Compound Intonation Units in Totoli

Postlexical prosody and the prosody-syntax interface

Christoph Bracks

Topics in Phonological Diversity 2

language science press

Topics in Phonological Diversity

Editors: Natalia Kuznetsova (Università Cattolica del Sacro Cuore, Cormac Anderson (Max Planck Institute for Evolutionary Anthropology, Leipzig), Shelece Easterday (University of Hawai'i, Mānoa)

In this series:

1. Tallman, Adam J. R., Sandra Auderset & Hiroto Uchihara (eds.). Constituency and convergence in the Americas.

2. Bracks, Christoph. Compound Intonation Units in Totoli: Postlexical prosody and the prosody-syntax interface.

Compound Intonation Units in Totoli

Postlexical prosody and the prosody-syntax interface

Christoph Bracks

language science press

Christoph Bracks. 2024. *Compound Intonation Units in Totoli: Postlexical prosody and the prosody-syntax interface* (Topics in Phonological Diversity 2). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/414

ISBN: 978-3-96110-436-9 (Digital)
978-3-98554-094-5 (Hardcover)

DOI: 10.5281/zenodo.10559070
Source code available from www.github.com/langsci/414
Errata: paperhive.org/documents/remote?type=langsci&id=414

Cover and concept of design: Ulrike Harbort
Typesetting: Christoph Bracks, Sebastian Nordhoff
Proofreading: Amir Ghorbanpour, Carrie Dyck, Christopher Straughn, Elliott Pearl, Jeroen van de Weijer, Katja Politt, Luciana Marques, Ludger Paschen, Raquel Benítez Burraco, Tom Bossuyt
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XƎLATEX

Language Science Press
xHain
Grünberger Str. 16
10243 Berlin, Germany
http://langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Contents

Acknowledgments

I am profoundly grateful and humbled by the many privileges I enjoy, the most significant of which is the opportunity to encounter numerous individuals who have, in diverse ways, made invaluable contributions to the realization of this book. I owe an immeasurable debt of gratitude to those who have shared their love and friendship, offered unwavering support, provided indispensable counsel and guidance, and so much more. Without them, this book would not have come to fruition. I am forever indebted to them all.

First and foremost, I extend my heartfelt gratitude to the people of Tolitoli who have generously opened their homes and hearts to me, often hosting me for several months.

Pertama-tama, saya mengucapkan terima kasih yang tulus kepada masyarakat Tolitoli yang telah dengan murah hati membuka rumah dan hati mereka untuk saya, menjadi tuan rumah saya selama beberapa bulan. Saya sangat berterima kasih kepada Datra Rusmin, Akram Hasan, dan Winarno, atas persahabatan mereka dan atas waktu yang mereka luangkan untuk membantu saya mewujudkan proyek ini, baik dengan menemani saya berkendara motor, mengetik data bersama saya di lantai, atau mengunjungi anggota komunitas lain untuk merekam data dalam waktu yang lama. Saya juga ingin mengucapkan terima kasih kepada 'Ima' Irma, yang bersama saya selama perjalanan berkendara motor, memberikan wawasan yang sangat berharga tentang kesehatan reproduksi dan hal-hal terkait dengan perencanaan keluarga. Terima kasih untuk banyak tawa yang kita bagi bersama. Saya sangat berterima kasih kepada Pak 'Onal' Ronal Noldi dan Bu Augustina Pali, yang selalu memberikan rumah yang nyaman untuk saya ketika saya membutuhkan tempat istirahat, dan cangkir, koin, dan koran mereka selalu siap saat saya sakit.

Selama tinggal di Totoli, beberapa keluarga membuka pintu mereka dan menyambut saya ke rumah mereka serta menawarkan tempat tidur. Bu 'Aji' Hasnawati menerima saya di rumahnya selama beberapa bulan dan saya sangat berterima kasih atas kebaikan hati dan keramahannya serta kelezatan *Kola Durian*-nya. Di Binontoan, Pak Kades Taufiq Hidayat dan Bu Kades Samira menyambut saya dan memberikan dukungan yang sangat berharga dan bahkan mengajari saya cara mandi yang benar. Saya bersyukur atas kemurahan hati dan keramahan mereka. Di Pinjan, Pak Ahmad Laterey dan istrinya Bu Bahira dengan ramah mengizinkan saya tinggal di desa mereka yang indah, yang tetap menjadi salah satu tempat yang paling berkesan yang pernah saya kunjungi. Di Nalu, saya berkesempatan tinggal di rumah Datra Hasan, Bu Arnia Tahir, dan Ibu Sarni Hasan. Saya berterima kasih atas kesabaran dan keramahan mereka karena mengajari saya lagu anak-anak yang paling berkesan *Cicak Cicak di Dinding*. Saya juga ingin mengakui kontribusi berharga dari banyak individu yang terlibat dalam proyek dokumentasi bahasa Totoli. Tidak mungkin untuk menyebutkan semua nama mereka, dan yang dapat saya lakukan hanyalah berterima kasih kepada mereka yang telah menghabiskan banyak waktu di depan kamera dan mikrofon. Saya ingin mengucapkan terima kasih khusus kepada Pak Rusmin DJ Sola, Pak Rusmin Datuamas, dan Pak Rustam Laterey, yang dengan murah hati berbagi pengetahuan mereka tentang dongeng dan cerita lain dengan saya. Terakhir, saya ingin mengungkapkan penghargaan mendalam kepada Pak Sumitro Pogi, yang dengan sabar duduk bersama saya selama sesi elicitation yang tak terhitung jumlahnya dan menjawab semua pertanyaan saya. Tanpa bantuan yang tak ternilai dari Sumitro Pogi, penelitian ini tidak akan mungkin dilakukan.

I would like to express my sincerest gratitude to my supervisor, Nikolaus P. Himmelmann, for his invaluable contribution to my academic growth. Without his unwavering confidence, guidance, and support, I would not be the linguist I am today. In 2017, Nikolaus offered me a PhD position in the Collaborative Research Center 1252 “Prominence in Language” at the University of Cologne, where I submitted my doctoral dissertation in 2019, of which this book is a revised version. During my time there, I had the privilege of undertaking several months of data collection and language documentation with the Totoli community, an experience that has profoundly shaped me and will remain with me always. Thank you for your confidence in me and for your support!

In particular, I wish to thank Maria Bardají i Farré, who accompanied me to Tolitoli. The laughter that we shared, while being laughed at, brought so much joy to my time in Tolitoli. Thank you. Through their friendship and through the many talks over drinks Henrike Frye, Alessia Cassarà, Fahime Same and Aung Si helped me through some of the more challenging moments of my writing process. Furthermore, I am deeply grateful to my colleagues Sonja Riesberg, Volker Unterladstetter, and Kurt Malcher, as well as the numerous colleagues in and around the Collaborative Research Center “Prominence in Language” and the various institutes at the University of Cologne. In particular, I want to thank my second supervisor Stefan Baumann and also Martine Grice for their invaluable comments, suggestions and support. Lastly, I would like to express my gratitude to the two anonymous referees and the editorial team of the Topics in Phonological Diversity series for their invaluable feedback and suggestions, which greatly enhanced the quality of this book.

Abbreviations and symbols

-	morpheme boundary
=	clitic boundary
.	syllable boundary
< >	infix
*	ungrammatical form
/ /	phonological representation
[ ]	phonetic representation
< >	in text: orthographic representation
|	in examples: phonetic boundary
σ	syllable
▷	link to the audio file[1]
1	first person
2	second person
3	third person
ACT	actor
AND	andative
APPL	applicative
AV	actor voice
COLL	collective
CPL	completive
DIST	distal
EXIST	existential
GEN	genitive
GER	gerund
HON	honorific article
INCPL	incompletive
INTJ	interjection
LOC	locative
LV	locative voice
MED	medial
N	nasal
NEG	negation
NMLZ	nominalizer
NP	noun phrase
NR	numeral
NRLS	nonrealis
P	plural
PE	plural exclusive
PI	plural inclusive
PN	personal name
POT	potentive
PP	prepositional phrase
PRX	proximative
QUOT	quotative
RCP	reciprocal
RDP	reduplication
REL	relative pronoun
RLS	realis
S	singular
SF	stem former
ST	stative
UV	undergoer voice
VEN	venitive

[1] Audio files corresponding to the referenced examples can be found on the Open Science Framework (OSF) platform (Bracks 2023). By clicking on the '▷' symbol, you will be redirected through a hyperlink to access the audio files associated with each specific example.

1 Introduction

The present book is an investigation into aspects of Totoli's prosody, intonation and the prosody-syntax interface. Totoli is an endangered Austronesian language of the Malayo-Polynesian group and this book is the first study of the intonation of Totoli and among the few investigations into the prosody and intonation of Austronesian languages in general. The investigation seeks to uphold maximal ecological validity (Cicourel 2007). To this end, the analysis is based on an extensive corpus of natural (semi-)spontaneous speech which is accessible through the *Language Archive Cologne* (Bracks et al. 2023). The study takes the prime structuring unit of speech – the Intonation Unit (IU) – as its principal unit of investigation and presents a thorough description of the IU, develops an intonational model thereof and investigates the syntactic units it contains. The proposed intonational model is supported by experimental evidence of both production and perception.

The results of the various approaches taken in this book show that Totoli falls under the category of Phrase Languages (Féry 2016). From what is known so far, Totoli shows no evidence of tonal specifications at the level of the word; the language does not make use of word stress nor of lexical tone. Prosodic prominence does not play a role in the marking of information-structural categories, and tonal specifications are assigned exclusively at the level of the intonational phrase and are associated with their right-edge boundary. Based on an investigation of tonal specifications and syntactic content of prosodic units of Totoli, I show that the data is best analyzed by assuming recursive embedding of IUs into Compound Intonation Units (CIU).

When working with un(der)researched languages, one faces the task of finding appropriate tools for tapping into the prosodic system of the language. Himmelmann (2006) and Himmelmann & Ladd (2008) argue that the study of the prosody of a language should best be supported with evidence obtained through different approaches. At best, it should contain an investigation of a substantial corpus of (semi-)spontaneous speech, the analysis of which is computer-aided and complemented with experimental evidence from production and perception.

The study presented here follows this approach. It is a combination of quantitative and qualitative analyses and is based on an extensive dataset collected by the author in the course of a language documentation project (Bracks et al. 2023). With a strongly data-driven approach, the study integrates a combination of experimental evidence from both production and perception with corpus-based evidence through descriptive and inferential statistics. The data used for this research was collected within the Collaborative Research Center 1252 "Prominence in Language", funded by the DFG (German Research Foundation), and was supported by the Ministry of Research and Technology of the Republic of Indonesia through the provision of a research permit.[1] During several field trips to Totoli from 2017-2019, 196 hours of video material of various genres were collected, including a 56-hour Child Language Corpus, 85 hours of elicitation recordings, and 31 hours of (semi-)spontaneous speech. At least 20 hours of these recordings are transcribed. The subset used in this study consists of 2h 19min of recordings of (semi-)spontaneous/naturalistic speech which will be further described and discussed in §3.1.1. It is essentially an extension of the first language documentation corpus (Leto et al. 2005–2010) and follows its glossing conventions and grammatical analysis.

I make a number of analytical proposals which are relevant to prosodic theory and typology in general. This research represents a significant advancement in our understanding of the nature of prosodic systems found in (Western) Austronesian languages and intonational systems in general. Additionally, the study adheres to the principle that research should be reproducible. Thus, all data is explicitly referenced in the text and made available at an online repository (Bracks et al. 2023). Furthermore, examples from the corpus in this book are represented by periograms, which utilize automatically smoothed and interpolated pitch contours that are enriched with periodic energy. Periograms are thus phonetic representations that modulate pitch trajectories with periodic energy, by integrating "relevant acoustic cues into a perceptually motivated representation of the pitch contour of an utterance" (Albert et al. 2018: 807). I followed Albert et al.'s (2022) workflow but modified the color code. Throughout this book, pitch curves are displayed as yellow lines overlaying blue lines that represent information about periodic energy, as indicated by modifications in the transparency and width of the line. Syllable boundaries in the periograms are represented by thin, gray lines, while thicker lines are used to indicate word boundaries.

Audio for all examples is provided alongside this book and is indicated by the "▷"-sign. The first line (in italics) gives the phonemic transcription, disregarding allophonic realizations. This includes the particular case of word-final /l/ and its allophonic realization as a length-feature in word-final position in a process of

[1] Research permit number: 197/SIO/FRP/E5/Dit.KI/VII/2018

word-final compensatory lengthening (see §1.2.2). The examples include a second line with the segmentable morphemes separated by hyphens. The third line contains translations and abbreviated grammatical category labels in small capitals, and the fourth line provides a free translation to English. Information on the files of recordings in the Totoli archive (Bracks et al. 2023) is given in a final line. The examples from Totoli in this book follow the Leipzig Glossing Rules (Haspelmath 2015) and all glosses, abbreviations and other symbols used are explained above.

The primary objective of the introductory section of this book is to provide an overview and essential information necessary to comprehend the main discussions of this book, which are presented in Chapters 2 and 3. These two chapters employ two distinct approaches to studying the prosody and intonation of Totoli, with Chapter 2 concentrating on experimental methods. The results of this chapter are subsequently integrated into the analysis presented in Chapter 3, which is based on corpus-based evidence and employs Intonation Units (IUs) as the primary unit of analysis. Chapter 3 is divided into three sections: §3.1 describes fundamental properties of IUs in the corpus, §3.2 develops an intonational model of IUs based on boundary tone events and the findings of the experiments presented in Chapter 2, and §3.3 examines the syntactic content of IUs and complex Compound IUs. Finally, Chapter 4 summarizes the results, explores their implications, and suggests future research avenues.

1.1 Information on the Totoli language

Totoli is an Austronesian language spoken in the Tolitoli regency (Kabupaten Tolitoli) of the Central Sulawesi province (Sulawesi Tengah) on the Indonesian island of Sulawesi. The linguistic area is divided into a southern region, primarily comprising the city of Tolitoli and surrounding villages, and a northern area consisting of the villages of Diule, Pinjan, Binontoan, Gio, and Lakuan (Tolitoli). Figure 1.1 shows the area where Totoli is spoken. The two linguistic areas are encircled.

