing of Prosody. Language and Cognitive Processes*. pp. 17-67.
- Beckman, Mary & E. Ayer & M. Gayle. 1997. *Guidelines for ToBI Labelling*. Copyright (1993). The Ohio State University Research Foundation. http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/ToBI/ToBI.1.html (accessed 02-01-2009)
- Beckman, Mary & Janet Pierrehumbert. 1986. *Intonational structure in Japanese and English*. *Phonology*, 3. pp. 255-309.
- Bell L. & J. Gustafson. 1999. *Utterance types in the August System*. Proceedings of the 'ESCA Workshop on Interactive Dialogue in Multi-Modal Systems'. pp. 81-84.
- Birch, Bruce. 2003. *A tour of Iwaidja intonation*. Paper presented at the 'Australian Linguistics Society Conference'. Newcastle, UK. September 26-28, 2003.
- Birch, Bruce. 2002. *The IP as domain of Syllabification*. Proceedings of 'Speech Prosody 2002'. Aix-en-Provence, France. April 11-13, 2002.
- Bishop, J.B. 2003. *Aspects of intonation and prosody in Biniŋ Gun-wok: an autosegmental-metrical analysis*. Department of Linguistics and Applied Linguistics, University of Melbourne: Ph D thesis.
- Bishop, J.B. & J. Fletcher. 2005. *Biniŋ Gun-Wok Intonation*. In Sun-Ah Jun (ed.), *Prosody and typology – The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. pp. 331-361.
- Blake, B. 1987. *Australian Aboriginal Grammar*. London: Croom Helm.
- Blanche-Benveniste, C. 1990. *Le français parlé, études grammaticales*. In *Macrosyntaxe* éd. du CNRS, coll. Sciences du langage, Paris.
- Bloomfield, Leonard. 1933. *Language*. Revised from 1914 edition. New York: Holt. 564 p.

- Blythe J. 2009. *Doing referring in Murriny Patha conversation*. Unpublished PhD dissertation, University of Sydney.
<http://hdl.handle.net/2123/5388>
- Blythe, J. 2009. *Prosodic person reference in Murriny Patha reported interaction*. In D Barth-Weingarten, N Dehé & A Wichman (eds.), *Where Prosody Meets Pragmatics*. Bingley, UK: Emerald. pp. 23-52
- Bolinger, D. D. 1986. *Intonation and Gesture*. *American Speech*, 58.2, pp. 156-174.
- Bolinger, D. D. 1972. *Intonation: Selected Readings*. Harmondsworth, Penguin Books.
- Bolinger, D. D. 1961. *Contrastive Accent and Contrastive Stress*. *Language* 37. pp.83-96.
- Bolt, Janet E. & W.G. Hoddinott & F.M. Kofod. 1971. *An elementary grammar of the Ngaliwuru language of the Northern Territory*. Canberra: Australian Institute of Aboriginal Studies (Unpublished Ms.).
- Braun, Bettina & Robert D. Ladd. 2003. *Prosodic correlates of contrastive and non-contrastive themes in German*. In EUROSPEECH-2003. pp. 789-792.
- Brazil, D.1997. *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press. First published in 1985 as Discourse Analysis Monograph No. 8, Bleak House Books and English Language Research, University of Birmingham.
- Brown, Gillian & George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Brown, Gillian & Karen L. Curie & Joanne Kenworthy. 1980. *Questions of intonation*. Baltimore: University Park Press Baltimore.
- Büring, Daniel. 2009. *Towards a typology of focus realization*. In M. Zimmermann & C. Féry (Eds.), *Information Structure: Theoretical, Typological, and Experimental Perspectives*. pp. 177-205.
- Büring, Daniel. 1999. *Focus -- Linguistic, Cognitive, and Computational Perspectives*. In Bosch, Peter & Rob van der Sandt (eds). CUP. pp.142-165 (revised version of the paper in Bosch, P. & R.v.d.Sandt (eds) (1994) *Focus & Natural Language Processing*. Vol.2. IBM Heidelberg. pp. 271-280.
- Büring, Daniel. 1997. *The Meaning of Topic and Focus - The 59th Street Bridge Accent*. London: Routledge.
- Busby, P. A. 1980. *The Distribution of Phonemes in Australian Aboriginal Languages*. *Pacific Linguistics*, 14 (A-60). pp. 73-139.

Butcher, A. (to appear). *Australian Aboriginal languages: consonant-salient phonologies and the 'place-of-articulation imperative'*.

Butcher, A. & J. Harrington. 2003. *An instrumental analysis of focus and juncture in Warlpiri*. Proceedings of the '15th International Conference of Phonetic Sciences'. Barcelona, Spain. pp. 321-324.

Calhoun, Sasha & Nissim Malvina & Mark Steedman & Jason Brenier. 2005. *A framework for annotating information structure in discourse*. In *Frontiers in Corpus Annotation II: Pie in the Sky*, ACL2005 Conference Workshop, Ann Arbor, Michigan, June 2005.

Cambier-Langeveld, T. 1997. *The domain of final lengthening in the production of Dutch*. In: J. Coerts & H. de Hoop (eds), *Linguistics in the Netherlands*.

Carmichael, Lesley. 2003. *Intonation: categories and continua*. Paper presented at the 19th Northwest Linguistics Conference, March 1-2, 2003, Victoria, BC, Canada.

Catford, J. 1964. *Phonation types: the classification of some laryngeal components of speech production*. In D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott.

Chadwick, Neil. 1997. *The Barkly and Jaminjung languages: a non-contiguous genetic grouping*. In Tryon, Darrell T. & Michael Walsh (eds), *Boundary rider: Studies in the lexicology and comparative linguistics of Australian languages*. Canberra: Pacific Linguistics. pp. 95-106.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: The University of Chicago Press.

Chafe, Wallace. 1988. *Punctuation and the Prosody of Written Language*. *Written Communication*, 5. pp. 395-426.

Chafe, Wallace. 1987. *Cognitive constraints on information flow*. *Coherence and Grounding in Discourse*, ed. by Russel Tomlin. Amsterdam : John Benjamins.

Chafe, Wallace. 1980. *The deployment of consciousness in the production of narrative*. In W. Chafe (ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Chafe, Wallace. 1976. *Givenness, contrastivness, definiteness, subjects, topics and point of view*. In C.N. Li (ed.), *Subject and Topic*. New York: Academic Press. pp. 25-55.