Himmelmann (1991: 18) calculates the ethnic population of the Totoli people to be approximately 25,000 people, but estimates that only about 30% of them – ca. 7,000 – are fluent speakers. Leto et al. (2005–2010) estimate a maximum of 5,000 fluent speakers. This number may have further declined over the last decade. In the city of Tolitoli and other villages in the southern linguistic region, it is safe to say that Totoli is never heard on the streets, and the everyday language of ethnic Totolis is Indonesian. Even in Totoli households, families speak almost exclusively in Indonesian. In contrast to the southern area, Totoli is more resilient

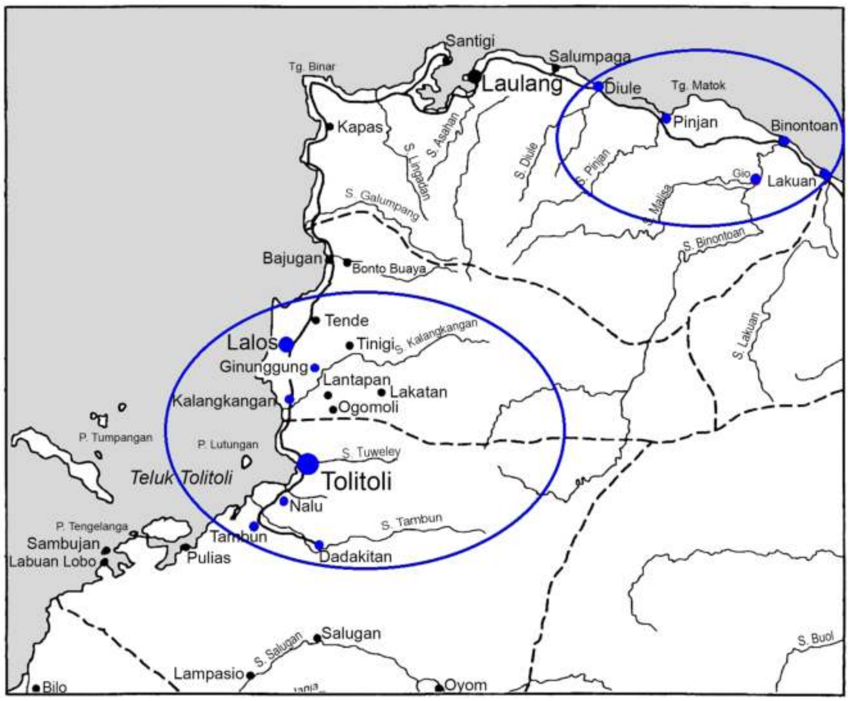

Figure 1.1: Area where Totoli is spoken, adapted from Himmelmann (1991: 29)

in the five northern villages and is, to varying degrees, still a language of everyday communication. In these villages, the great majority of ethnic Totolis are fluent speakers of Totoli, although in many areas of everyday life, Indonesian is the preferred language. Children are raised entirely in Indonesian and they are not actively taught Totoli by their parents. The only exception is the village of Gio. Interestingly, Himmelmann (1991: 28) lists the village of Gio as a Dondo-speaking settlement belonging to the village of Binontoan. Many Totoli families from Binontoan and Lakuan (Tolitoli) have moved to Gio since the 1990s, which resulted in a growth in the population of the village. As of 2019, 839 inhabitants live in 189 households in Gio (BPS Tolitoli 2019: 7). Although it was formerly a Dondo-speaking settlement, Dondo is now almost never heard there. Some of the original Dondo-speaking inhabitants still speak the language; their children, however, speak Totoli. It is now the only village where Totoli is the preferred language in almost all domains, and it is the only place where infants are reared almost exclusively in Totoli. Furthermore, it is now considered by other Totolis as the stronghold of the Totoli language, culture and music, and musicians from Gio practicing the verbal art of Lelegesan (Riesberg 2019, Bracks & Moss 2022) are frequently invited to perform in other Totoli-speaking villages, as well as in the city of Totoli. The community has become increasingly aware of the endangered state of their language. Some young speakers have successfully promoted Totoli through short films and other content on social media and streaming platforms. Furthermore, the mayor of Binontoan village has established an improvised TV channel that primarily focuses on topics related to Totoli, such as music and festivities. The channel features recordings of recent events captured on cellphones and other devices.

Totoli is primarily a spoken language and is rarely used in written communication, resulting in the absence of a standardized orthography. However, some community members occasionally write in Totoli on social media or cellphone messenger apps, using the orthography of the Indonesian language. In this book, examples from Totoli are presented in phonemic transcription.

1.2 Aspects of the segmental phonology of Totoli

As a necessary precursor to the subsequent chapters, I provide a brief description of the fundamental aspects of Totoli's segmental phonology relevant to this study. For a more detailed description, consult Bracks (submitted). The main focus here is on the phoneme inventory (§1.2.1), along with a brief commentary on phonotactics and general patterns of word structure (§1.2.3). Additionally, the topic of vowel length and related processes in Totoli is explored in greater detail (§1.2.2), as it is pertinent to the ensuing exposition of Totoli's prosody and intonation.

1.2.1 Phoneme inventory

The phoneme inventory of Totoli consists of 18 consonants and 5 vowels. Seven consonant phonemes have been introduced through loanwords, mainly from Indonesian and Arabic.

The consonant phonemes are shown in Table 1.1, with the 7 marginal phonemes indicated in brackets.

The vowel phonemes are shown in Table 1.2.

The degree of allophonic variation in phoneme realization is generally limited, except for the phoneme /l/. The following section on vowel length contains a detailed explanation of this exception, as it is of importance to this study.

Table 1.1: Consonant phonemes of Totoli

	bilabial		alveolar		palatal		velar		glottal
stop	p	b	t	d	(ʧ)	(ʤ)	k	g	
fricative			s						(h)
nasal	m		n		(ɲ)		ŋ		
lateral			l						
trill			(r)						
approximant	(w)				(j)				

Table 1.2: Vowel phonemes of Totoli

	front	central	back
close	i		u
mid	ɛ		ɔ
open		a	

1.2.2 Vowel length

Totoli phonetically distinguishes between long and short vowels. Both can occur word-initially, word-medially and word-finally. Some lexical roots inherently involve a long vowel and other long vowels are the result of affixation.

Table 1.3 gives examples of long vowels in each position as they occur in roots and through affixation.

In addition to the above, final long vowels can occur through a process of compensatory lengthening. The lateral phoneme /l/ has three allophonic realizations: [l] after front vowels, [ɹ] after back vowels, and a length feature on the preceding vowel in word-final position. Table 1.4 shows examples for the different allophonic realizations of /l/:

Evidence for analyzing final lengthening as an allophone of /l/ comes from the “reappearance” of [ɹ] or [l] when suffixes are added to such bases. Note that when clitics are added, no [ɹ] or [l] appears and the vowel remains lengthened. Three examples from the corpus are given in (1)–(2).

In example (1), the base *sumbɔl* ‘life’ occurs unsuffixed. The phoneme /l/ is realized in its word-final allophone as a length feature on the preceding vowel. In example (2), the same base is followed by the enclitic *=mɔ* ‘CPL’, so /l/ is also realized as a length feature on the preceding vowel. In example (3), the same base is followed by a suffix with initial vowel /a/ and hence the /l/ is realized as [ɹ].

Table 1.3: Bases with, or processes leading to long vowels

	EXAMPLE	POSITION OF LONG VOWEL	TRANSLATION/ GLOSS
Lexical bases with long vowel	tik**ɔɔ**	#__V:#	'neck'
	t**ii**ŋ	#_V:_#	'hear'
	ɛɛŋ	#V:__#	'saliva'
Long vowels through affixation	manaŋk**ii**	#__V:#	mɔN-saŋki-i AV-carry-APPL
	kulɔb**aa**nai	#_V:_#	ku=lɔba-an =ai 1S.ACT=inform-APPL=VEN
	iitaanna	#V:__#	i-ita-an=na RLS-see-UV=3S.GEN

Table 1.4: Allophones of lateral /l/

[l] after front vowels	lɛlɛan	[lɛlḛan]	'bridge'
	siisiligna	[si:sɪlɪgna]	'looking at him/her'
[ɹ] after back vowels	balɛ	[baɹɛ]	'house'
	nɔlumulas	[nɔɹumuɹas]	'scatter'
	tuutulu	[tu:tuɹu]	'sleeping'
compensatory lengthening in word-final position CV[l]# ⟶ CV:#	ampil	[ampi:]	'side/twin'
	mɔnɔndʒɔl	[mɔnɔndʒɔ:]	'regret'

(1) *[mɔsumbɔ: ana]*
mɔ-sumbɔl ana
AV-live MED
'(it is) alive' (lifestory_RDA_1.160) ▹

(2) *[nɔsumbɔ:mɔ]*
nɔ-sumbɔl=mɔ
AV.RLS-grow=CPL
'alive' (monkey_turtle.130) ▹

(3) *[nɔsumbɔ**ɹa**ngai]*
nɔ-sumbɔl-an=ga=ai
AV.RLS-grow-APPL=?=VEN
'they grow again' (tau_bentee.033) ▷

The regular omission of final laterals in loanwords provides further support for this analysis: Malay *kapal* 'ship' > Totoli [*kapa:*] . Throughout this work, the first line of examples gives a phonological representation of the examples. Hence, the lateral /l/ in final position is represented as /l/ but phonetically realized as a length feature.

The case of word-final long vowels is important for the subsequent discussion of tonal patterns, presented in Chapter 3. I show that right-edge boundary-tone complexes are usually associated with the final and the prefinal syllable of an IU. If the final syllable involves a long vowel, the tonal pattern is realized on the final long syllable exclusively. This is illustrated and discussed in §3.2.3, examples (13) and (14).

1.2.3 Phonotactics and word structure

Most words in Totoli follow a strict CV-pattern. Consonant clusters are rare with the exception of homorganic nasal stop clusters, a common phenomenon in languages that otherwise exhibit a rather strict CV-structure (Downing & Mtenje 2017, Downing 2005, Reid 2000, Riehl 2008, Herbert 1986). In Totoli, such sequences occur word-initially and word-medially but not in word-final position. Frequently they arise from a process commonly known as "nasal substitution" in the Austronesianist literature (Blust 2004, Pater 2004). In the examples, "nasal substitution" is represented by a capital N on the second line. Furthermore, Totoli makes use of geminates, which occur word-initially and word-medially, but not in word-final position. Some lexical roots involve geminates but frequently result from reduction processes of $C_xV_yC_xV_y$ sequences whereby the first vowel is dropped, yielding $C_xC_xV_y$. Other heterorganic consonant sequences are very rare. Only few lexical bases involve such consonant sequences. Across clitic-boundaries, however, they are allowed but are also very infrequent. Another major morphophonological process in Totoli is vowel harmony in prefixation. It is always regressive, being restricted to prefixes containing the vowel /o/ in their citation form, such as *mɔN-*, *nɔN-*, *mɔ-*, *nɔ-*, *mɔg-*, *nɔg-*, *pɔ-*, *pɔg-*, and *kɔ=*. The vowels of these prefixes occur as /ɔ/ when they precede bases containing /ɔ/, /u/, or /i/ in their first syllable. However, when the first syllable of the base contains /ɛ/, the prefix vowel is realized as /ɛ/, and when it contains /a/, the prefix vowel is realized as /a/.

Additionally, reduplication is a common morphophonological process in both verbal and nominal morphology in Totoli. This process encompasses various forms, all of which are represented by a single label, RDP, in the glossing of examples.

The aforementioned discussion provides a concise overview of the fundamental aspects of segmental phonology necessary for comprehending the discussion on Totoli's prosody and intonation in Chapters 2 and 3.

1.3 Research on the prosody of Austronesian languages

Little is known so far about the prosody of Austronesian languages, a fact also acknowledged by Himmelmann & Kaufman's (2020) chapter on the state of the research on the prosody and intonation of Austronesian languages. Himmelmann (2018: 348) proposes a model of the basic structure of the Intonation Unit (hereafter: IU) in Austronesian languages of Indonesia and East Timor. More thorough phonetic studies that have been conducted on Austronesian languages of Indonesia in recent years suggest that many languages in the area may lack word-level prominence and that tonal targets are primarily assigned at the phrase level. Indonesian/Malay as one of the major languages in the region has stirred debate about "stress" placement and its existence (for a summary see Goedemans & van Zanten 2007: 28–9). For Indonesian, as well as for many other Austronesian languages, the position of word stress is often claimed to be on the penultimate syllable of a word. Analyzing this claim on the assumption that speakers of Indonesian as a second language show a strong L1 influence, Goedemans & van Zanten (2007: 42) compared Indonesian spoken by Toba Batak speakers with that of Javanese speakers. They found that Toba Batak speakers produce the penultimate syllable of IU-final words in focus condition with higher intensity, longer duration and a rise in F0. Speakers of Indonesian with Javanese as their first language, however, produce the words in the same condition only with a rise in F0, whilst duration and intensity are not affected. They conclude that Indonesian spoken by Toba Batak speakers exhibits prominence on the level of the word as well as the phrase. For speakers with a Javanese background, however, they "only found evidence for prominence at the phrase level (in the form of pitch movements)" (Goedemans & van Zanten 2007: 45). The results found for the Indonesian of Javanese L1 speakers are in fact similar to what has been reported about the Indonesian/Malay variety Ambonese Malay, spoken in Eastern Indonesia on the Maluku Islands. Analyzing IU-final F0 movements in Ambon Malay, Maskikit-Essed & Gussenhoven (2016: 382) found no association of the timing of IU-final boundary-tone complexes with any syllable. Moreover, focus condition did not reveal any systematic effect on the range and shape of the pitch on IU-final words. Hence the authors opt for analyzing IU-final tone complexes as floating boundary tones, since such an analysis assumes neither word stress nor pitch accent, whether associated lexically or postlexically (Maskikit-Essed & Gussenhoven 2016: 356). They conclude that IU-final boundary-tone complexes may instead signal the function of sentences. The absence of word prosody and the assignment of tone complexes to boundaries of prosodic domains fit the characteristics of what Féry (2016) labels Phrase Languages.

In addition to the studies of phonetic correlates of stress in Austronesian languages, only a small number of analyses of the intonation of Austronesian languages exist: Himmelmann (2010) presents a description of the intonation of Waima'a, spoken in East Timor; Maskikit-Essed & Gussenhoven (2016) describe the two most common IU-final pitch melodies of Ambonese Malay; Stoel (2006) proposes a concise description of the intonation of Banyumas Javanese. These studies are based on a set of target phrases or question-answer pairs, the realization of which has been taken as generalizable over the intonational system of the language as a whole. Such an approach may be suitable for the description of the major aspects of the intonation of a language. However, the frequency distribution of patterns and also less frequently used intonation patterns may only be observed in a corpus study, covering different communicative events. Possibly the only study conducted on the intonation of an Indonesian language which is primarily based on a corpus of spoken spontaneous discourse is that of Stoel (2005) on Manado Malay.

1.4 The units of spoken speech

This research is based on the analysis of a large corpus of speech. The first hurdle one faces when dealing with corpora of (semi-)spontaneous speech is the identification and segmentation of the data into tangible units (Himmelmann 2006, Edwards & Lampter 1993). Speech can be segmented into various units of different sizes, though most studies recognize the Intonation Unit (IU) (Chafe 1994) as the basic unit into which discourse and the flow of speech is structured.

The IU has been discussed under a variety of other names such as the Tone Group (Halliday 1967), the Tone Unit (Crystal 1976), the Intonation/Intonational Phrase (Selkirk 1986, Nespor & Vogel 1983, Pierrehumbert 1980, Ladd 2008, Gussenhoven & Chen 2020, Jun 2005c, 2014a), and the Breath Group (Lieberman 1966, Lieberman et al. 1970). Details of the definition of the various terms vary. Underlying this study is the basic definition of the IU by Chafe (1987: 22):

> An intonation unit is a sequence of words combined under a single, coherent intonation contour, usually preceded by a pause.

The coherent intonation contour is the defining characteristic of an Intonation Unit. A number of features have been identified which contribute to the perceived single, coherent intonation contour. Other features, on the other hand, delimit a speech segment and indicate the boundary of an IU. The criteria discussed pertain mainly to pitch, rhythm, and voice quality, but non-prosodic features have also been identified, such as the end of a turn/the change of speaker, inhalation, and lexical boundary markers (Schuetze-Coburn 1994: 93–155, Himmelmann 2006: 260–270, Du Bois et al. 1992, 1993, and Cruttenden 1997: 29–39).

Tao (1996: 52) mentions that discourse particles also proved a reliable criterion for the identification of IU boundaries in Mandarin Chinese, as they correlate highly with IU boundaries. Strictly speaking, however, prosodic clitics are a syntactic criterion and, as such, should not be used to identify any prosodic unit (see the discussion in §3.3). A single IU-boundary feature alone does not suffice to reliably detect an IU boundary, and hardly any IU exhibits all of the boundary cues:

> The relative importance of the cues may differ — pitch reset, for example, is arguably more central than tempo modulation — but none alone defines an IU boundary per se; rather, a conjunction of cues is usually required for an IU to be perceived. One can say that the prototypical IU exhibits all of these cues, yet seldom are all actually present in any given instance. That is, most IUs deviate from the prototype to some degree. Thus, a given IU may exhibit pitch reset and a definite contour, but none of the other features. (Schuetze-Coburn et al. 1991: 217)

Hence, the IU is defined in terms of a prototype and "the more features that coalesce at any point, the stronger ('more prototypical') the boundary will be, but an IU boundary may also be perceived when only one or two features occur" (Schuetze-Coburn et al. 1991: 227).

Many discourse-oriented linguists report on difficulties with identifying IU boundaries and comment on the sometimes tedious nature of the task. While Brown et al. (2015: 46) report "constant difficulty in identifying tone groups in spontaneous speech", a great many other linguists working with discourse data admit that, no matter how difficult the task, with "practice and appropriate guidance, however, one should be able to attain a reasonably high degree of intertranscriber reliability" (Du Bois et al. 1992: 112). On that matter, Chafe (1994) comments that "in a better world they would be as important a part of the training of a linguist as the ability to transcribe vowels and consonants" (see also comments from Schuetze-Coburn 1994: 165, Crystal 1976: 206, and Cruttenden 1997: 29). Based on his hands-on experience of working with various corpora, Himmelmann (2006: 261) reports that an estimated proportion of 80-90% of IUs are rather unproblematically identifiable, which also reflects my own experience.

Despite the many remarks about the difficult nature of the transcription process, studies have shown that even naive, untrained speakers perform remarkably well in segmenting discourse. Kreckel (1981) conducted an experiment in which untrained, native English-speaking, participants were presented with a written transcript and a corresponding (English) audio recording. Participants were asked to mark 'message' boundaries on the transcript. The results showed that the participants segmented speech into Intonation Units (i.e. 'tone groups') with a high degree of interrater agreement. Furthermore, participants gave priority to prosodic cues over syntactic ones.

In recent years, these findings have been confirmed by a number of studies using the Rapid Prosody Transcription method (RPT; Cole & Shattuck-Hufnagel 2016). Its boundary-marking task is similar to the method used by Kreckel (1981), and the results obtained from a number of typologically unrelated languages show a high degree of interrater agreement on the placement of boundaries. In §2.1.5, I give an overview on agreement results from different studies and compare these to the results obtained from an RPT study of Totoli (cf. Figure 2.5). However, results from the boundary-marking task obtained from RPT experiments are usually not discussed as evidence for the perception of IU boundaries. Yet, in §3.1.3, I correlate the results from the boundary-rating task of the RPT with IU boundaries which occur within rated speech segments. The results show that naive listeners can indeed reliably identify IU boundaries. While the universality of prosodic units below the IU are subject to debate (Bickel & Zúñiga 2017, Schiering et al. 2010), the IU is widely accepted.

The Intonation Unit as a discourse-structuring unit has been successfully employed by studies on a variety of typologically unrelated languages. If prosodic cues that delimit IUs are similar across languages, then listeners should be able to identify IU boundaries even in languages they are not familiar with. In this regard, Ford & Thompson (1996: 174) briefly commented that trained transcribers can reliably identify IU boundaries in an unknown language with a precision of 85-90%. This observation has been put to the test only by Himmelmann et al.

(2018), which investigated the inter-transcriber agreement of IU segmentation done by trained transcribers on familiar and unfamiliar languages from different language families (German, Papuan Malay, Wooi, Yali). The results showed statistically significant inter-transcriber agreement on the placement of IU boundaries, which led the authors to postulate the Universal Phonetic IP Hypothesis (UPIPH). The hypothesis claims

> […] that all natural languages make use of the same kinds of phonetic cues for IPs, and that these cues can be perceived by speaker-hearers even in unfamiliar languages. […] We believe that it is quite likely that phonological IPs are part of the prosodic system of all natural languages. If this is the case, IPs would be a prime example of a universally attested *phonological* category. (Himmelmann et al. 2018: 239–240)

The UPIPH is a strong claim and further data from different languages is needed to substantiate it. Furthermore, an investigation into the comparison of the various cues for IP boundaries may yield interesting cross-linguistic similarities and/or differences. However, with supporting evidence from a variety of languages, it appears that all speakers organize their speech into Intonation Units, which are perceived as such by the listener.

As will become evident from the analysis of tonal patterns of segmented IUs of the corpus in §3.2, we have to assume recursive embedding of Intonation Units into Compound Intonation Units in Totoli. While some segmentable stretches of speech of the corpus occur as simple, singleton IUs, others occur as complex, Compound IUs, all of which are subsumed under the label CIU.

2 Experimental approaches to the prosody of Totoli

Himmelmann & Ladd (2008: 250) summarize that sentence-level prosody is typically employed in marking sentence modality, phrasing, and prominence. In this chapter, I present experimental evidence that focuses particularly, though not exclusively, on the latter, i.e. the role of prominence in the prosody of Totoli. The chapter consists of two sections. In the first section, I present the results of a Rapid Prosody Transcription (RPT) experiment (§2.1). This setup has proven particularly useful for obtaining a first impression of the prosody of a little-known language. The study's results are complemented by two focus marking experiments, which constitute the second section of this chapter (§2.2). This section explores whether prosodic prominence is used to mark the information-structural category focus. To ensure adherence to the fundamental principle of complete reproducibility, R scripts and raw data can be readily downloaded from Bracks (2023).

2.1 Investigating the role of prosodic prominence through a Rapid Prosody Transcription experiment

As a first step towards understanding the prosody of Totoli, I conducted a Rapid Prosody Transcription (RPT) experiment to gain preliminary insights into the role of prosodic prominence in the language's intonation. The RPT method is a simple and relatively quick tool that captures listeners' perception of boundaries and prominences (Mo et al. 2008, Cole et al. 2014). A description of the method is given by Cole & Shattuck-Hufnagel (2016: 11).

> [It] draws on linguistic theories of prosody (or intonation) in recognizing prominence and phrasing as two separate dimensions of prosodic form, and as such RPT can be used within any theoretical framework that recognizes prominence and phrasing, as a means of tapping into ordinary listeners' subjective impression of prominences and boundaries in speech.

In an RPT experiment, speakers are presented with speech samples and a transcription thereof and are asked to identify perceived prominences and boundaries, based on their auditory impression of the recording. The task does not require any experience in prosodic transcription or linguistic knowledge. The RPT method has been employed in a number of studies on well-researched languages such as English (Mo et al. 2008, Cole et al. 2010), German (Riesberg et al. 2020), Estonian (Ots & Asu 2019), and Korean (You 2012).

Crucial in the choice of RPT as experimental approach towards the prosody of Totoli taken here is Cole & Shattuck-Hufnagel's note that it is particularly useful for "populations not easily accessed from the university communities where most prosody researchers reside [and it] opens the door to obtaining prosody judgments from minority linguistic communities, from elderly people and those in rural communities, and from communities of language learners" (2016: 12). Riesberg et al. (2018, 2020) followed up on this suggestion and successfully employed the RPT method in a study on Papuan Malay.

In this light, RPT provides a suitable setup for this investigation here.

2.1.1 Materials

For the experiment, speech samples were taken from recordings of Pear Story (Chafe 1980) retellings from the Totoli language corpus (see §3.1). Five different speakers were selected based on the quality of their retelling in terms of smooth speaking flow and naturalness. A total of 71 speech samples were taken from the recordings, each ranging between 1.37 and 6.73 seconds in length and comprising between one and three CIUs.

Table 2.1 gives an overview of the number of CIUs the speech samples contain.

Table 2.1: Number of CIUs contained in the samples used in the RPT experiment

Number of speech samples	Number of CIUs in speech sample
41	1
26	2
4	3
total = 71	

Table 2.2 gives an overview of the number of words and the duration of samples used in the RPT experiment.

Table 2.2: Duration (in seconds) and number of words in speech samples

	min	max	mean	median	sd	total
number of words	4	13	6.83	6	2.32	484
duration in seconds	1.37	6.73	2.78	2.68	1.03	197.26

Further information about the speakers of the stimuli is given in §A.1.1 of the Appendix.

The speech samples were presented without any punctuation and used the local orthography. An example is discussed below (see example (1) and its realization in Figure 2.2).

2.1.2 Participants

Twenty native Totoli speakers were recruited for the experiment: 12 male and 8 female — M_{Age} = 30.05; $Range_{Age}$ = 18–45. Participants were required to be fluent in Totoli and possess good computer skills. All participants reported being born and raised in the Tolitoli regency (Kabupaten Tolitoli) and raised with Totoli as their first language. Additionally, they are also fluent speakers of the spoken variety of Indonesian/Malay in the region, and to varying degrees, Standard Indonesian. Further information regarding the participants can be found in §A.1.2 of the Appendix.

Totoli is an endangered language and, as such, the recruitment of participants is challenging. Consequently, all 20 participants in the experiment were asked to perform both boundary and prominence judgments. To control for potential task order effects, I followed the approach of Mo et al. (2008) and divided the participants into two groups. The first group rated prominences before boundaries, while the second group completed the tasks in the reverse order.

2.1.3 Procedure

In an RPT experiment, participants are presented with speech samples along with a transcription of the recording and are asked to identify perceived prominences and boundaries based on their auditory impressions. It is noteworthy that the task does not require any experience in prosodic transcription or linguistic knowledge.

The stimuli were presented via the LMEDS web interface (Cole et al. 2017, Mahrt 2016: 206). Since Indonesian is the national language and the medium of formal education, the instructions were given in Indonesian, as participants would find it highly unnatural to receive instructions in Totoli. To maintain consistency and comparability, the instructions and examples were taken from Riesberg et al. (2018: 409–411) and reprinted in §A.1.3 and §A.1.4 of the Appendix.

Boundaries were briefly explained to participants as a tool employed by speakers to chunk some words together or separate others (see §A.1.3 of the Appendix). An example of grouped numbers in a long telephone number was given:

229 100 2999

A second example was given which was equivalent to:[1]

"I eat, Father." vs. *"I eat father."*

The concept of *prominence*, on the other hand, has no exact equivalent in Indonesian (compare also Cole & Shattuck-Hufnagel 2016: 29). Riesberg et al. (2018: 409) describe prominences as a way in which speakers make some words stand out (Indonesian: *menonjol* 'to stand out') and state that this can usually be heard or felt by the listener. The exact wording and a translation to English is reprinted in §A.1.4 of the Appendix.

Two Indonesian examples were presented to the participants; their English translations are reprinted here:

1) She sees **a cow**

2) She sees a **cow** and a horse **eating** grass

In the LMEDS web interface, speakers click on a word: in the boundary-marking task, a vertical bar (|) appears after a selected word to indicate a boundary; in the prominence-marking task, the selected word appears in **bold**. Participants listened to the audio exactly twice. Selection of words (i.e. placement of boundaries or prominences) was permitted only after participants had listened to the speech sample at least once. No time constraint was given for marking either the boundaries or prominences for a respective speech sample. Participants were told explicitly that they were free to mark as many or as few boundaries or prominent words as they wanted. They were also told that they could change their minds,

[1]The Indonesian example given was: *Bapak saya sudah datang*. 'My father is already at home.' vs. *Bapak, saya sudah datang*. 'Father, I am already home.'; cf. Riesberg et al. (2018: 409–412).

selecting and deselecting words freely before continuing to the next speech sample.

The prominence- and boundary-rating tasks are illustrated in example (1), taken from the speech samples used in the experiment. The glossing and translation is included here for the reader only and boundary and prominence marking is arbitrarily chosen for the illustration of the task. In the experiments, participants were presented with the transcript — i.e. the first line — only.

(1) *isakɛmɔ* ***ulaŋ*** *dɛi sapɛda maŋana ia* ***nallakɔmɔ***
isakɛmɔ ulaŋ | *dɛi sapɛda* | *maŋana ia nallakɔmɔ*

ni-sakɛ-0=mɔ	ulaŋ	dɛi	sapɛda	maŋana	ia	nɔ-RDP-lakɔ=mɔ
RLS-put.up-UV=CPL	again	LOC	bike	child	PRX	AV.RLS-RDP-walk=CPL

'(after he) puts it on his bicycle again, the child walks off'

(pearstory_9_FAH.039-40) ▹

Before the Totoli data was presented to the participants, they completed a training run with four Indonesian speech samples taken from Pear Story retellings. Participants had no prior experience in participating in an experiment and the trial runs in a language they are most familiar with as written medium was deemed necessary so that they could get accustomed to the task. Riesberg et al. (2020) showed that participants are very sensitive to language specific cues in the marking of prominences, even in languages they are not familiar with. Based on these results, I do not expect any influence of the trial runs in Indonesian, although a potential influence on the overall result cannot be excluded.

2.1.4 Analysis

Participants rated 71 speech samples. Boundaries placed after the last word of a given speech sample were discarded, as no judgment was needed there. Following Cole et al. (2010: 304), I calculated boundary scores (b-scores) and prominence scores (p-scores) for each word, representing the proportion of speakers who marked the respective word as prominent or as preceding a boundary.

In Figure 2.1, the results of both tasks are illustrated for the speech sample presented above in example (1). The speech sample consists of two CIUs.

Figure 2.1 shows that most speakers perceived a boundary following *sapɛda* 'bicycle' (b-score = 90), which coincides with the location of a CIU boundary determined by my analysis (see §1.4). Similarly noticeable are the relatively high

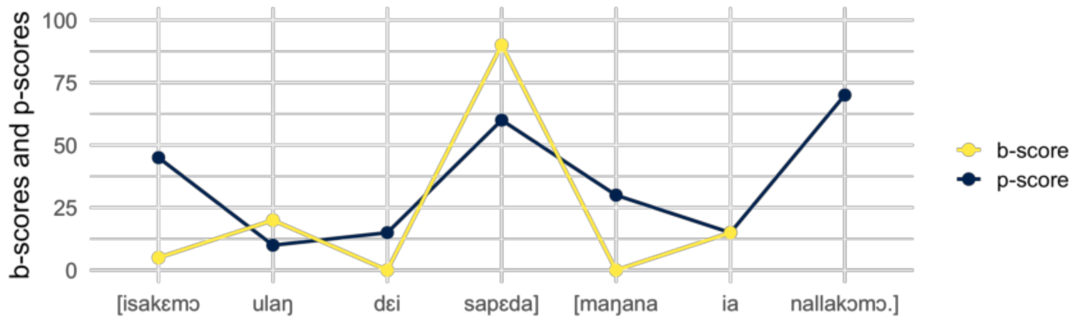

Figure 2.1: p-scores and b-scores for example (1), squared brackets indicate CIU boundaries.

p-scores for the first word of the first CIU, *isakɛmɔ* 'put up' (p-score = 45), and the last word of the second CIU, *nallakɔmɔ* 'to go' (p-score = 70).

Figure 2.2 shows the periogram with pitch track in semitones (st) of example (1). The pitch contour is given in yellow. The blue line in the background gives information on periodic energy, represented by the width and transparency of the line. The prominence ratings are indicated in red and the boundary ratings are indicated in black.