- Chafe, Wallace. 1970. *Meaning and the Structure of Language*. Chicago: The University of Chicago Press.
- Chapman, Jeffrey. 2007. *Some Aspects of Intonation in Warlpiri*. Paper presented at the 'Australian Languages workshop'. University of Sydney, Pearl Beach, Australia. March 16-18 2007.
- Chavarría, S & T.Yoon & J. Cole & M.Hasegawa-Johnson. 2004. *Acoustic differentiation of ip and IP boundary levels: comparison of l- and l-l% in the switchboard corpus*. In SP-2004. pp. 333-336.
- Chen, S.-W. & B. Wang. & Y. Xu. 2009. *Closely related languages, different ways of realizing focus*. Proceedings of 'Interspeech 2009'. pp. 1007-1010.
- Chen, Y. & Y. Xu. 2006. *Production of weak elements in speech - Evidence from f0 patterns of neutral tone in standard Chinese*. *Phonetica*, 63. pp. 47-75.
- Chen, A. & C. Gussenhoven & A. Rietveld. 2004. *Language-specificity in perception of paralinguistic intonational meaning*. *Language and Speech*, 47(4). pp. 311-350.
- Choi, J.-Y. 2003. *Pause length and speech rate as durational cues for prosody markers*. *Journal of the Acoustical Society of America*, 114(4). p. 2395.
- Clark, J. & C.Yallop & J. Fletcher. 2007. *An introduction to phonetics and phonology*. (3rd Edition), Oxford: Blackwell.
- Cole, Jennifer & Mo Yoonsook & Mark Hasegawa-Johnson. 2008. *Frequency and repetition effects outweigh phonetic detail in prominence perception*. Paper presented at the 11th Laboratory Phonology (LabPhon 11), Wellington, New Zealand (June 30-July 2, 2008).
- Cook, Clare & Jeff Muehlbauer. 2005. *Acoustic cues to Nehiyawewin constituency*. *Acoustical Society of America Journal*, Vol. 117, Issue 4. pp. 2491-2491.
- Couper-Kuhlen, Elizabeth. 2001. *Intonation and discourse: Current views from within*. In *Handbook of Discourse Analysis*, Deborah Schiffrin, Deborah Tannen, Heidi Hamilton (eds.) Oxford: Blackwell.
- Couper-Kuhlen, Elizabeth. 1986. *An Introduction to English Prosody*. *Forschung & Studium Anglistik 1*, Tübingen: Max Niemeyer and London: Edward Arnold.
- Cresti, Emanuela & Massimo Moneglia (eds.). 2005. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*, Amsterdam/Philadelphia: Benjamins.

- Croft, William. 1995. *Intonation units and grammatical structure*. Linguistics, 33: 839-882.
- Cruttenden, Alan. 1997. *Intonation*. 2nd edition. New York: Cambridge University Press.
- Crystal, D. 2003. *A dictionary of linguistics and phonetics*. Oxford: Blackwell. 5th edn.
- Crystal, D. 1995. *The Cambridge Encyclopaedia of Linguistics*. Cambridge: Cambridge University Press.
- Crystal, D. 1979, *Prosodic development*. In P. Fletcher and M. Garman (eds.), *Language Acquisition*. Cambridge: Cambridge University Press.
- Crystal, D. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- d'Alessandro, C. 2006. *Voice source parameters and prosodic analysis*. In Stefan Sudhoff & Denisa Leternovà & Roland Meyer & Sandra Pappert & Petra Augurzky & Ina Mleinek & Nicole Richter & Johannes Schliesser (eds), *Method in Empirical Prosody Research*. Berlin, New York: de Gruyter. pp. 63-87.
- Delattre, Pierre. 1966. *Les dix intonations de base du français*. French Review, 40.1-14.
- Dellwo, Volker & Petra Wagner. 2005. *English and French native language influence on the perception of syllable prominence in German*. In PSP2005.
- De Vries, Mark. 2007. *Dislocation and backgrounding*. In Bettelou Los & Marjo van Koppen (eds), *Linguistics in the Netherlands 2007*. Amsterdam, Philadelphia: Benjamins. pp 236-248.
- Di Cristo, Albert & Daniel Hirst. 1998. *Intonation systems : a survey of twenty languages*. Cambridge: Cambridge University Press.
- Dik, Simon. C. 1997. *The Theory of Functional Grammar*. Berlin: Mouton de Gruyter. pp. 313-338.
- Dik, Simon C. 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris.
- Dilley, L. & M. Breen & M. Bolivar & J. Kraemer & E. Gibson. 2006. *A comparison of inter-transcriber reliability for two systems of prosodic annotation: RaP (Rhythm and Pitch) and ToBI (Tones and Break Indices)*. In *Proceedings of the International Conference on Spoken Language Processing*, Pittsburgh, PA.

Dipper, Stefanie & Michael Götze & Stavros Skopeteas. 2007. *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure*, Vol. 7 of Interdisciplinary Studies on Information Structure (ISIS). Universitätsverlag Potsdam, Potsdam, Germany.

Dixon, R. M. W. 2002. *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1972. *The Dyrbal Language of North Queensland*. Cambridge: Cambridge University Press.

Downing, L. 2006. *The Prosody and Syntax of Focus in Chitumbuka*. ZAS Papers in Linguistics, 43. pp. 55-79.

Drubig, Hans Bernhard & W. Schaffar. 2001. *Focus constructions*. In Haspelmath, M. & E. König & W. Oesterreicher & W. Raible (eds), *Language Typology and Language Universals (Handbook)*. Berlin: de Gruyter. pp. 1079-1104.

DuBois, John W. & Stephan Schuetze-Coburn. 1993. *Representing hierarchy: Constituent structure for discourse databases*. In Jane A. Edwards & Martin D. Lampert (eds), *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum. pp. 221-260.

DuBois, John W. & Susanna Cumming & Stephan Schuetze-Coburn & Danae Paolino (eds). 1992. *Discourse Transcription*. Volume 4 of the Santa Barbara Papers in Linguistics.

Dutoit, Thierry. 1997. *An introduction to text-to-speech synthesis*. Kluwer Academic Publishers, Norwell, MA.

Eades, Diana. 1988. *They don't speak an Aboriginal language, or do they?*. In Keen, Ian (ed.), *Being Black: Aboriginal Cultures in Settled Australia*. Canberra: Aboriginal Studies Press. pp. 97-117.

Eipstein, Melissa, 2002. *Voice Quality and Prosody in English*. University of California, Los Angeles. Ph.D Dissertation
<http://www.linguistics.ucla.edu/people/grads/melissa/>

El Zarka, Dina. (2008). *Leading, linking, and closing tones and tunes in Egyptian Arabic - what a simple intonation system tells us about the nature of intonation*. To appear in: *Perspectives on Arabic Linguistics XXII, XXIII*. Amsterdam/Philadelphia: John Benjamins.

Evans, Nicholas. 2009. *Some problems in the typology of reported speech: a canonical approach*. Paper presented at the symposium 'Creating Infrastructure for Canonical Typology'. University of Surrey, UK. January 2009.