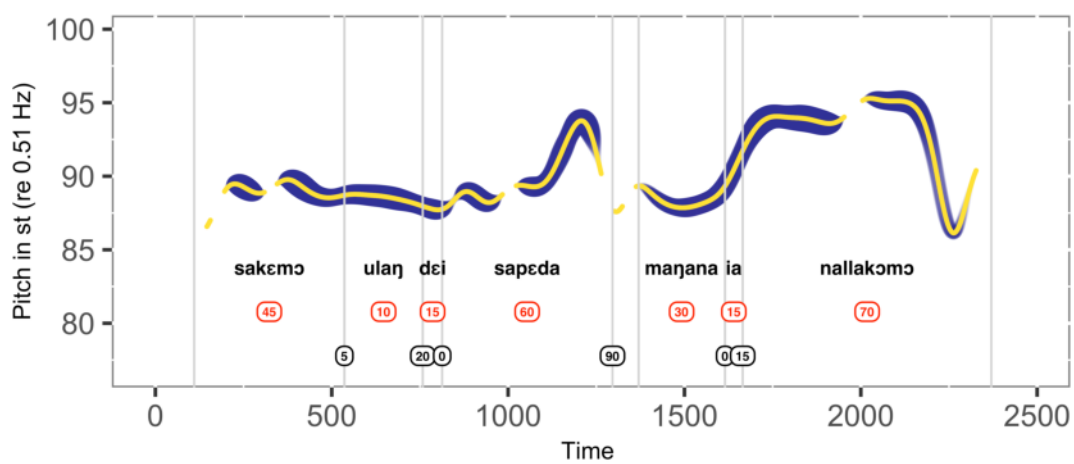

Figure 2.2: Periogram with pitch track (in st) for example (1): order of tiers from top to bottom is word, p-score, b-score; speaker FAH.

Inspection of the pitch contour given in Figure 2.2 and its comparison with the respective boundary and prominence ratings in Figure 2.1 shows that boundary ratings, prominence ratings and pitch rises appear to largely coincide.

To analyze the degree of agreement between participants, I calculated Fleiss' kappa and Cohen's kappa (κ) coefficients (Fleiss 1971, Cohen 1960). Cohen's κ measures the agreement of judgments between two participants over all rated items, thus providing $(n^2 - n)/2$ values for n participants. Fleiss' κ is a measure that provides a single figure indicating the overall agreement among all participants. Kappa values range from 0 to (−)1. A value of 0 indicates agreement at chance level and a positive value indicates agreement above chance level (Cramer & Howitt 2004: 83).

First, I discuss the Fleiss' kappa and Cohen's kappa for Totoli before relating the results more meaningfully to those reported for other studies.

2.1.4.1 Fleiss' kappa coefficients

Table 2.3 shows the Fleiss' κ coefficients for the boundary-rating task and the prominence-rating task. It provides values for all raters together, as well as separately for the group that rated boundaries first and then prominences and vice versa.

Table 2.3: Fleiss' kappa: the difference between rated subjects equals the 71 discarded words in stimulus-final position for boundary ratings.

prominences			all	boundaries first	prominences first
	Stimuli	=	484	484	484
	Raters	=	20	10	10
	Kappa	=	**0.143**	0.138	0.165
boundaries			all	boundaries first	prominences first
	Stimuli	=	413	413	413
	Raters	=	20	10	10
	Kappa	=	**0.485**	0.497	0.506

Here, I interpret the kappa values only in relation to each other. Comparing the values between the two tasks gives information about the extent to which the raters agree on the placement of prominences in comparison to the placement of boundaries. The kappa values in Table 2.3 are substantially lower for judgments of prominence placement ($\kappa = 0.143$) than for judgments of boundary placement ($\kappa = 0.485$).

The comparison of kappa values between the two groups, i.e. those who rated boundaries first and then prominences and vice versa, provides information about the influence of the task order. The comparison shows that, similar to the findings of Mo et al. (2008: 736), the influence of order of tasks does not have a strong effect (prominence placement: κ = 0.138 and 0.165; boundary placement: 0.497 and 0.506).

In sum, the participants agreed substantially more on the placement of boundaries than they did on the placement of prominences. The task order, however, had only a marginal effect on the kappa values.

To further analyze the degree of agreement between individual pairs of speakers, Cohen's kappa coefficients were calculated.

2.1.4.2 Cohen's kappa coefficients

The distribution of Cohen's kappa coefficients for pairwise interrater agreement between speakers are shown below in the violin plot in Figure 2.3.

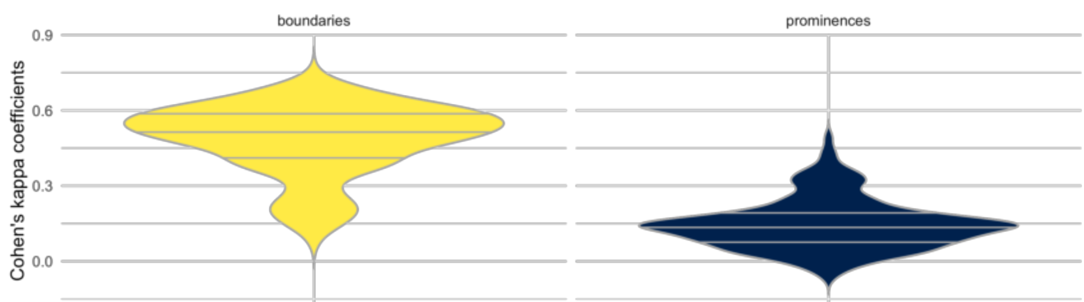

Figure 2.3: Distribution of Cohen's kappa coefficients of 190 rater pairs: mean boundaries = 0.49, median boundaries = 0.52; mean prominences = 0.15, median prominences = 0.14

Figure 2.3 shows that the pairwise agreement on boundaries is substantially higher than agreement for prominences. Landis & Koch (1977: 165) propose agreement bins for the classification of pairwise interrater agreement values. Figure 2.4 shows the frequency distribution across agreement bins: <0.00 = *poor*; 0.00–0.20 = *slight*; 0.21–0.40 = *fair*; 0.41–0.60 = *moderate*; 0.61–0.8 = *substantial*; 0.81–1.00 = *almost perfect*.

In the case of boundaries, no pairs are found in the *poor*, 6.84% in the *slight* and 16.32% in the *fair* categories, with the majority being in the category *moderate* (57.37%). Some are even in the category *substantial* (19.47%). As for prominences, 79.48% of all pairs are found in the categories *poor* and *slight*. *Moderate* is only attested for three pairs (1.58%) and non fall into the category *substantial* or *almost perfect*.

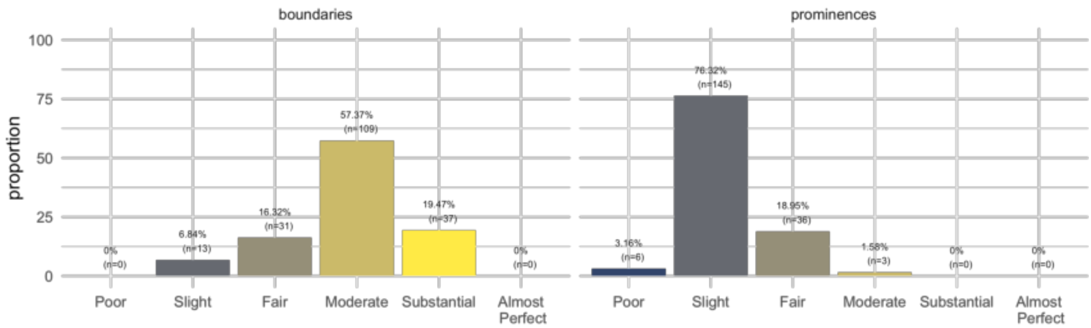

Figure 2.4: Frequency distribution of Cohen's kappa coefficients in agreement bins according to Landis & Koch (1977: 165): total numbers are indicated in brackets.

2.1.5 Discussion

The Fleiss' κ and Cohen's κ coefficients above show that participants generally agree only very little on the judgment of which word in a given speech sample is prominent. They agree considerably better on the placement of boundaries. This is especially evident when considering the pairwise interrater agreements of the Cohen's κ.

As measurements of interrater agreement, Cohen's κ and Fleiss' κ have been used in a number of RPT studies on different languages, providing a growing body of literature for comparison. Figures 2.5 and 2.6 compare the results obtained for Totoli with those reported by other studies. There are more studies that report on values for agreements on prominence ratings than on boundary ratings. Note, however, that the different studies vary with regard to the speech samples used, which may limit their comparability.

In the studies of American English (Mo et al. 2008, Cole et al. 2010), excerpts from the Buckeye Corpus of spontaneous conversations (Pitt et al. 2007) were used for the experiment. Similarly, the study of Papuan Malay (Riesberg et al. 2018, 2020) used samples of spontaneous speech, obtained from recordings of the Pear Film (Chafe 1980) and the Tangram Task. The study on Estonian (Ots & Asu 2019) used excerpts of the Phonetic Corpus of Estonian Spontaneous Speech. You (2012) used answers obtained from two native speakers replying to a set of questions in Korean. Lastly, data on German are reported by Baumann & Winter (2018), who used sentences read out loud for the RPT experiment. The study of Totoli is most comparable to that on Papuan Malay by Riesberg et al. (2020) regarding the experimental setup and material used.

Figure 2.5 shows the Fleiss' κ coefficients of the boundary-rating task reported in the studies mentioned above, comparing these with the results obtained from Totoli.

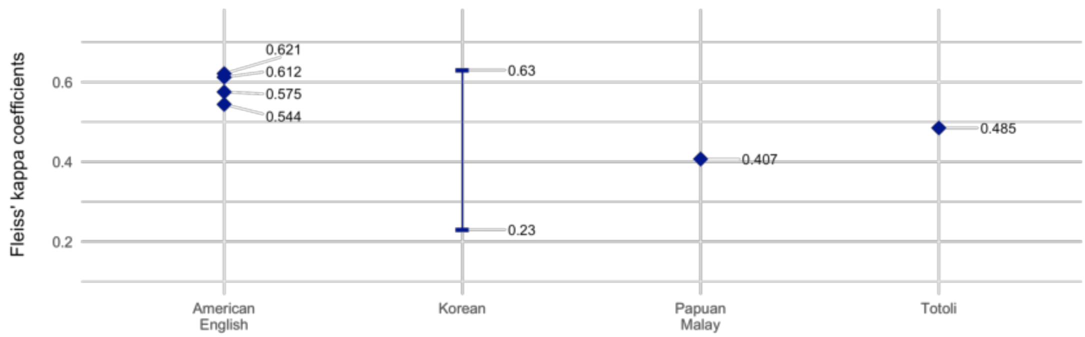

Figure 2.5: Fleiss' κ for boundary ratings using the RPT method: dots indicate single reported values, lines indicate reported range.

Table 2.6 shows the Fleiss' κ coefficients reported in various studies for the prominence-rating task.

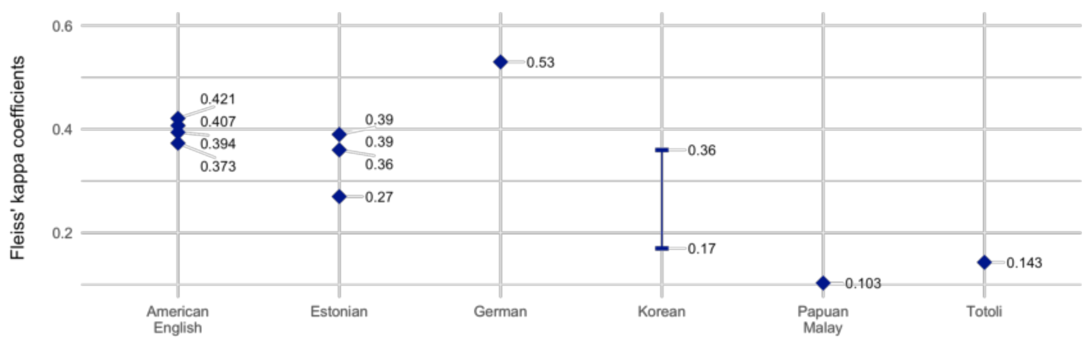

Figure 2.6: Fleiss' κ for prominence ratings using the RPT method: dots indicate single reported values, lines indicate reported range.

Across the languages, Fleiss' κ coefficients vary substantially for prominence ratings. The highest results for agreement were found for German with a Fleiss' κ coefficient of 0.53 (Baumann & Winter 2018). The lowest results were reported for Papuan Malay (Riesberg et al. 2018, 2020) with a Fleiss' κ coefficient of only 0.103. The results obtained for Totoli ($\kappa = 0.143$) are similar to that of Papuan Malay, although slightly higher.

In sum, the comparison of prominence and boundary ratings across the various studies on various languages shows a higher degree of agreement for boundary rating than for prominence rating. With regard to boundary rating, the results for Totoli are similar to those reported by other studies (see Figure 2.5). With regard to prominences, however, Totoli listeners show comparatively low agreement values that are most similar to those reported for Papuan Malay (see Table 2.6). The results point to the fact that prosodic prominence may not be a relevant category in the prosodic system of Totoli.

This is in fact similar to what Riesberg et al. (2018) report for Papuan Malay, where participants show comparatively low agreement on prominence placement and appear to perceive prominences mainly at boundaries. Riesberg et al. (2020: 2) caution that the results obtained from an RPT study, such as the above, "cannot establish 'facts' of the type 'language X makes use of pitch accents' or 'speakers of language X hear durational differences as marking lexical stress.'" Therefore, I conducted two further experiments, reported in §2.2 that examine the assumption about a possible lack of prosodic prominence marking in Totoli.

2.2 Investigating the role of prosodic prominence through a focus marking experiment

In order to further investigate the role of prosodic prominence in Totoli, I conducted a production and perception experiment which examines whether focus is acoustically prominent in the marking of information-structural categories, such as focus.

Studies on West Germanic languages show a fine-grained distinction between various focus structures, with the most pronounced difference in production between background and narrow focus (Mücke & Grice 2014, Baumann et al. 2006, Swerts et al. 2002, Kaland et al. 2023, Lee et al. 2015, Kember et al. 2021). Thus, if prosodic prominence were to play a role in the marking of focus in Totoli, I expect it to be observable in constructions such as those in Figure 2.7.

Most European languages make use of postlexical pitch accents as a means to express information-structural categories (see among others Gussenhoven 1984, Ladd 2008, Jun 2014a, Bolinger 1986, Halliday 1967, Grice et al. 2020, Wagner 2012). Thus far, little is known about the role of prosodic prominence in the marking of information-structural categories in Austronesian languages. Summarizing what is known so far about Austronesian languages of Indonesia, Himmelmann (2018: 347) comments that "it seems likely that prosodic prominence does not have a major role to play in marking information-structural categories".