- Evans, Nicholas. 2003. *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*. Canberra: Pacific Linguistics.
- Ewing, M. 2005. *Hierarchical constituency in conversational language: The case of Cirebon Javanese*. *Studies in Language* 29.1. pp. 89-112.
- Féry, C. & G. Fanselow & M. Krifka. 2007. *The Notions of Information Structure*. In *Interdisciplinary Studies on Information Structure 6, 2007*. pp. 1-12.
- Féry, C. & Ishihara. 2009. *How Focus and Givenness Shapes Prosody*. In Zimmermann, M. & Féry, C. (eds), *Information structure from different perspectives*. Oxford: Oxford University Press. pp. 36-63.
- Flemming, E. 2008. *The role of pitch range in focus marking*. Slides from a talk given at the Workshop on Information Structure and Prosody, Studiecentrum Soeterbeeck.
- Fletcher, Janet. 2007. *Intonation in Dalabon*. Paper presented at 'Intonation of Fieldwork and Less-studied Languages' workshop, ICPhS Saarbruecken, August 5th, 2007.
- Fletcher, Janet & Nicholas Evans & Belinda Ross. 2004. *Pausing strategies and prosodic boundaries in Dalabon*. *Proceedings of the 10th Australian International Conference on Speech Science & Technology*. Macquarie University, Sydney, December 8 to 10, 2004. pp. 436-439.
- Fletcher, Janet & Nicholas Evans & Erich Round. 2002. *Left-edge tonal events in Kayardild (Australian) - a typological perspective*. In *SP-2002*. pp. 295-298.
- Fletcher, Janet & Nicholas Evans. 2002. *An acoustic intonational study of intonational prominence in two Australian languages*. *Journal of the International Phonetic Association*, 32. pp. 123-140.
- Fougeron, C. 1999. *Prosodically conditioned articulatory variations : a review*. *U.C.L.A. Working Papers in Phonetics*, 97. pp. 1-74.
- Fox, Anthony. 2000. *Prosodic Features and Prosodic Structure : The Phonology of Suprasegmentals*. *Oxford Linguistics*. Oxford University Press, Oxford.
- Fried, Mirjam & Jan-Ola Östman. 2004. *Construction Grammar: a thumbnail sketch*. In M. Fried & J-O. Östman (eds), *Construction Grammar in a cross-language perspective*. Amsterdam: Benjamins. pp. 11-86.
- Fry, D. B. 1958. *Experiments in the perception of stress*. *Language and Speech* 1. pp. 126-152.

- Fujisaki, H. 1983. *Dynamic characteristics of voice fundamental frequency in speech and singing*. *The Production of Speech*. P. F. MacNeilage, New York: Springer-Verlag. pp.39-55.
- Gaby, Alice. 2008. *Rebuilding Australia's Linguistic Profile: Recent Developments in Research on Australian Aboriginal Languages*. *Language and Linguistics Compass*, Vol. 2, Issue 1. pp. 211-233.
- Gauthier, B. & R. Shi. & Y. Xu. 2007. *Learning phonetic categories by tracking movements*. *Cognition*, 103. pp. 80-106.
- Genetti, Carol. In press. *Syntax and prosody: Interacting coding systems in Dolakha Newar*. *Proceedings of the Thirteenth Annual Meeting of the Southeast Asian Linguistics Society*, Iwasaki, Shoichi (ed.). Canberra: Pacific Linguistics.
- Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. Mouton Grammar Library. Berlin / New York: Mouton de Gruyter
- Gibbon, Dafydd & Richter Helmut. 1984. *Intonation, Accent and Rhythm*. *Studies in Discourse Phonology*. Berlin: de Gruyter.
- Givón, T. 1983. *Topic continuity in discourse: An introduction*. In Givón, T. (ed), *Topic Continuity in Discourse*. Amsterdam: Benjamins. pp. 1–42.
- Goldman-Eisler, F. 1968. *Psycholinguistics: experiments in spontaneous speech*. London, New York: Academic Press.
- Gordon, M. 1999. *The intonational structure of Chickasaw*. *Proceedings of the '14th International Congress of Phonetic Sciences'*. pp. 1993-1996.
- Gordon, M & P. Ladefoged. 2001. *Phonation types: a cross-linguistic overview*, *Journal of Phonetics* 29. pp. 383-406.
- Götze, M. & T. Weskott. & C. Endriss. & I. Fiedler & S. Hinterwimmer & S. Petrova & A. Schwarz & S. Skopeteas & R. Stoel. 2007. *Information Structure*. In Dipper, S. & Götze, M. & Skopeteas, S. (eds), *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure*. Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS). Potsdam: Universitätsverlag Potsdam. pp. 147-187.
- Grice, Martine & Stefan Baumann. 2007. *An Introduction to Intonation – Functions and Models*. In Trouvain, Jürgen & Ulrike Gut (eds.): *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin, New York: De Gruyter (= Trends in Linguistics. Studies and Monographs [TiLSM] 186). pp.25-51.

- Grosjean, F. & L. Grosjean & H. Lane. 1979. *The patterns of silence: Performance structures in sentence production*. Cognitive Psychology 11 pp. 58-51.
- Gundel, Jeanette K. 1988. *Universals of topic-comment-structure*. In Hammond, M. & E. Moravcsik & J.R. Wirth & Jessica R (eds), *Studies in syntactic typology*. Amsterdam: Benjamins. pp. 209-239.
- Gundel, Jeanette K. 1985. *Shared knowledge and topicality*. Journal of Pragmatics, 9.1. pp. 83-97.
- Gundel, Jeanette. K. 1978. *Stress, pronominalization and the given-new distinction*. University of Hawaii Working Papers in Linguistics 10.2 pp. 1-13.
- Gundel, Jeanette. K. & N. Hedberg & R. Zacharski. 1997. *Topic-comment structure, syntactic structure, and prosodic tune*. Workshop on 'Prosody and Grammar in Interaction'. Helsinki. July 1997.
- Gussenhoven, Carlos, 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos. 2002. *Intonation and interpretation: phonetics and phonology*. In B.Bel & I. Marlien (eds), *Proceedings of the 'Speech Prosody 2002 Conference'*. Aix-en-Provence. pp. 47-57.
- Gussenhoven, Carlos. 1999. *On the limits of focus projection in English*. In Focus: Linguistic, Cognitive, and Computational Perspectives, Peter Bosch & Rob van der Sandt (eds). Cambridge: Cambridge University Press. pp. 43-55.
- Hale, Kenneth. 1992. *Basic word order in two "free word order" languages*. In D. Payne (ed.), *Pragmatics of word order flexibility*. Amsterdam: Benjamins.
- Hale, Kenneth. 1983. *Warlpiri and the grammar of non-configurational languages*. Natural Language and Linguistic Theory, 1, 5-47.
- Hale, Kenneth & Elisabeth, Selkirk. 1987. *Government and tonal phrasing in Papago*. Phonology Yearbook, 4. pp.151-183.
- Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London: Arnold, 1985. 2nd ed.
- Halliday M.A.K. 1967. *Notes on Transitivity and Theme in English, Part 2*. Journal of Linguistics 3. pp. 199-244.
- Halliday, M. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton de Gruyter.

- Harrington, J.M. & M. Tabain (eds). 2006. *Speech Production: Models, Phonetic Processes and Techniques*. New York: Psychology Press.
- 't Hart, J. & R. Collier & A. Cohen. 1990. *A perceptual Study of Intonation — An experimental phonetic approach to speech melody*. Cambridge, Cambridge University Press.
- Harvey, Mark David. 2008. *Proto Mirndi: A Discontinuous Language Family in Northern Australia*. Canberra: Pacific Linguistics ACT.
- Haspelmath, Martin & Ekkehard König & Wulf Oesterreicher & Wolfgang Raibl. 2001. *Language typology and language universals: An international handbook*. (Handbücher zur Sprach- und Kommunikationswissenschaft) Vol. 1-2. Berlin: Mouton de Gruyter. 1856 pages.
- Hayes, Bruce. 1989. *The prosodic hierarchy in meter*. In Paul Kiparsky, and Gilbert Youmans (eds). *Phonetics and phonology, volume 1: Rhythm and meter*. Orlando, FL: Academic Press. pp.201-260.
- Heath, Jeffrey G. 1984. *Functional grammar of Nunggubuyu*. Canberra: AIAS/Atlantic Highlands, New Jersey: Humanities Press.
- Hedberg, Nancy & Juan M. Sosa. 2007. *The Prosody of Topic and Focus in Spontaneous English Dialogue*. In Chungmin Lee & Mathew Gordon & Daniel Burin (eds), *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*. Dordrecht: Springer. pp. 101-120.
- Hellmuth S., Kügler & R. Singer. 2007. *Intonational patterns, tonal alignment and focus in Mawng*. Paper presented at ICPhS 2007 Satellite Workshop 'Intonational phonology: understudied and fieldwork languages'.
- Hellmuth, Sam & Frank Kügler & Ruth Singer. 2007. *Quantitative investigation of intonation in an endangered language*. In Peter K. Austin & Oliver Bond & David Nathan (eds), *Proceedings of Conference on Language Documentation and Linguistic Theory*. London: SOAS. pp. 123-132.
- Himmelman, Nikolaus P. 2006. *Prosody in language documentation*. In Gippert, Jost & Himmelman, Nikolaus P. & Mosel, Ulrike (eds), *Essentials of language documentation*. Berlin: Mouton de Gruyter. pp. 163-181.
- Hirschberg, J. & J. Pierrehumbert. 1986. *Intonational Structuring of Discourse*. *Proceedings of the 24th Meeting of the Association for Computational Linguistics*. pp.136-144.
- Hirst, D.J. 2005. *Form and function in the representation of speech prosody*. *Speech Communication*, 46 (3-4). pp.334-347.
- Hirst, D. & A. Di Cristo & R. Espresser. 2000. *Levels of description and levels of representation in the analysis of intonation*. In Hornes M. (ed.)

Prosody : Theory and Experiments. Dordrecht, Netherlands : Kluwer Academic. pp. 51-87.

House, David. 2004. *Final rises and Swedish question intonation*. Proceedings of FONETIK 2004, Dept. of Linguistics, Stockholm University.

Hyman, Larry M. 1990. *Boundary Tonology and the Prosodic Hierarchy*. In Inkelas and Zec. pp.109-125.

Ingram, J. & M. Laughren. 2007. *Gradients of lenition in Warlpiri rhotics and phonetic conditions of sound change*. Presentation at the Workshop 'On The Phonetics And Phonology Of Australian Languages'. 3rd and 4th December 2007 La Trobe University, Melbourne.

Inkelas, Sharon & Draga Zec. 1995. *Syntax-phonology interface*. In The Handbook of Phonological Theory, ed. John Goldsmith. Oxford: Blackwell.

Inkelas, Sharon & Draga Zec. 1990. *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.

Iwasaki, Shoichi & Hongyin Tao. 1993. *A comparative study of the structure of the intonation unit in English, Japanese, and Mandarin Chinese*. Paper presented at the annual meeting of the 'Linguistic Society of America'. Los Angeles, California.

Jackendoff, Ray, 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jannedy, S. 2007. *Prosodic Focus in Vietnamese*. In Ishihara, S. & Jannedy, S. & Schwarz, A. (eds), Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 8 . Potsdam: Universitätsverlag Potsdam. pp. 209-230.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Karlsson, A. & D. House. & D. Tayanin. 2008. *Recognizing phrase and utterance as prosodic units in non-tonal dialects of Kammu*. In Proceedings, FONETIK 2008. Department of Linguistics, University of Gothenburg. pp. 89-92.

King, Heather B. 1994. *The interrogative intonation in Dyirbal*. Proceedings of the 'Australasian International Conference on Speech Science and Technology'. Perth, December 6-8, 1994. pp.144-149.

Kingdon, R. 1958. *The groundwork of English intonation*. London: Longman.

- Kiss, E. Katalín. 1998. *Identificational focus versus information focus*. Language 74,2. pp. 245-273.
- Kneidler, C. 2001. *Phonology: Critical concept in Linguistics*. Routledge: London.
- Knowles, G. 1987. *Patterns of Spoken English: an Introduction to English Phonetics*. London: Longman.
- Knowles, G. & B. Williams & L. Taylor. 1996. *A corpus of formal British English speech: the Lancaster/IBM Spoken English Corpus*. London: Longman.
- Kohler, Klaus J. 2008. *The perception of prominence patterns*. *Phonetica*, 65. pp. 257-269.
- Kohler, Klaus J. 2006. *What is emphasis and how is it coded?* Proceedings of 'Speech Prosody 2006'. Dresden, Germany. May 2-5, 2006.
- Kohler, Klaus, J. 2004. *Prosody Revisited: FUNCTION, TIME, and the LISTENER in Intonational Phonology*. In: B. Bel, I. Marlien (eds.): Proc. International Conference: Speech Prosody 2004, Nara/Japan. pp.171-174.
- Kohler, Klaus .J. 1991. *A model of German intonation*. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 25. pp. 295-360.
- König, E. 1991. *The meaning of focus particles*. London: Routledge.
- Kreidler, Charles W. (ed). 2001. *Phonology: Critical Concepts*. London: Routledge.
- Krifka, Manfred. 2006. *Interdisciplinary Studies on Information Structure*. In Féry, Fanselow and Krifka (eds.).
- Krivokapić, J. 2007. *Prosodic planning: Effects of phrasal length and complexity on pause duration*. *Journal of Phonetics*, 35. pp. 162-179.
- Kügler, Frank. 2008. *The role of duration as a phonetic correlate of focus*. In SP-2008, pp. 591-594.
- Ladd, D. Robert. 2008 (and 1996, 1st ed.). *Intonational Phonology, second edition*. Cambridge University Press.
- Ladd, D. Robert. 2001. *Intonation*. In Haspelmath, Martin & Ekkehard König & Wulf Oesterreicher & Wolfgang Raible (eds), *Language Typology and Language Universals: ein internationales Handbuch*, Vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter. pp. 1380-1390.