Inasmuch as is known from typologically diverse languages, phrasing may be an alternative strategy to mark focus:

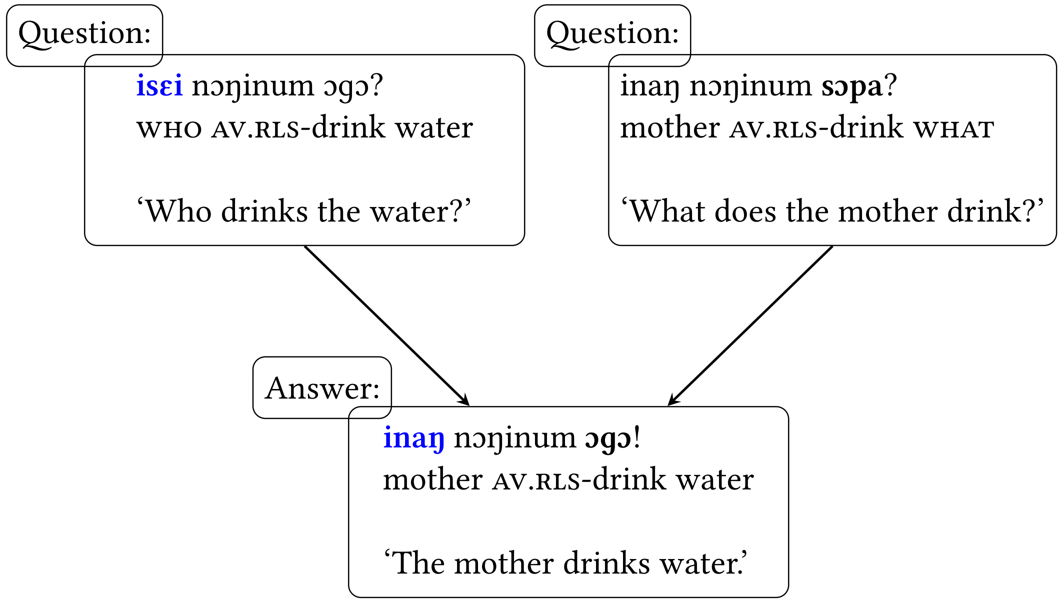

Figure 2.7: Example of question-answer pairs with a syntactically identical answer

> [..] the function of postlexical pitch accent in English and other Germanic languages (such as marking focus or disambiguating an ambiguous string) is performed by placing words in the same or different prosodic units, i.e. prosodic phrasing in Japanese and Korean. (Jun 2005b: 414)

As exemplified in Figure 2.7 above, Totoli allows a syntactically identical clause as an answer to different wh-questions that trigger narrow focus on either the subject or the object. This provides a suitable testing ground for the investigation of focus marking in Totoli and the role of prosodic prominence thereto.

Here, I present an experiment that examines the role of prosodic prominence in Totoli by searching for prosodic cues in the marking of information-structural categories. That is, I investigate whether the information structure category of focus is acoustically prominent in identical constructions that were uttered as answers to questions triggering either subject or object focus.

2.2.1 Materials

I recorded a set of question-answer (QA) pairs of different focus types (explicated below), taken from Skopeteas et al. (2006: 206–220). I selected those of narrow focus type for further analysis, including the examples in Figure 2.7. Different types of narrow focus have been identified; the QA pairs under discussion here correspond to what has been called "information focus" (Krifka 2008).

I recorded the QA pairs with 6 different speakers (2 females and 4 males): M_{Age} = 31.15; $Range_{Age}$ = 26–61. Further information on the recorded speakers is given in §A.2.1 of the Appendix.

The QA pairs were presented to the speakers one by one in a PowerPoint presentation. The recruited speakers had already been recorded beforehand during the Totoli documentation project and, therefore, were comfortable with the recording setting. All recordings were done in cooperation with Datra Hassan (DT; see Appendix §A.2.1), native speaker of Totoli and member of the Totoli documentation team. He uttered the questions and the other speakers spoke the answers. The set of QA pairs was recorded twice with each speaker. In the first round, the speakers could familiarize themselves with the task and the different QA pairs. The recordings of this round were not used for the analysis here. Before the second round, the speakers were instructed to listen attentively to the question to which they were answering. This was done to ensure that, when uttering the answers, speakers were fully aware of the foci triggered by the questions. Speakers were allowed to immediately repeat a QA pair if they judged their production to be unnatural or erroneous.

Both Datra and the speakers were wearing head-mounted AKG C520 condenser microphones, attached to a Zoom Q8 audio/video recorder. Recordings were done at a sampling rate of 44.1 kHz in a 16-bit mono format.

The nine selected QA pairs involved four different answer clauses. Three of the answer clauses were transitive constructions of the structure *subject-verb-object* (SVO) and one of the answer clauses was a ditransitive construction of the structure *subject-verb-object-indirect object* ($SVOO_i$). Each of the three transitive clauses occurred twice: once as the answer to a wh-question triggering narrow focus on the subject and once as the answer with narrow focus on the object. The ditransitive clause occurred three times: as the answer to wh-questions triggering narrow focus on the subject, on the object, and on the indirect object.

For a full account of the set of QA pairs, see §A.2.3 of the Appendix.

In the first step (§2.2.2), I analyzed whether the constituents differ acoustically with regard to focus condition. In the second step (§2.2.3), I conducted a perception experiment with the goal of investigating whether listeners perceptually distinguish answer pairs such as those in Figure 2.7. The results of this analysis are explicated below.

2.2.2 Acoustic analysis

Various prosodic cues have been correlated to prosodic prominence (see among others Baumann et al. 2006, Maskikit-Essed & Gussenhoven 2016, Lee et al. 2015,

Table 2.4: Type and number of constructions: in-focus constituents are indicated in uppercase letters and colored in blue, constituents that are not in focus are indicated in lowercase letters and colored in yellow.

TYPE	WH-QUESTION	ANSWER TYPE	NUMBER OF CONSTRUCTIONS
transitive	subject focus	Svo	3×
	object focus	svO	
ditransitive	subject focus	Svoo$_i$	1×
	object focus	sv**O**o$_i$	
	indirect object focus	svo**O**$_i$	

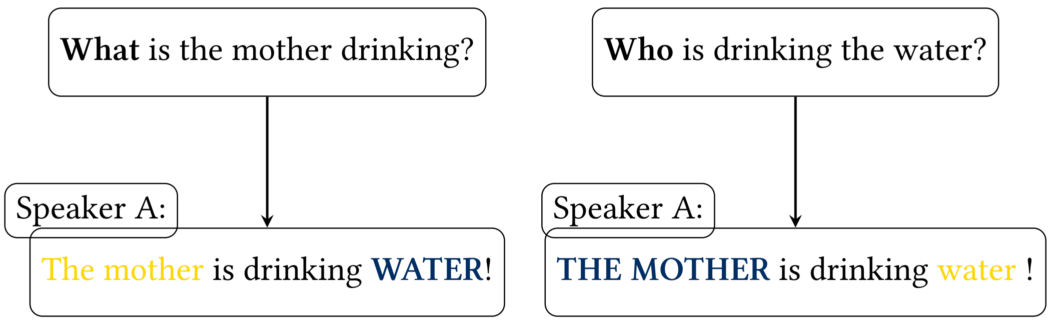

Figure 2.8: Example of pairwise analysis of answer pairs: in-focus constituents are indicated in uppercase letters and colored in blue; constituents that are not in focus are indicated in lowercase letters and colored in yellow.

Arnhold & Kyröläinen 2017). For the analysis here, I investigate duration, pitch, and intensity.

To do so, I first present and discuss data obtained from one randomly selected speaker. Following this, I present and discuss data on the two constituents *ɔgɔ* 'water' and *inaŋ* 'mother' as these occur in two different answer clauses — once in clause-initial subject position and once in clause-final object position — and in the two different focus conditions. The two clauses are given in examples (2) and (3).

(2) *inaŋ nɔŋinum ɔgɔ*
inaŋ nɔN-inum ɔgɔ
mother AV.RLS-drink water
'The mother is drinking water.'

(3) *ɔgɔ niinum inaŋ*
ɔgɔ ni-inum inaŋ
water UV.RLS-drink mother
'The mother is drinking water.'
(approx. 'The water is being drunk by the mother.')

The two constructions are a good illustration of the influence of the focus condition on the realization of the constituents in different clause positions, i.e. initial position and final position. The discussion of phonetic parameters is supported by a statistical analysis.

I ran mixed effects models with the respective parameter as dependent variable and focus as independent variable, using the *lme4* package (Bates et al. 2015) in R software (Team 2017). I included random effects for the speakers, position of the segment and the segments (random intercepts and random slopes). Furthermore, I included the valency of the constructions as a control variable because of the unequal number of observations that I obtained (three recordings for ditransitive constructions and two for transitive constructions).

2.2.2.1 Pitch and focus

To investigate the effect of focus condition on pitch contour, I measured F0 values for each constituent of the answer clauses in 30 time steps using Praat software (Boersma & Weenink 2023). It should be noted that a constituent may consist of a single word, such as *dɛuk* 'dog', or multiple words, such as *maŋana dɔlagɔ* 'the girl'. In a subsequent step, I transformed F0 values to semitones using the *HzToSemitones* command of the Soundgen package (Anikin 2019) in R software (Team 2017), with the frequency of the reference value set to 1.

Consider first the pitch contours in Figure 2.9. It displays pitch contours for nine different target constituents produced by a randomly selected speaker (speaker SRN; see Table A.3 of the Appendix). Each pitch contour is labeled according to its focus condition (indicated by color) and grammatical role (indicated by a capital letter). For instance, svO denotes a constituent that serves as the object of a transitive clause, while svoO denotes a constituent that serves as the indirect object of a ditransitive clause.

Constituents in clause-initial and clause-medial position show very similar pitch contours in both focus conditions. For example, the pitch contours of *buuk* (svOo) 'book', *maŋana dɔlagɔ* (Svoo) 'girl', *ɔgɔ* (Svo) 'water', and *dɛuk* (Svo) 'dog' demonstrate no distinguishable influence of the focus condition on their tonal realization, including both the shape of the pitch contours and the pitch range.

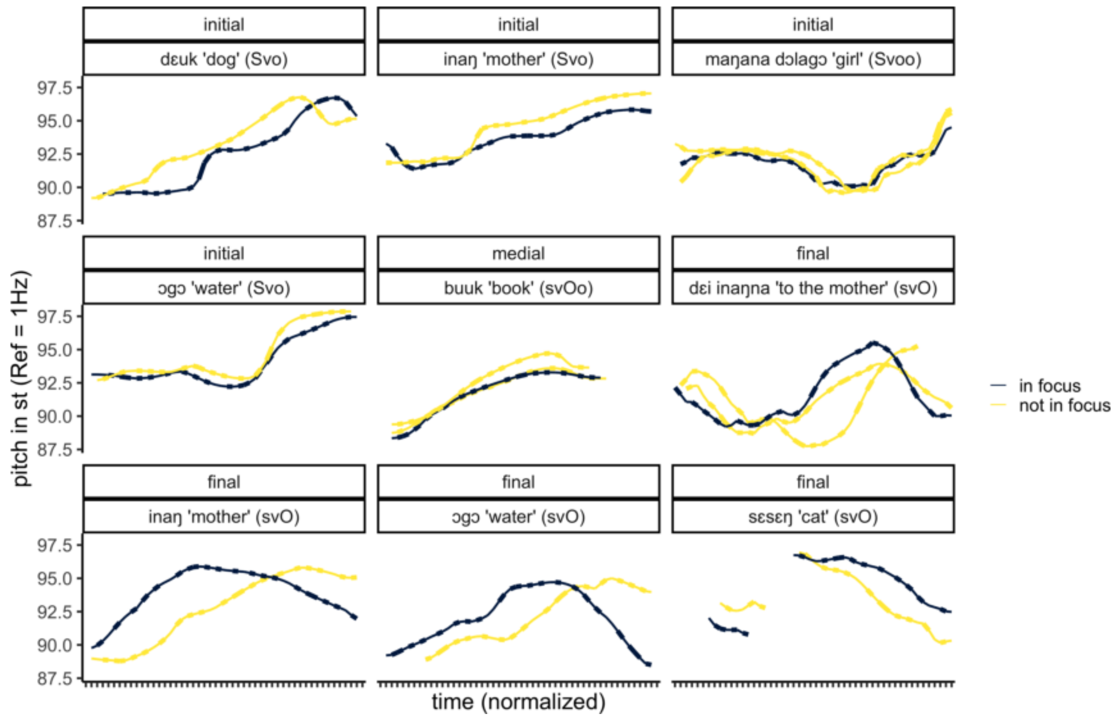

Figure 2.9: Pitch in st (Ref = 1 Hz) for the target constituents of one randomly selected speaker (SRN): focus condition is indicated by color, for ditransitive constructions, recordings have two elements in the non-focus condition, and one in focus condition, for transitive constructions, each focus condition has one recording; position and clause structure are indicated in brackets, time scale is normalized, phrase position is indicated above the target constituents, position is indicated above the target constituents, position and clause structure is indicated in brackets.

Noticeably greater variation is attested for elements in final position, such as *dɛi inaŋna* (svoO) 'to the mother', *sɛsɛŋ* (svO) 'cat', *ɔgɔ* (svO) 'water', and *inaŋ* (svO) 'mother'. These elements exhibit more diverse pitch contours, indicating that their tonal realization is more sensitive to the focus condition.

The latter two occur in both clause-final position (svO) and clause-initial position (Svo), cf. example (2) and (3) and in both focus conditions (cf. Figure 2.7).

Figure 2.10 shows these two constituents as produced by all speakers. The left-hand column shows their realizations in clause-initial position and the right-hand column shows their realizations in clause-final position. The focus condition is indicated by color.

Once again, the pitch contours of the two constituents in clause-initial position exhibit very similar shapes across the two focus conditions. However, the constituents *ɔgɔ* 'water' and *inaŋ* 'mother' in initial position display variations

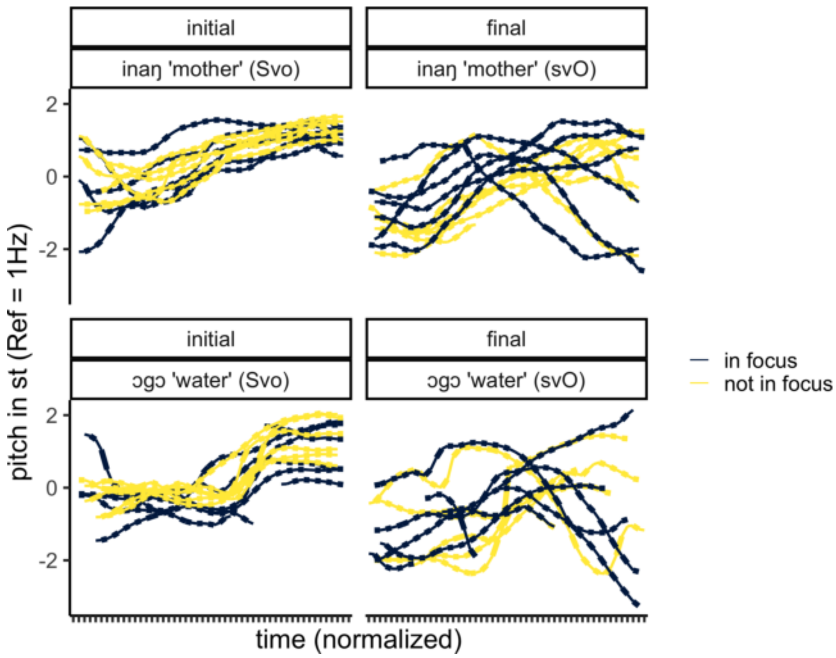

Figure 2.10: Pitch in st (Ref = 1 Hz) of the two constituents *inaŋ* 'mother' and *ɔgɔ* 'water' in the two focus conditions for all speakers: focus condition is indicated by color, position and clause structure is indicated in brackets, time scale is normalized.

in the timing of the pitch rise. Specifically, the constituent *ɔgɔ* 'water' exhibits an initial fall or level tone, a steep rise, and a final pitch plateau, while the constituent *inaŋ* 'mother' shows a relatively continuous rise. Further investigation is required to determine how the segmental material of the two constituents influences the different pitch contours in initial position. Furthermore, constituents in final position, in the right-hand column, exhibit much greater variability. Notably, a final rising boundary tone and a final rise-fall boundary tone are clearly visible (see §3.2.1). The different boundary tones may potentially be correlated with focus marking and could serve as a cue to indicate the focus condition of clause-final constituents. Table 2.5 correlates the IU-final boundary tones used with each focus condition.