- Ladd, D. Robert. 1986. *Intonational phrasing: The case for recursive prosodic structure*. *Phonology*, 3. pp. 311-340.
- Ladd, D. R. & A. Cutler. 1983. *Models and measurements in the study of prosody*. In Cutler & Ladd (eds.) *Prosody: Models and Measurements*. Springer-Verlag.
- Ladefoged, Peter. 2003. *Phonetic Data Analysis, An introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.
- Ladefoged, Peter. 1993. *A course in phonetics*. 3rd edition. Fort Worth, TX: Harcourt, Brace, and Jovanovich.
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Lambrecht, Knud. 2001. *Dislocation*. In Martin Haspelmath & Ekkehard König & Wulf esterreicher & Wolfgang Raible (eds), *Language Typology and Language Universals. An International Handbook*. Vol. 2. Series Handbücher zur Sprach- und Kommunikationswissenschaft. Berlin, New York: Walter de Gruyter.
- Lambrecht, Knud. 1994. *Information structure and sentence form*. Topic, focus and the mental representations of discourse referents. Cambridge: Cambridge University Press.
- Laver J. 1994. *Principles of phonetics*. Cambridge University Press.
- Laughren. M. et al. 1996. *A Learner's Guide to Warlpiri*. Alice Springs: IAD Press.
- Lehiste, I. 1975. *The phonetic structure of paragraphs*. In A. Cohen & S.G. Nooteboom (eds), *Structure and Process in Speech Perception*. Springer-Verlag. pp. 195-203.
- Lehiste, I. 1973. *Rhythmic units and syntactic units in production and perception*. *J. Acoust. Soc. Amer.* Vol. 54, Issue 5. pp. 1228-1234.
- Lehiste, I. 1970. *Suprasegmentals*. MIT Press, Cambridge, Mass.
- Liu, F. & Y. Xu. 2005. *Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation*. *Phonetica* 62. pp. 70-87.
- Liu, F. & Y. Xu. 2007. *Question intonation as affected by word stress and focus in English*. In *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken. pp. 1189-1192.
- Louwerse, M.M. & J. Jeuniaux & B. Zhang & J. Wu & M.E. Hoque (to appear). *The interaction between information and intonation structure: Prosodic marking of theme and rheme*. In V. Sloutsky & B. Love & K.

- McRae (eds), Proceedings of the 29th Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum.
- Makarova, V. 2001. *Perceptual correlates of sentence type intonation in Russian and Japanese*. Journal of Phonetics, 29. pp. 137-154.
- Martin, Philippe. 2000. *L'intonation en parole spontanée*. Revue Française de Linguistique Appliquée, Vol. IV-2. Paris. pp. 57-76.
- Martin, Philippe. 1982. *Phonetic Realizations of Prosodic Contours in French*. Speech Communication 1. pp. 283-294.
- Maslova, Elena & Giuliano Bernini. 2006. *Sentence Topic in the Language of Europe and Beyond*. In Bernini, G. & Schwarz, M. (eds), Pragmatic Organization of Discourse in the Languages of Europe. Berlin, New York: Mouton de Gruyter. Vol. 20-8. pp. 67-120.
- Mayer, M. 1944 (1969). *Frog, where are you?* New York: Dial Books.
- McCarthy, John J. & Prince, Alan S. 1993. *Generalized Alignment*. In Geert Booij & Jaap van Marle (eds), Yearbook of Morphology. Dordrecht: Kluwer. pp. 79-153.
- McDonough, Joyce. 2002. *The Prosody of Interrogative and Focus Constructions in Navajo*. In A. Carnie & H. Harley (eds), Formal Approaches to Functional Phenomena. Amsterdam: Benjamins.
- McGregor, William. 2004. *The Languages of the Kimberley, Western Australia*. London, New York: Taylor & Francis.
- McGregor, William. 1990. *A functional grammar of Gooniyandi*. Amsterdam: Benjamins.
- Meakins, Felicity. 2010. *The development of asymmetrical serial verb constructions in an Australian mixed language*. Linguistic Typology 14 (1). pp.1-38.
- Miller, G. A. 1956. *The magical number seven, plus or minus two: Some limits on our capacity for processing information*. Psychological Review 63 (2). pp. 81-97.
- Miller, J. E. & Weinert, Regina. 1998. *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Oxford University Press.
- Mo, Y. 2008. *Duration and intensity as perceptual cues to naïve listeners' prominence and boundary perception*. Paper presented at 'Speech Prosody 2008'. Campinas, Brazil.
- Molnár, Valéria. 2002. *Contrast from a contrastive perspective*. In H. Hallelgard & Stig Johansson & Bergljot Behrens & Cathrine Fabricius-

Hansen (eds), Information structure in a cross-linguistic perspective. Amsterdam, New York: Rodopi. pp. 147-161.

Munro, Jennifer M. 2000. *Kriol on the move. A case of language spread and shift in Northern Australia*. In Jeff Siegel (ed.), *Processes of language contact: Studies from Australia and the South Pacific*. Quebec: Collection Champs linguistiques. pp. 245-270.

Mushin, Ilana 2005. *Word order pragmatics and narrative functions in Garrwa*. Australian Journal of Linguistics, 25 2. pp. 253-273.

Mushin, I. & B. Baker (eds). 2008. *Discourse and grammar in Australian languages*. Amsterdam: Benjamins.

Mushin, Ilana & Jane Simpson. 2005. *Word order and discourse prominence in six Australian Aboriginal languages*. Presented at 'Australian Linguistics Society Conference'. Monash University. 28-30 September.

Nagano-Madsen, Y. & A-C. Bredvad-Jensen. 1995. *Intonational phrasing in West Greenlandic Eskimo*. Department of Linguistics and Phonetics, Lund University. Working Papers, 44. pp. 129-144.

Nakagawa, N. & N. Nagaya. 2008. *Information Structure and Intonation of Right-Dislocation Sentences in Japanese*. Kyoto University Linguistic Research, 27. pp. 1-22

Neeleman, A. & E. Titov & H. van de Koot & R. Vermeulen. 2009. *A Syntactic Typology of Topic, Focus and Contrast*. In Van Craenenbroeck, J. (ed.) *Alternatives to Cartography*. Studies in Generative Grammar series. Mouton de Gruyter.

Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Ní Chasaide, A & C. Gobl. 2004. *Voice quality and f0 in prosody: towards a holistic account*. Keynote Talk. Proceedings of the '2nd International Conference on Speech Prosody'. Nara, Japan. pp.189-196.

Ní Chasaide, A. & C. Gobl. 1995. Voice source variation. In W. Hardcastle & J. Laver (eds.), *The handbook of phonetic sciences*, pp. 427-61. Cambridge, MA: Blackwells.