An inspection of the boundary tone of the 54 CIUs revealed that out of 30 instances of a constituent in clause-final position not in focus, 12 were produced with a final rise-fall contour, while 18 had a final rise contour. For constituents in final position and in focus, out of 24 instances, 10 had a final rise-fall contour, and 14 had a final rise contour. Due to the small sample size, a correlation between final pitch movement and focus condition cannot be conclusively established. However, while it cannot be completely ruled out, if focus were indeed

Table 2.5: IU-final boundary tones per focus condition (n=54 constituents in final position)

	final rise-fall	final rise	total
not in focus	12	18	30
in focus	10	14	24
	22	32	=54

expressed by the CIU-final boundary complex, a stronger correlation would be expected. Hence, the findings suggest that the choice of final boundary tone does not indicate the presence or absence of focus.

Inasmuch as focus can be expressed by prosodic phrasing in the constructions used here, it involves variation in the chunking of Compound IUs into embedded IUs (cf. §3.2). In general, a subject in preverbal position is chunked into its own IU, and variation pertains mainly to chunking the verb and a following object NP into either one IU or two separate IUs, i.e. $[\text{ɔgɔ}]_{\text{IU}}$ $[\text{niinum inaŋ}]_{\text{IU}}$ vs. $[\text{ɔgɔ}]_{\text{IU}}$ $[\text{niinum}]_{\text{IU}}$ $[\text{inaŋ}]_{\text{IU}}$. These two possible realizations are shown in Figure 2.11 and Figure 2.12, respectively.

In both realizations the subject *ɔgɔ* 'water' is chunked into its own IU, clearly visible by the pitch rise on the last syllable *.gɔ*. In Figure 2.11, the verb *niinum* 'is being drunk' and the object NP *inaŋ* 'mother' are chunked into one IU. In Figure 2.12, however, the verb occurs in a separate IU, indicated by the pitch rise on the final syllable *.num*.

Example (3) and its two realizations in Figure 2.11 and Figure 2.12 indicate that speakers have a certain freedom with regard to chunking. Of 54 clauses, however, the realization in Figure 2.12 is the only instance in which a speaker chunks the verb into a separate IU. All other instances lump the verb in an IU together with the following object NP, as in Figure 2.11. Hence, in the recorded target clauses, focus is not expressed through phrasing or chunking of constituents into IUs.

To statistically test the influence of the focus condition on pitch, I calculated the pitch minimum, pitch maximum, and pitch range as the difference between the maximum and minimum pitch values for each target constituent (subject, direct object, indirect object) in all the target clauses. Mean F0 values were measured for each labeled vowel in the respective constituents to avoid octave jumps. F0 values were then converted to semitones to control for speaker-dependent vocal range.

The outcome of the mixed effects model (see §2.2.2) is given in Table 2.6.

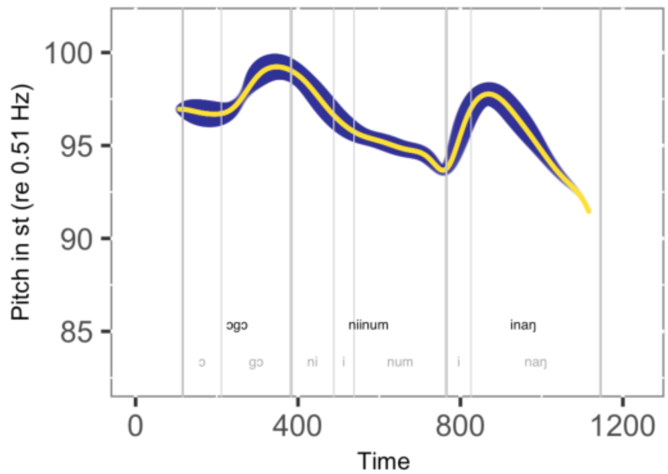

Figure 2.11: Periogram with pitch track (in st) for example (3) with narrow focus on the final constituent *inaŋ* 'mother': speaker IFS ▹

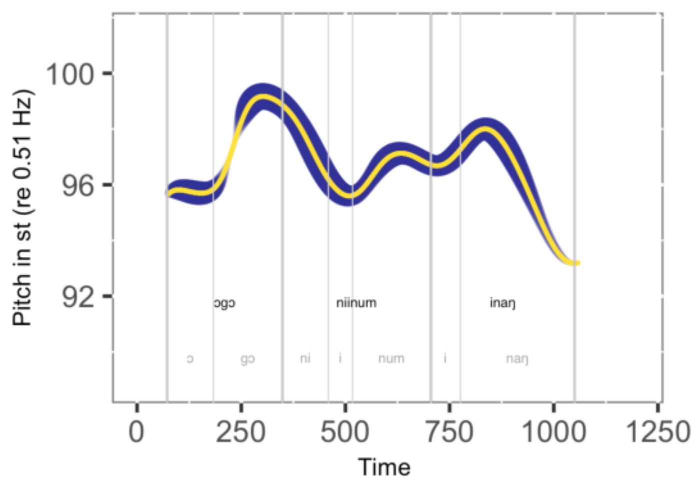

Figure 2.12: Periogram with pitch track (in st) for example (3) with narrow focus on the initial constituent *ɔgɔ* 'water': speaker IFS ▹

Table 2.6: Results of the mixed effects model with focus as independent variable and pitch range, pitch minimum and pitch maximum as dependent variables

	Estimate	SE	t	Pr(> $\|t\|$)
range	0.313	0.39	−0.793	0.442
max	0.028	0.830	0.034	0.976
min	0.657	0.942	0.697	0.555

The results of the model in Table 2.6 show that there is no statistically significant effect of focus condition on pitch range, pitch maximum, nor on pitch minimum. Furthermore, the results indicate that the target constituents are produced with a reduced pitch range (−0.313 st), and higher pitch maximum (+0.028 st) and minimum (+0.657 st) when in focus. The effects are negligible, considering that the *just noticeable difference* (*j.n.d.*) is estimated to be around 1.5–2 st ('t Hart 1981, 't Hart et al. 1990: 29).

I conclude that the focus condition has no discernible impact on the pitch contour, the pitch minimum and maximum, and pitch range.

2.2.2.2 Duration and focus

To discover any potential effect of focus on duration, I measured the duration from the onset until the end of each labeled constituent.

Similar to the above, I plotted the duration of the nine different constituents in both focus conditions produced by one randomly selected speaker (speaker ZHRM; cf. Table A.3 of the Appendix). These are given in Figure 2.13. No apparent effect of focus condition on duration is discernible.

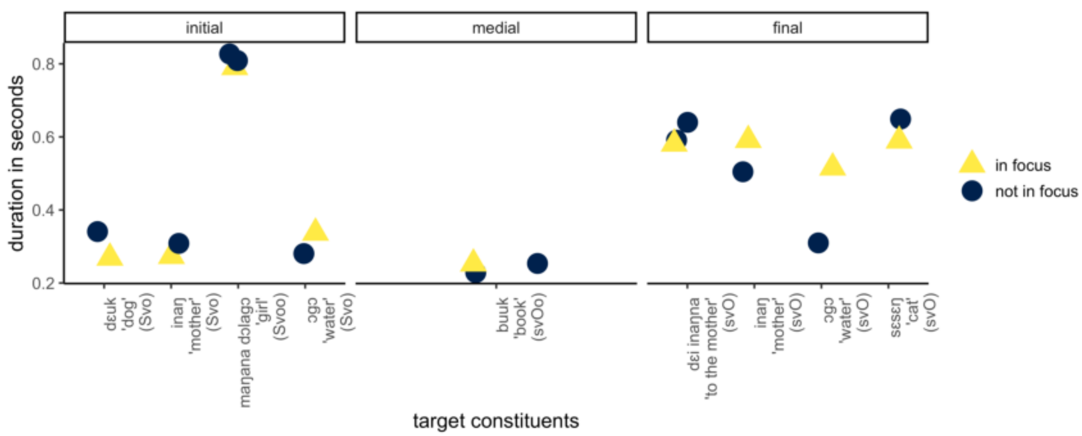

Figure 2.13: Duration (in seconds) and focus condition of the target words for one randomly selected speaker (speaker ZHRM): for ditransitive constructions, two elements in the non-focus condition, and one in focus condition were recorded; for transitive constructions, each focus condition has one recording; position and clause structure are indicated in brackets.

The constituents *buuk* (svOo) 'book', *maŋana dɔlagɔ* (Svoo)'the girl' and *inaŋ* (Svo) 'mother' show very similar values for both focus conditions. With regard to the constituent *ɔgɔ* (Svo/svO) 'water', duration is longer when in focus, especially when in object position (svO). On the other hand, *sɛsɛŋ* (svO) 'cat' shows shorter duration when in focus.

To get a better picture of any systematicity across speakers, I plotted durational values for the two words *inaŋ* 'mother' and *ɔgɔ* 'water' produced by all speakers in Figure 2.14.

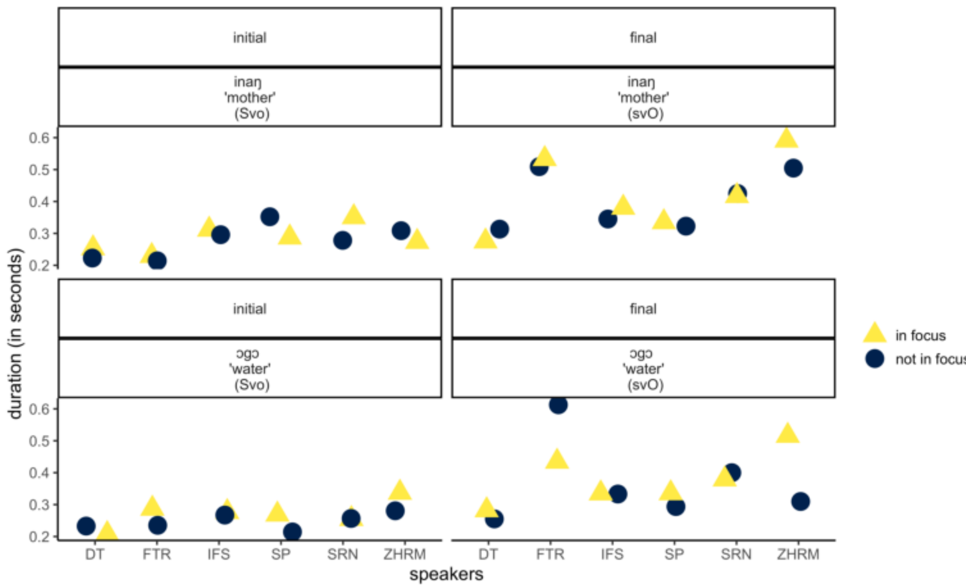

Figure 2.14: Duration (in seconds) and focus condition of the two constituents *inaŋ* 'mother' and *ɔgɔ* 'water' in the two focus conditions for all speakers; position and clause structure are indicated in brackets.

Figure 2.14 yields results similar to those obtained from the inspection of pitch contours in Figure 2.10. There appears to be a substantial effect of clause position. Elements in clause-final position are longer than elements in clause-initial position. This is indeed expected (cf. utterance-final lengthening/preboundary lengthening; Turk & Shattuck-Hufnagel 2020, Byrd et al. 2006, Berkovits 1993). However, the focus condition does not yield any apparent effect. The results of the mixed effects model (see §2.2.2) are given in Table 2.7.

Table 2.7: Estimate of the group difference of the mixed effects model with focus as independent variable and duration as dependent variable

Estimate	SE	t	$\Pr(>\|t\|)$
0.009	0.014	0.646	0.523

The results of the mixed effects model show that the effect of focus condition on duration is not significant ($p = 0.523$).

2.2.2.3 Intensity and focus

Lastly, I investigated whether there is any potential effect of the focus condition on intensity. I measured the mean intensity in decibels (dB) of each labeled constituent from the onset until the end of the constituent. As explained above, all recordings were done with head-mounted microphones in order to keep the distance from the microphone to the mouth constant (see §2.2.1). This is a necessary prerequisite for taking into account intensity. No normalization of data is needed as I am visually comparing individual speaker variation and the statistical model includes random effects for speakers.

Figure 2.15 shows the intensity values of the nine different constituents in both focus conditions as produced by one randomly selected speaker (speaker SP; cf. Table A.3 of the Appendix).

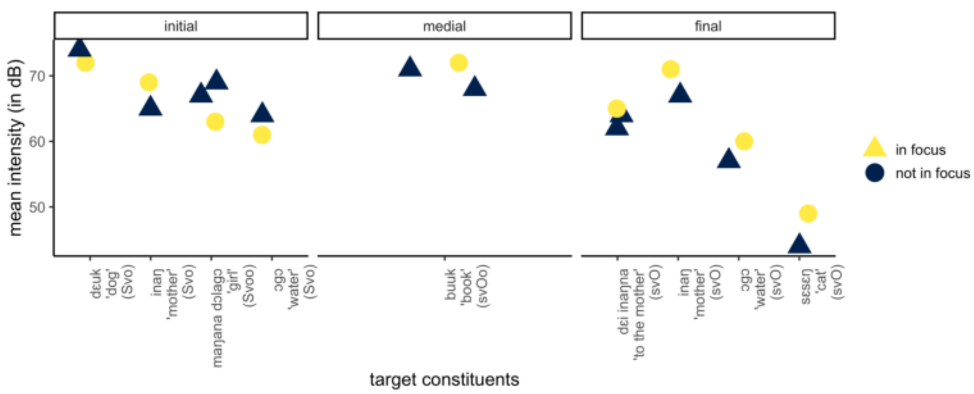

Figure 2.15: Intensity (in dB) and focus condition of the target words by one randomly selected speaker (SP): for ditransitive constructions, two elements of the non-focus condition, and one in-focus condition were recorded; for transitive constructions, each focus condition has one recording; position and clause structure are indicated in brackets.

Figure 2.15 shows that this speaker tends to produce the constituents in focus with a slightly lower mean intensity, with the exception of *dɛuk* (Svo) 'dog' and *ɔgɔ* (svO) 'water'.

To see whether this trend can be observed for other speakers as well, I plotted mean intensity values for the two constituents *inaŋ* 'mother' and *ɔgɔ* 'water' of all speakers in Figure 2.16.

Figure 2.16 shows that the constituent *ɔgɔ* (svO) 'water' appears to generally be uttered with a slightly higher mean intensity when in focus. In initial position,

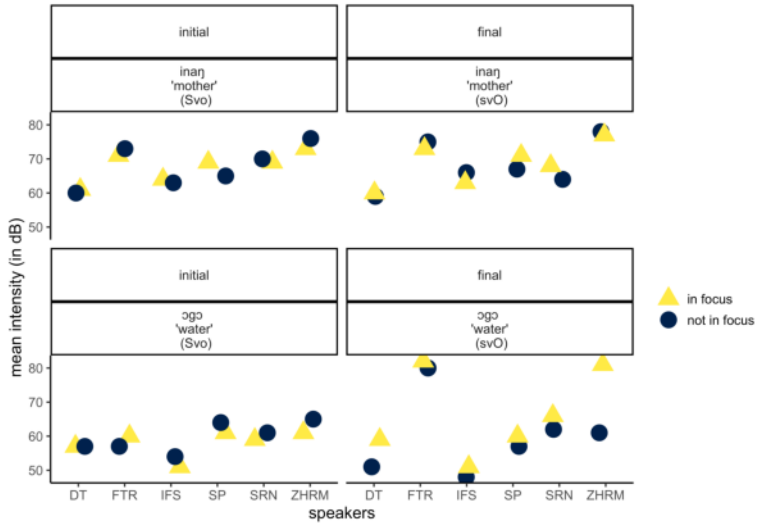

Figure 2.16: Intensity (in dB) for the two constituents *inaŋ* 'mother' and *ɔgɔ* 'water' in the two focus conditions; position and clause structure are indicated in brackets.

mean intensity tends to be lower when in focus (with one exception). No such trend is found for the constituent *inaŋ* 'mother'.

Again, also with regard to mean intensity, no clear effect of the focus condition is discernible. The results of the mixed effects model (see §2.2.2) are given in Table 2.8.

Table 2.8: Results of the mixed effects model with focus as independent variable and intensity as dependent variable

Estimate	SE	t	Pr(> |t|)
0.601	0.909	0.661	0.518

The results of the mixed effects model show that there is no statistically significant effect of focus on intensity.

2.2.2.4 Summary

The results of the acoustic analysis suggest that, in the set of clauses analyzed, focus is not prosodically encoded by the parameters tested, i.e. by pitch range, minimum and maximum, CIU-final boundary-tone complexes, phrasing, duration or mean intensity.

However, the focus condition may be expressed by other means not tested here, for example spectral tilt, variations in the tonal realization of constituents, vowel quality etc. To exclude this possibility, I conducted a perception experiment to see whether native speakers distinguish between different focus conditions.

2.2.3 Perception experiment

In the perception experiment, participants listened to two QA pairs. For the question, the same recording was used in both. The answers, however, although syntactically identical, were previously recorded as answers to different wh-questions that trigger different foci. Hence, in one of the two QA pairs, the answer was the one originally uttered in response to that particular question. In the other QA pair, there was a mismatch (Figure 2.18). In a two-alternative forced-choice experiment, participants were asked to identify the correctly matched QA pair. I expected this to be a particularly easy task if focus is encoded in the answer clauses.

2.2.3.1 Participants

Twenty participants were recruited from the Totoli community, with the prerequisite that they were fluent speakers of the language and sufficiently knowledgeable with computers: M_{Age} = 32.25; $Range_{Age}$ = 20–46. All speakers stated that they were born and raised in the Tolitoli regency (Kabupaten Tolitoli) and were raised with Totoli. As is the default in the area, they are also fluent speakers of the local variety of Indonesian/Malay and, to varying degrees, of Standard Indonesian. Further information about the participants is given in §A.2.2 of the Appendix.

2.2.3.2 Procedure

Participants listened to one correctly paired QA pair and one incorrectly paired QA pair.

The experiment was run on a laptop, using the OpenSesame platform (Mathôt et al. 2012). Participants were told that they would hear two QA pairs, of which one was correct and one was incorrect. Their task was to choose the QA pair they perceived as correct. They listened to the two QA pairs twice before making their choice. The task was repeated 72 times per participant (20 participants × 72 choices = 1440). Stimuli were presented in random order and the visual order of the two choices was randomized as well. Figure 2.17 illustrates the experimental setup.

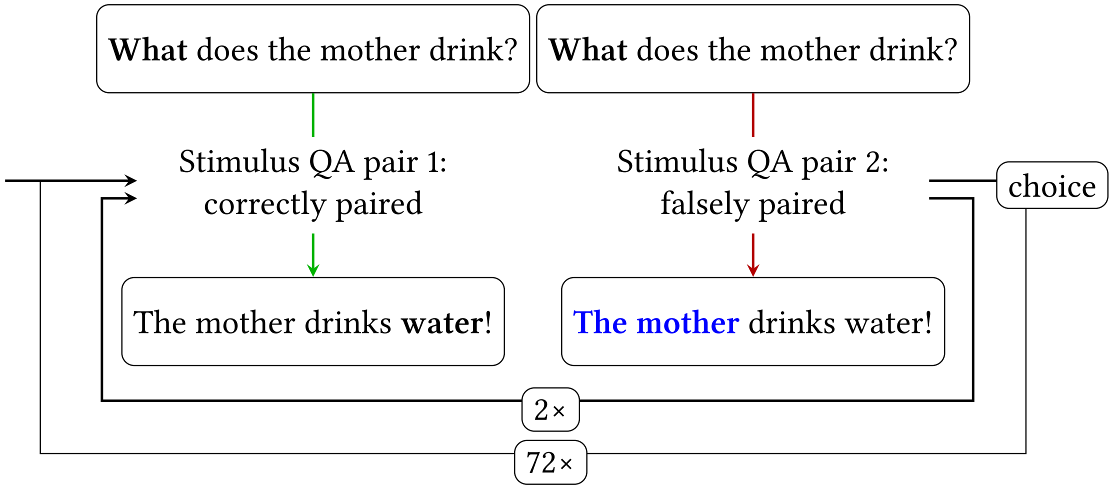

Figure 2.17: Illustration of experiment procedure

2.2.3.3 Stimulus preparation

I used two recordings of each transitive clause: one was previously uttered as the answer to the wh-question triggering focus on the subject, the other was uttered as the answer to the wh-question triggering focus on the object. For the ditransitive clauses, I had three recordings per speaker, each uttered in response to a different question that triggered focus on each of the three constituents. The constructions were summarized in Table 2.4 in §2.2.1. These constructions were recorded with 6 different speakers, yielding 56 recorded QA pairs of which the answers were cut out.

For the experiment, I combined the same question with two different answers, yielding two QA pairs.

1. One QA pair consisted of a question paired with an answer that had been previously recorded as the answer to the same question. For instance, a wh-question that triggered focus on the subject was paired with an answer that had been previously recorded in response to a wh-question that also triggered focus on the subject. This resulted in a correctly paired QA pair.

2. The other QA pair was composed of the same question paired with an answer that had been previously recorded as the answer to a question triggering a different focus. For example, a wh-question that triggers focus on the subject was paired with an answer that had been previously recorded as the answer to a wh-question triggering focus on the object. This means that it was an incorrectly paired QA pair.

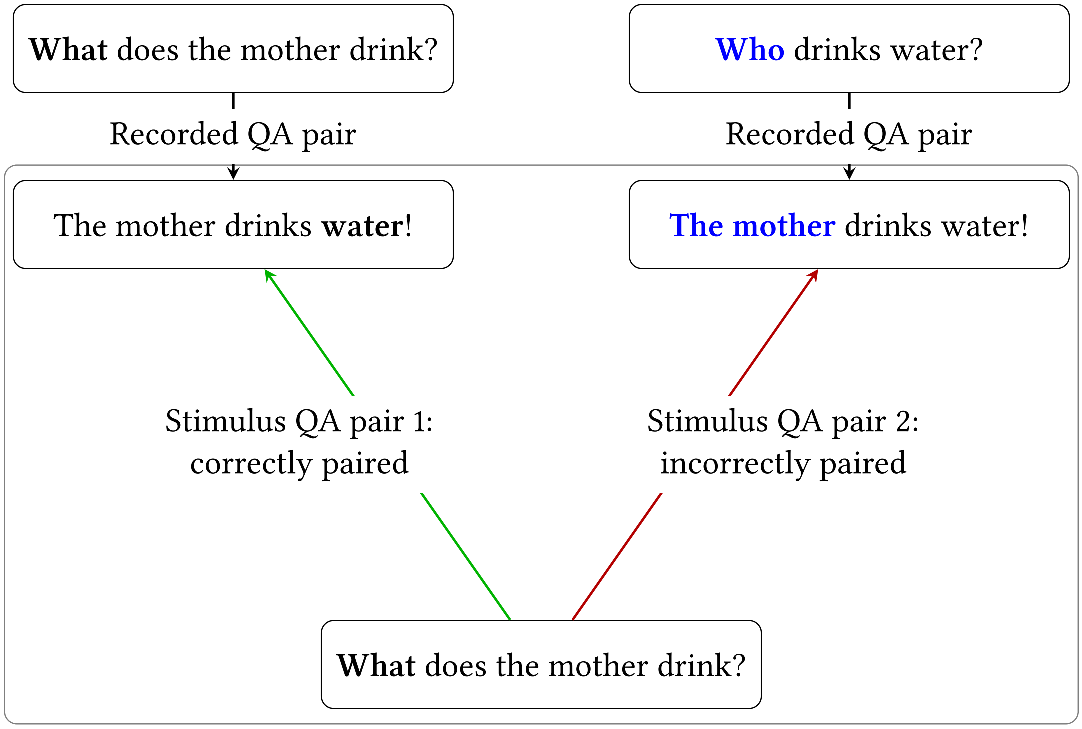

Figure 2.18: Example of pairing of questions and answers

The pairing of the question with the two answers is exemplified in Figure 2.18

For the pairing of questions and answers in the perception experiment, all recorded answers from the QA pairs of §2.2.1 were used. The pairing was automatically generated and the order was randomized using the OpenSesame platform (Mathôt et al. 2012). To control for speaker variation as a potential factor influencing the choice of the correct QA pair, both choices of answers were taken from recordings of the same person. However, the recorded speaker of the question was always different from the one who recorded the answer to create a natural situation where the questioner and answerer are different individuals.

2.2.3.4 Results

The question at hand is whether participants exhibit a significant preference for the correctly paired QA pair over the incorrectly paired items (chance level = 50%). If so, it can be assumed that the answer clauses carry prosodic cues that encode information about focus. In other words, if the prosody of Totoli encodes focus, then participants should show a preference for the correctly paired QA pairs. The distribution of question-answer assignments is depicted in Figure 2.19, which reveals that 52% of participants selected the correctly paired QA pair while 48% selected the incorrectly paired QA pair.

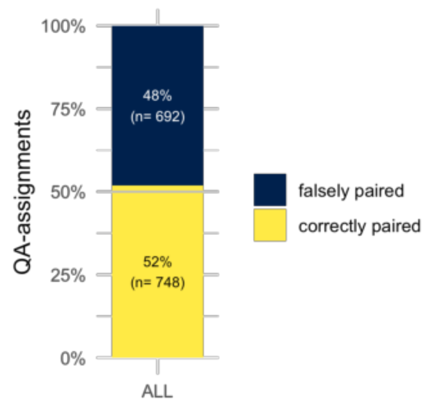

Figure 2.19: Distribution of question-answer assignments: horizontal line indicates chance level 50%, $n = 1440$.

In order to determine the significance of the participants' tendency to choose the correct answer, I conducted a logistic mixed effects model with the choice of QA pair as the dependent variable. The model did not include any predictors. In this case, the intercept is the only parameter of interest, as it measures the probability of choosing the correct answer, with an intercept of 0 on the logit scale representing 50%. As random intercepts, I included the rater and recorded speaker, as well as the focus type (see Table 2.4)). I used *glmer*() with *family = binomial* of the *lme4* package (Bates et al. 2015) in R (Team 2017).

glmer(), with family = binomial,

The estimate of the intercept was very close to 0 (0.077), indicating only a slight preference for the correctly paired QA pair. This result was not significantly different from 0 ($p = 0.183$; SE $= 0.058$; $z = 1.332$). Based on these findings, I concluded that the tendency to select the correctly paired QA pair was likely due to chance.

To further investigate participant performance, I analyzed whether there were any observable effects. Specifically, I wanted to determine whether some participants performed differently than others, whether participant performance depended on the recorded speaker, and whether performance varied depending on the position and grammatical role of the constituent in focus.

2.2.3.4.1 Effect of rater

First, I analyzed the distribution of the question-answer assignment across participants. The results are shown in Figure 2.20.

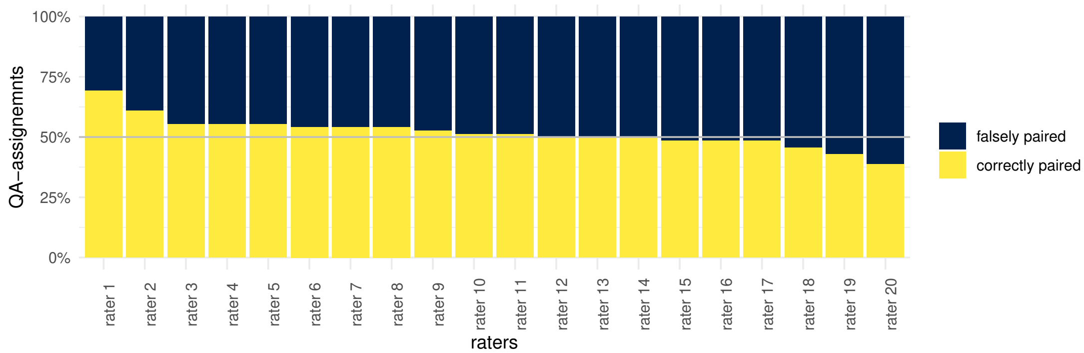

Figure 2.20: Distribution of question-answer assignments by rater: horizontal line indicates chance level 50%.

Figure 2.20 shows that the performance is comparable across participants.

One participant (rater 1) showed a particularly high frequency of selecting the correctly paired QA pair (in 69% of all instances), while another participant (rater 20) showed a higher preference for the incorrectly paired QA pair (in 39% of all instances).

2.2.3.4.2 Effect of recorded speakers

Second, I analyzed the distribution of the question-answer assignment according to the recorded speaker of the answers of the QA pairs. The results are plotted in Figure 2.21. Note that the QA pairs were paired such that the answer in both was uttered by the same speaker (cf. §2.2.3.3).

Figure 2.21 shows that there is only slight variation in task performance depending on the recorded speaker. The differences appear to be negligible.

2.2.3.4.3 Effect of position

Thirdly, I investigated the distribution of the question-answer assignment according to the position in the clause of the constituent that is focused by the question. The results are plotted in Figure 2.22.

Figure 2.22 shows that the position in the clause of the constituent that is focused by the question does not affect performance in choosing the correctly paired QA pair.

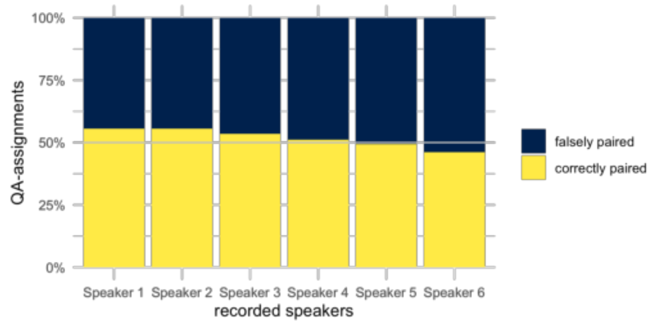

Figure 2.21: Distribution of question-answer assignments by recorded speaker: vertical horizontal line indicates chance level 50%.

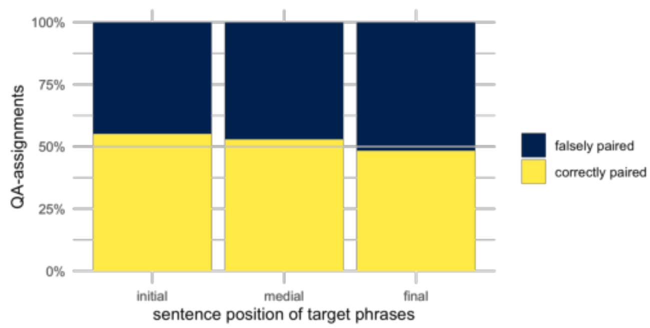

Figure 2.22: Distribution of question-answer assignments by position of the in-focus constituent in the clause: medial position means the direct object of ditransitive constructions (svOo), final position means the indirect object of ditransitive constructions and the object of the transitive constructions (svO/svoO); horizontal line indicates chance level 50%.

2.2.3.4.4 Summary

The data exploration examined three factors that may have influenced the question-answer assignment performance: rater, recorded speaker, and position in the clause of the constituent on which the question triggers focus. Despite some variability, particularly with individual raters, the analysis revealed that none of these factors had a significant impact on participants' performance.