Nilsenova, M. 2006. *Rises and falls. Studies in the semantics and pragmatics of intonation*. Ph.D. thesis, University of Amsterdam.

Nolan, Francis. 2003. *Intonational equivalence: an experimental evaluation of pitch scales*. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona. pp. 771-774.

- Norcliffe, E. & T.F. Jaeger. 2005. *Accent-free Prosodic Phrases? Accents and Phrasing in the Post-Nuclear Domain*. Proceedings of 'Interspeech 2005.
- O'Connor, J. D. & G.F. Arnold. 1973. *Intonation of colloquial English*. Longman Group Ltd, Bristol.
- Ohala, J. & A.Dunn & R. Sprouse. 2004. *Prosody and phonology*. In B. Bel & I. Marlien (Eds.), *Speech prosody 2004*, Nara, Japan. ISCA Archive. pp 161-163.
- Oliveira, M. 2003. *Pitch reset as a cue for narrative segmentation*. In Proceedings of Prosodic Interfaces 2003.
- Oliveira, M. 2000. *Prosodic features in spontaneous narratives*. Simon Fraser University, Vancouver, Canada. Ph.D. Thesis.
- Palmer, H. 1922. *English intonation, with systematic exercises*. Cambridge: Heffer.
- Patil, Umesh & Frank Kuegler & Anja Gollrad & Gerrit Kentner & Shravan Vasishth, Shravan & Caroline Féry. 2008. *Focus, word order and intonation in Hindi*. Journal of South Asian Linguistics.
- Patz, Elisabeth. 2002. *A grammar of the Kuku Yalanji language of north Queensland*. Pacific linguistics. 527 pages.
- Pensalfini, R. J. 2002. *A Grammar of Jungulu, an Aboriginal Language of the Northern Territory*. Pacific Linguistics. 536 pages.
- Pensalfini, R. J. 1999. *Suffix Coherence and Stress in Australian Languages*. In: John Henderson, Proceedings of the 1999 Annual Conference of the Australian Linguistics Society. University of Western Australia, Perth. 28th September - 2nd October, 1999.
- Pentland, C. & M. Laughren. 2004. *Distinguishing Prosodic Word and Phonological Word in Warlpiri: Prosodic Constituency in Morphologically Complex Words*. In Proceedings of the 2004 Conference of the Australian Linguistic Society.
- Pentland C et al. 1999. *An acoustic analysis of stress in Warlpiri*. ALS conference. Perth.
- Peperkamp, S. 1999. *Prosodic words*. GLOT International 4.4. pp.15-16.
- Pierrehumbert, J.B. 2000. *Tonal elements and their alignment*. In Prosody: Theory and Experiment — Studies Presented to Gösta Bruce. M. Horne. (eds.) London: Kluwer Academic Publishers. pp. 11-36.

Pierrehumbert, J.B. 1980. *The phonology and phonetics of English intonation*. PhD dissertation, MIT. Published 1988 by Indiana University Linguistics Club.

Pierrehumbert, J. & Hirschberg, J. 1990. *The Meaning of Intonational Contours in the Interpretation of Discourse*. In P. Cohen & J. Morgan & M. Pollack (eds), *Intentions in Communication*. Cambridge, MA: MIT Press. pp. 271-312.

Pierrehumbert, J. & M. Beckman. 1988. *Japanese Tone Structure*. MIT Press.

Pike, K. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.

Poiré, P. 2006. *La perception des proéminences et le codage prosodique*. Bulletin PFC 6. pp. 69-79.

Prom-on, S. & Y. Xu & B. Thipakorn. 2009. *Modeling tone and intonation in Mandarin and English as a process of target approximation*. Journal of the Acoustical Society of America, 125. pp. 405-424.

Redi, L. & S. Shattuck-Hufnagel. 2001. *Variation in the rate of glottalization in normal speakers*. Journal of Phonetics 29. pp. 407-42.

Rialland, A. 2009. *African "lax" question prosody: its realisation and its geographical distribution*. Lingua, 119. pp. 928-949.

Ritz, Julia & Stefanie Dipper, & Michael Götze. 2008. *Annotation of Information Structure: An Evaluation Across Different Types of Texts*. In Proceedings of the 6th LREC-2008 conference. Marrakech, Morocco.

Roach, Peter. 2000. *Techniques for the Phonetic Description of Emotional Speech*. In Proceedings of the ISCA Workshop on Speech and Emotion. Newcastle, Northern Ireland. September 2000. pp 53-59.

Ross, Belinda. 2003. *The phonological / grammatical mismatch in the Dalabon word* Unpublished Honours Thesis, Department of Linguistics & Applied Linguistics, University of Melbourne.

Rossi, Mario. 2000. *Intonation: Past, Present, Future. Intonation: Analysis, Modelling and Technology*. Ed. A. Botinis. Kluwer Academic, Dordrecht; Boston; London. pp.13-52.

Rossi, Mario. 1997. *Intonation: past, present, future*. In INT-1997. pp.1-10.

Saeed, John I. 2003. *Semantics*. Blackwell Publishers; 2nd edition.

Safarova, M. & P. Muller & L. Prevot. 2005. *The Discourse Meaning of IP-final Rises in French*. In Proceedings of DIALOR. Nancy, France.

- Sasse, Hans-Jürgen. 2006. *Theticity*. In G. Bernini & M. Schwarz (eds), *Pragmatic Organization of Discourse in the Languages of Europe*. Berlin: Mouton de Gruyter. pp. 255-308.
- Sasse, Hans-Jürgen. 1987. *The thetic-categorical distinction revisited*. *Linguistics*, 25. pp. 511-580.
- Schaden, Gerhard. 2005. *Vocatives: A Note on Addressee-Management*. U. Penn Working Papers in Linguistics, Vol. 14.2.
- Schiering, René & Kristine A. Hildebrandt & Balthasar Bickel. 2007. *The prosodic word is not universal*. Ms., University of Leipzig & University of Manchester.
- Schuetze-Coburn, S. & M. Shapley & E. Weber. 1991. *Units of intonation in discourse: A comparison of acoustic and auditory analyses*. *Language and Speech* 34. pp. 207-234.
- Schultze-Berndt, Eva. (to appear). *Recent grammatical borrowing into an Australian Aboriginal language: The case of Jaminjung and Kriol*. In Y. Matras & J. Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, Berlin: Mouton de Gruyter, 363-386.
- Schultze-Berndt, Eva. (to appear). *Grammatical properties and classification of three-participant predicates in Jaminjung*. In A. Malchukov, M. Haspelmath & B. Comrie (eds.), *Ditransitives*. Berlin: Mouton de Gruyter.
- Schultze-Berndt, Eva. 2008. *What do “do” verbs do? Towards an account of the semantic diversity of generalised action verbs and its implications for grammaticalisation*. In E. Verhoeven, S. Skopeteas, Y-M Shin, Y. Nishina & J. Helmbrecht (eds.), *Studies on Grammaticalization*. Berlin: Mouton de Gruyter. pp. 185-208.
- Schultze-Berndt, Eva. 2007. *Making sense of complex verbs: on the semantics and argument structure of closed-class verbs and coverbs in Jaminjung*. In M. Bowerman & P. Brown (eds.), *Cross-linguistic perspectives on Argument Structure. Implications for learnability*. Hillsdale, NJ: Erlbaum. pp. 69-88.
- Schultze-Berndt, Eva. 2006a. *Sketch of a Jaminjung Grammar of Space*. In S.C. Levinson & D.P. Wilkins (eds.), *Grammars of Space*. Cambridge: Cambridge University Press. pp.63-113.
- Schultze-Berndt, Eva. 2006b. *Taking a closer look at function verbs: lexicon, grammar, or both?* In F.K. Ameka, A. Dench, and N. Evans (eds.), *Catching Language. The Standing Challenge of Grammar Writing*. Berlin: de Gruyter. pp. 359-391.

- Schultze-Berndt, Eva. 2006c. *Secondary Predicates in Australian Languages*. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, Volume IV. Malden: Blackwell. pp. 180-208.
- Schultze-Berndt, Eva. 2003. *Preverbs as an open word class in Northern Australian languages: synchronic and diachronic correlates*. In Geert Booij & Jaap van Marle (Hgg), *Yearbook of Morphology 2003*. Dordrecht: Kluwer. pp.145-177.
- Schultze-Berndt, Eva. 2002a. *Converging patterns of information structure in the Victoria River District (Northern Australia)*. Paper presented at the conference on 'Linguistic areas, convergence and language change'. Manchester, UK. November 23rd, 2002.
- Schultze-Berndt, Eva. 2002b. *Constructions in Language Description*. *Functions of language* 9. pp. 267-308.
- Schultze-Berndt, Eva. 2002c. *Grammaticalized restrictives on adverbials and secondary predicates: evidence from Australian languages*. *Australian Journal of Linguistics* 22,2. pp.231-264.
- Schultze-Berndt, Eva. 2001. *Ideophone-like characteristics of uninflected predicates in Jaminjung (Australia)*. In F.K.E. Voeltz & C. Kilian-Hatz (Hgg), *Ideophones*. Amsterdam: Benjamins. pp. 355-373.
- Schultze-Berndt, Eva. 2000. *Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language*. Nijmegen: PhD dissertation, University of Nijmegen.
- Schultze-Berndt, Eva. & Nikolaus P.Himmelman. 2004. *Depictive secondary predicates in crosslinguistic perspective*. *Linguistic Typology* 8,1. pp. 59-131.
- Searle, J. 1975. *A taxonomy of illocutionary acts*. In K. Gunderson (ed), *Language, Mind and Knowledge*. Minneapolis, MN: University of Minnesota Press. pp.344–369.
- Searle, John & Daniel Vanderveken. 1985. *Theory of Speech Acts of & Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.
- Selkirk, Elisabeth O. 2003. *Sentence Phonology*. *International Encyclopedia of Linguistics*. Oxford: Oxford University Press. 2nd edition.
- Selkirk, Elisabeth O. 1986. *On derived domains in sentence phonology*. *Phonology Yearbook*, 3. pp. 371-405.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax: the relation between sound and structure*. Chicago: MIT Press.

- Selkirk, Elisabeth O. 1978. *On prosodic structure and its relation to syntactic structure*. In T. Fretheim (ed.), *Nordis Prosody II*. Trondheim: TAPIR.
- Selkirk, Elisabeth O. & Koichi Tateishi. 1988. *Constraints on Minor Phrase formation in Japanese*. CLS 24. pp. 316-336.
- Selting, M. 2000. *The construction of units in conversational talk*. *Language in Society*, 29. pp. 477-517.
- Sgall, Petr. 1999. *Remarks on Sentence Prosody and Topic-Focus Articulation*. Springer-Verlag, Berlin Heidelberg. TSD 1999. pp.139-145.
- Sharpe, Margaret. 1972. *Alawa phonology and grammar: Australian Aboriginal Studies* 37. Linguistic Series 15. Canberra: AIAS.
- Shattuck-Hufnagel, Stefanie. 2000. *Phrase-level phonology in speech production planning: Evidence for the role of prosodic structure*. In Merle Horne (ed.), *Prosody: Theory and Experiment – Studies presented to Gösta Bruce*. Kluwer: Academic Publishers. pp. 201-229.
- Shattuck-Hufnagel, S. & A. Turk. 1996. *A prosody tutorial for investigators of auditory sentence processing*. *Journal of Psycholinguistic Research*, 25. pp. 193-247.
- Silverman, K. & M. Beckman & J. Pierrehumbert & M. Ostendorf & C.W.S Wightman & P. Price et al. 1992. *ToBI: A standard scheme for labeling prosody*. In *Proceedings of the 2nd International Conference on Spoken Language Processing*. Banff. pp. 867-879.
- Simpson, Jane. 2004. *Word order in Australian languages, 160 years on*. Invited paper, Lexical Functional Grammar Conference, 10 July 2004. Christchurch, New Zealand.
- Singer, Ruth. 2006. *Expression of Information Structure in Mawng: Intonation and Focus*. *Proceedings of the 'Australian Linguistics Society Conference*.
- Skopeteas, S. & I. Fiedler & S. Hellmuth & A. Schwarz & R. Stoel & G. Fanselow. & C. Féry & M. Krifka. 2006. *Questionnaire on Information Structure (QUIS)*. In *Interdisciplinary Studies on Information Structure (ISIS)*, Vol. 4. Universitätsverlag Potsdam, Germany.
- Sluijter, A. M. C. & V. J. van Heuven. 1996. *Spectral balance as an acoustic correlate of linguistic stress*. *Journal of the Acoustical Society of America*, 100. pp. 2471-2485.
- Steedman, Mark. 2003. *Information-Structural Semantics for English Intonation*. *LSA Summer Institute Workshop on Topic and Focus*. Santa Barbara, July 2001.

Steedman, Mark. 1991. *Structure and Intonation*. Language, 68. pp. 260-296.

Sudhoff, Stefan & Denisa Lenertová & Roland Meyer & Sandra Pappert & Petra Augurzky & Ina Mleinek & Nicole Richter & Johannes Schließer. 2006. Hg. *Methods in Empirical Prosody Research*. Berlin, New York: Mouton de Gruyter.

Sun, X. and Y. Xu. 2002. *Perceived pitch of synthesized voice with alternate cycles*. Journal of Voice 16. pp. 443-459.

Swerts, Marc. 1997. *Prosodic features at discourse boundaries of different strength*. Journal of Acoustic Society of America, Vol. 101, No. 1. pp. 514–521.

Swerts, Marc. 1995. *Combining statistical and prosodic analyses of spontaneous discourse segmentation*. Proceedings of the 8th International Congress of Phonetic Sciences, Stockholm, August 1995: vol. 4 – Stockholm. pp. 208-211.

Taylor, Paul. 2009. *Text-to-Speech Synthesis*. Cambridge: Cambridge University Press.

Taylor, Paul. 2000. *Analysis and Synthesis of Intonation using Tilt Model*. Journal of the Acoustical Society of America. Vol 107 3, pp. 1697-1714.

Trager, H. & H. Smith. 1951. *An outline of English structure*. Studies in Linguistics Occasional Papers 3. Washington D.C.: American Council of Learned Societies.

Tseng, C. & F. Chou. 1999. *A prosodic labeling system for Mandarin Chinese speech database*. Proceedings of the '14th International Congress of Phonetic Science'. San Francisco, USA. pp. 2379-2382.

Tsunoda, Tasaku. 2005. *Language Endangerment and Language Revitalisation*. Berlin, New York: Mouton de Gruyter.

Tsunoda, Tasaku. 1981. *The Djaru Language of Kimberley, Western Australia*. Pacific Linguistics. Series B, No 78.

Turk, A. & s. Shattuck-Hufnagel. 2000. *Word-boundary-related duration patterns in English*. Journal of Phonetics, 28. pp. 397-440.

Turk, A. & S. Nakai & M. Sugahara. 2006. *Acoustic segment durations in prosodic research: a practical guide*. In Sudhoff, Stefan & Denisa Lenertová & Roland Meyer & Sandra Pappert & Petra Augurzky & Ina Mleinek & Nicole Richter & Johannes Schliesser (eds), *Methods in Empirical Prosody Research*. Berlin, New York: De Gruyter. pp. 1-28.

- Tuttle, Siri G. & Olga Lovick. 2007. *Intonational Marking of Discourse Units in Two Dena'ina Narratives*. Cahiers de linguistique française 28.
- Umbach, C. 2001. *Contrast and Contrastive Topic*. Proceedings of the 'ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics'. Helsinki.
- Vaissière, J. 1983. *Language-Independent Prosodic Features*. In A. Cutler & R. Ladd (eds), *Prosody: Models and Measurements*. Springer Verlag. pp. 53-65
- vanValin Jr, Robert D. & Randy J. LaPolla. 1997. *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Van Dijk, T. 1976. *Discourse topics and sentence topics*. Papers in Slavic Philology, 1, 1977. pp 49-61.
- Van der Wal, J. 2008. *Agreement inthetic sentences in Bantu and Romance*. In C. De Cat & K. Demuth (eds), *The Bantu-Romance Connection. A comparative investigation of verbal agreement, DPs and information structure*. Amsterdam: Benjamins. pp. 323-350.
- Van Heuven, V.J. & J. Haan. 2000. Phonetic correlates of statement versus question intonation in Dutch. In A. Botinis (ed.), *Intonation: Analysis, Modelling and Technology*. Kluwer, Dordrecht/Boston/London. pp. 119-144.
- Von Heusinger, Klaus. 1999. *Intonation and Information Structure. The Representation of Focus in Phonology and Semantics*. Habilitationsschrift. Universität Konstanz.
- Wagner, Michael. 2005. *Prosody and recursion*. MIT. PhD Thesis.
- Walsh, Michael. 1971. *Jaminjung Dictionary*. Unpublished Ms.
- Wang, B. & Y. Xu. 2006. *Prosodic encoding of topic and focus in Mandarin*. In Proceedings of Speech Prosody 2006, Dresden, Germany: PS3-12_0172.
- Wells, R. 1945. *The pitch phonemes of English*. Language, 21. pp. 27-40.
- Werner, S. & E. Keller. 1994. *Prosodic aspects of speech*. In *Fundamentals of Speech: Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges* ed. E. Keller, 23-40. Chichester: John Wiley.
- Wichmann. Ann. 2000. *Intonation in Text and Discourse*. London: Pearson Education.
- Wightman, C. 2002. *ToBI or not ToBI*. Preceedings of the International Conference on Speech Prosody 2002, Aixen-Provence: France.

- Wightman, C. W. & S. Shattuck-Hufnagel & M. Ostendorf & P.J. Price. 1992. *Segmental durations in the vicinity of prosodic phrase boundaries*. J. Acoust. Soc. Am., Vol. 91, No. 3. pp. 1707-17.
- Wu, W. L. & Y. Xu. 2010. *Prosodic Focus in Hong Kong Cantonese without Post-focus Compression*. In Proceedings of Speech Prosody 2010, Chicago.
- Xu, Yi. 2009. *Timing and coordination in tone and intonation - An articulatory-functional perspective*. Lingua, 119 pp. 906-927.
- Xu, Y. 2005. *Speech Melody as Articulatorily Implemented Communicative Functions*. Speech Communication, 46 pp. 220-251.
- Xu, Yi. 1999. *Effects of tone and focus on the formation and alignment of F0 contours*. Journal of Phonetics, 27 pp. 55-105.
- Xu, Yi. 1997. *Contextual tonal variations in Mandarin*. Journal of Phonetics 25. pp. 61-83.
- Xu, Y. & F. Liu. 2007. *Determining the temporal interval of segments with the help of F0 contours*. Journal of Phonetics, 35. pp. 398-420.
- Xu, Y. & F. Liu. 2006. *Tonal alignment, syllable structure and coarticulation: Toward an integrated model*. Italian Journal of Linguistics, 18 pp. 125-159.
- Xu, Y. & C.X. Xu. 2005. *Phonetic realization of focus in English declarative intonation*. Journal of Phonetics, 33. pp. 159-197
- Xu, Y. & Wang, M. (submitted). *The nature of syllable organization - Evidence from F0 and duration patterns in Mandarin*.
- Xu, Y. & Wang, Q. E. 2001. *Pitch targets and their realization: Evidence from Mandarin Chinese*. Speech Communication, 33. pp. 319-337.
- Yang, Li-chiung. 2004. *Duration and Pauses as Cues to Discourse Boundaries in Speech*. Proceedings of the '2nd International Conference on Speech Prosody'. Nara, Japan. March 23-26, 2004. pp. 267-270.
- Yoon, Tae-Jin & Xiaodan Zhuang & Jennifer Cole & Mark Hasegawa-Johnson. 2008. *Voice Quality Dependent Speech Recognition*. In Shu-Chuan Tseng (ed.) Linguistic Patterns in Spontaneous Speech (Language and Linguistics Monograph Series), Academia Sinica.
- Zellner, B. 1994. *Pauses and the temporal structure of speech*. In E. Keller (ed.), Fundamentals of speech synthesis and speech recognition. Chichester: John Wiley. pp. 41-62.

Zimmermann, M. 2007. *Contrastive Focus*. In Féry, C. & Fanselow, G. & Krifka, M. (eds), Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6. Pots. pp. 147-160.

Zwicky, Arnold M. 1974. *Hey, whatsyourname!* CLS 10.787-801.