2.2.4 Discussion

In this section, I presented an experimental analysis of the interaction of prosody and focus in Totoli. To this end, I presented an experiment that examines the role of prosodic prominence in Totoli by searching for prosodic cues in the marking of information-structural categories.

The topic was first approached by an investigation that analyzed whether the information-structural category of focus is acoustically prominent in identical constructions that were uttered as answers to questions triggering either subject or object focus (§2.2.2). The following section described a perception experiment to investigate whether native speakers distinguish between different focus conditions (§2.2.3).

In order to investigate the effect of focus condition on the production of syntactically identical constructions, I analyzed the phonetic parameters related to pitch, duration, and intensity (§2.2.2.1–§2.2.2.3). However, no significant effect of the focus condition was observed in any of these parameters. To exclude the possibility of focus condition being expressed by other means not tested here, I conducted a perception experiment to see whether native speakers distinguish between different focus conditions (§2.2.3). However, the results of the perception experiment revealed no significant difference in the perception of syntactically identical clauses recorded as answers to questions with different foci. Based on these findings, I conclude that focus is not prosodically encoded in the set of clauses analyzed.

As the present investigation focuses on a controlled experimental analysis of focus marking in Totoli, it is important to note that in less controlled situations, speakers of Totoli use syntactic means and concomitant prosodic phrasing to express focus. While an in-depth investigation of this focus marking strategy is beyond the scope of this study, two examples are provided for illustration purposes. These instances are taken from a recording of an adapted version of the Anima elicitation game described in Skopeteas et al. (2006: 99–107), which was originally designed to elicit different focus types. In this game, participants are shown pictures and then asked questions about them. Examples (4) and (5) illustrate focus marking in Totoli in this less controlled setting.

In example (4), the speaker answers to a question asking about the patient of the situation ("*In front of the well*: <u>What</u> is the man pushing?"; adapted from Skopeteas et al. 2006: 103). The speaker replies with a cleft construction in order to mark the focus on the patient *ɔtɔ* 'car'. The focused constituent occurs in sentence-initial position and is followed by a free relative clause, which is introduced by the relative particle *anu* 'REL'. Figure 2.23 shows the periogram with pitch track (in st) of example (4).

(4) a. *ɔtɔ*
ɔtɔ
car
'(it is a) car'

b. *anu laalau suludan tau mɔanɛ dɛi dulak ɔgɔbbun ana*
anu laa-lau sulud-an tau mɔanɛ dɛi dulak ɔgɔbbun ana
REL RDP-presently push-APPL person male LOC front well MED
'that is currently pushed by the man in front of the well'
(QUIS-focus_SP.041-42) ▹

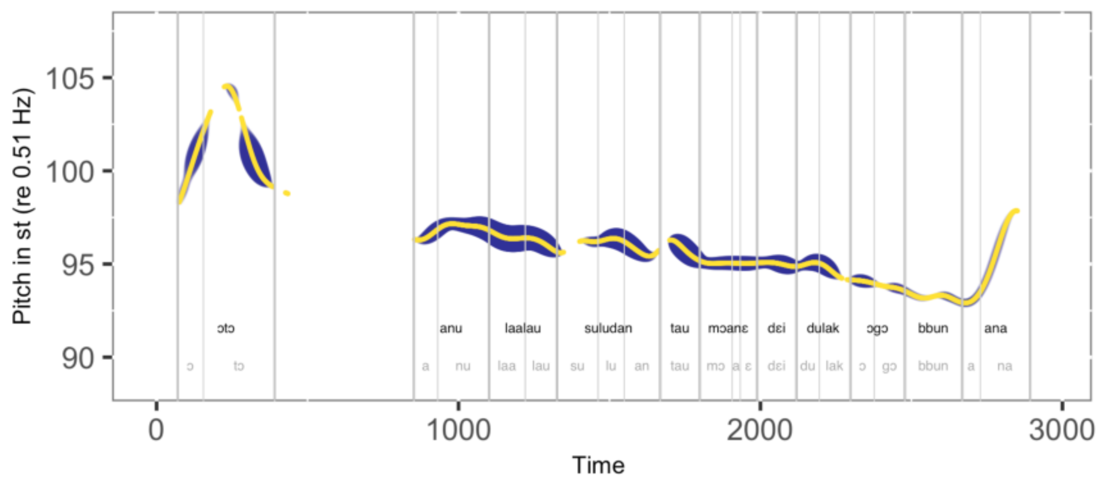

Figure 2.23: Periogram with pitch track (in st) for example (4), speaker SP

The pitch contour in Figure 2.23 depicts that the focused constituent *ɔtɔ* 'car' is parsed into its own prosodic phrase, marked by a final rise-fall boundary-tone complex with a high target located at the beginning of the ultimate syllable. The relative clause *anu laalau suludan tau mɔanɛ dɛi dulak ɔgɔbbun ana* 'which is currently pushed by the man in front of the well' is pronounced as one prosodic phrase with a final boundary-marking tonal complex, consisting of a low target located at the boundary between the penultimate and ultimate syllable, followed by a final high tone.

Another example is provided in (5), where the speaker responds to a question about the agent of the situation ("In front of the well: Who is pushing the car?"; cf. Skopeteas et al. 2006: 103). Similar to the previous example, the speaker uses a cleft construction to mark the focus on the agent *mɔanɛ* 'man', which is placed in the sentence-initial position, followed by a free relative clause introduced by the relative particle *anu* 'REL'. Figure 2.24 shows the periogram with pitch track (in st) of example (5).

(5) *mɔanɛ anu lau mɔnuludan ɔtɔ*

mɔanɛ	anu	lau	mɔN-sulud-an	ɔtɔ
man	REL	presently	AV.RLS-push-APPL	car

'It is a man who is currently pushing the car.' (QUIS-focus_SP.10) ▹

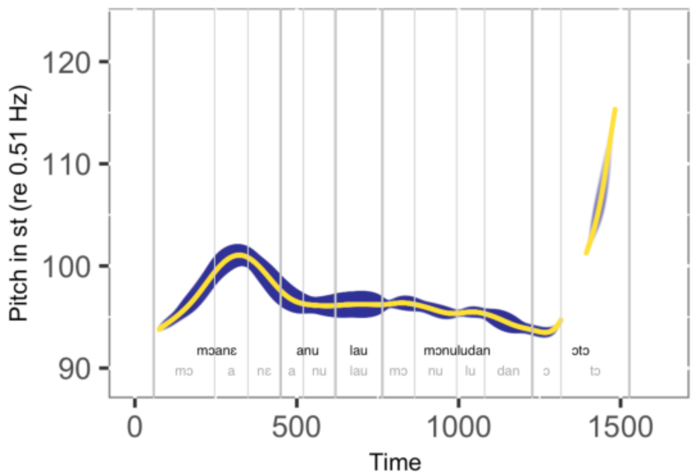

Figure 2.24: Periogram with pitch track (in st) for example (5), speaker SP

The pitch contour in Figure 2.24 shows that the prosodic realization is similar to that of example (4) above. The focus constituent *mɔane* 'man' is chunked into its own prosodic phrase, clearly demarcated by a prosodic boundary in the form of a rise-fall boundary-tone complex with a high target located at the boundary between the penultimate and the ultimate syllable. The relative clause *anu lau mɔnuludan ɔtɔ* 'who is currently pushing the car' is uttered as one prosodic phrase with a final boundary-marking tonal complex that consists of a low target located at the boundary between the penultimate and the ultimate syllable, followed by a final high tone. No pause or pitch reset occurs between the focus constituent and the relative clause. In cleft constructions such as example (4), the fronted constituent is necessarily chunked into its own prosodic phrase and therefore has to be demarcated with a phrase-final boundary tone.

The prosodic marking appears to be concomitant to the syntactic marking of focus. Crucial to this discussion is the fact that in a controlled environment, Totoli allows syntactically identical SVO constructions as answers to different wh-questions that trigger narrow focus on either the subject or the object. Yet, in such constructions where the focus structure is not syntactically encoded, it is also not acoustically prominent.

With regard to purely prosodic strategies of marking focus, Lee et al. (2015: 4754) comment that it "is less commonly recognized that purely prosodic marking of focus may be much weaker in some languages than in others, to the extent that purely prosodic focus may be nearly absent as a general mechanism for communication of information structure". Totoli apparently presents such a case.

2.3 Conclusion

The aim of this chapter was to investigate the role of prosodic prominence in the intonation of Totoli. The RPT experiment in §2.1 showed that participants generally do not agree on the judgment of prominences. Hence, similar to Papuan Malay, the results for Totoli make it "doubtful whether prosodic prominence can be usefully distinguished from boundary marking in this language" (Riesberg et al. 2018: 389). This was further tested by a subsequent focus marking experiment that included a production and a perception part, with the question whether speakers of Totoli employ purely prosodic means to mark SVO sentences with different focus structures. However, speakers of Totoli do not use purely prosodic means to express focus.

The results from the RPT experiment and the focus marking experiments show that prosodic prominence does not play a role in the prosodic system of Totoli. In other words, Totoli does not employ postlexical pitch accents to mark focus, which is similar to other Austronesian languages in the region (cf. Maskikit-Essed & Gussenhoven 2016 on Ambonese Malay, Riesberg et al. 2018 on Papuan Malay, Goedemans & van Zanten 2007 on Standard Indonesian). One of the few studies that specifically investigated phrasal prominence and the realization of focus is Maskikit-Essed & Gussenhoven (2016), who conducted a study on Ambonese Malay. They elicited scripted speech in the form of question-answer mini-dialogues that were controlled for focus condition. In the target words analyzed, they came to the conclusion that "in effect, this means that Ambonese Malay does not express focus in its prosody" (Maskikit-Essed & Gussenhoven 2016: 383) and that there is no evidence for stress in general in that language. Furthermore, they hypothesize that their findings may actually apply also to other Malay varieties Maskikit-Essed & Gussenhoven (2016: 391). The absence of (post)lexical stress, however, may not only be a feature found in many Indonesian/Malay varieties but may likewise be found in many other languages of the region, as the results from Totoli suggest.

The experimental results obtained here are highly relevant for the following chapter, in which I turn to an investigation of the tonal realizations of CIUs of a corpus of Totoli (§3.2.1–§3.2.4), and in which I propose an intonational model of the CIU (§3.2).

3 Corpus-based approaches to the intonation of Totoli

In this chapter, I investigate the segmentable prosodic units as they occur in the corpus of (semi-)spontaneous speech of Totoli. Based on the analysis of tonal patterns in §3.2, I conclude that in Totoli we have to assume recursive embedding of IUs into complex Compound IUs rather than IUs that are parsed into lower-level prosodic units. Hence, the label CIU here. The CIU in my analysis is hence equivalent to the label IU in other studies (Croft 1995, 2007, Tao 1996, Park 2002, Schuetze-Coburn et al. 1991, Schuetze-Coburn 1994, Matsumoto 2003, Iwasaki 1996, Iwasaki & Tao 1993, Wouk 2008). For the sake of clarity and readability, I use the label CIU for both complex CIUs that consist of several embedded IUs and also for singleton IUs (see Figure 3.4).

In §3.1, I describe some of the fundamental properties of the CIU in the corpus, including their categorization, distribution, and length. Section §3.2 presents an intonational model and discusses the tonal specifications of boundaries of prosodic units. Finally, in §3.3, I investigate the prosody-syntax interface by examining the syntactic content of CIUs as a whole (§3.3.1) and the embedded IUs of CIUs in particular (§3.3.2).

3.1 Properties of the Compound Intonation Unit in Totoli in a cross-linguistic perspective

From the discussion in §1.4, it is clear that the IU is "the spoken-language analyst's most popular unit of choice for analysis" (Croft 1995: 841) and it is considered the basic unit structuring discourse. Furthermore, it is locally managed and "different sizes of IUs are used in different interactional contexts" (Park 2002: 674). This section explores some properties of the 3226 CIUs in the corpus of Totoli and compares these with data reported for other languages.

3.1.1 The corpus

The corpus consists of 21 selected recordings that were segmented and annotated by me according to the criteria described in §1.4. All recordings were made using head-mounted microphones worn by the consultants, with an additional recording on the built-in camera microphone. Video and audio were recorded with a Zoom Q8 audio/video recorder with two external AKG C520 head-mounted condenser microphones at a sampling rate of 48 kHz.

I distinguish between conversational and monological data, as they are often theorized to differ substantially in various ways, such as "information pressure" (Du Bois 1987: 836). In this seciton, frequency distributions of various phenomena are displayed for the entire corpus, as well as for conversational and monological data separately. Inclusion of both types of data ensures that certain prosodic phenomena that may occur infrequently, if at all, in either genre are accounted for. Analyzing these data categories individually allows us to investigate how various types of genres influence the frequency of tonal events and syntactic structures. By taking this approach, we can draw more nuanced and detailed conclusions about the relationship between prosodic and syntactic phenomena and discourse genre.

The corpus includes recordings from 15 speakers: 11 male and 4 female. Further information on the speakers are given in Table A.6 and Table A.5 in the Appendix. Table 3.1 below gives an overview of the recordings.

Table 3.1: Overview of recordings

interactivity type	*n* CIUs	duration
conversational	1327	00:40:08
monological	1899	01:30:08
total	3226	02:10:16

Read, elicited, or otherwise highly planned speech such as data obtained from a laboratory setup are not included in the corpus. I used this type of data only in the experiments on focus marking, which are described in §2.2.

Obtaining natural conversational data while speakers are wearing head-mounted microphones is rather difficult, especially when working with endangered languages such as Totoli. To overcome this difficulty, I used two different elicitation games to obtain the conversational data:

The Animal Game is an elicitation game described in Skopeteas et al. (2006: 111–117) with the original purpose of eliciting narrow/contrastive focus. A stack of cards with photos is divided equally between two speakers who take turns in describing the different pictures on the cards. I also used the game in a monological setting with one speaker only.

Man and Tree & Space Games is a classic elicitation game, originally designed to explore spatial reference in field settings (Levinson et al. 1992). It has proven to be a very interactive task where two participants are each presented with identical sets of 12 cards displaying different items. Without seeing the interlocutor's stack of photos, both participants have to describe the relevant details of a certain card to find an exact match. Once they are certain they have found the matching one, they put the card aside. After all cards have been described, they are checked to see whether or not the cards match. Participants usually, but not obligatorily, take turns in describing. The game involves four rounds.

The monological data comprise recordings of various genres. All data were recorded in a face-to-face situation with a local member of the Totoli documentation team, with both parties wearing head-mounted microphones. The recordings were of different types:

The Pear Story is a short movie, designed by Chafe (1980). The corpus contains several recordings of retellings of the film. Participants watch the movie first and are then asked to narrate the story-line.

Anima is an elicitation game described in Skopeteas et al. (2006: 99–107) with the original purpose being to elicit different focus types. In the game, participants are asked a set of questions about the photos they are seeing.

The Animal Game is an elicitation game described in Skopeteas et al. (2006: 111–117) for eliciting narrow/contrastive focus. The speaker receives a stack of cards with photos on them and then simply describes the different pictures on the cards.

Stories and folktales were recorded with three different speakers. These include, firstly, a recording of a lengthy account of a particularly memorable period of the speaker's life, and, secondly, five folk tales.

Explanatory texts were obtained from two different speakers, each describing an important cultural event, namely wedding traditions and a special ritualized singing game called Lelegesan. In this game, two or more singers

> "spontaneously produce as many rhyming two-liners as possible" (Riesberg 2019: 83).

It is important to note that the language studied in this research is endangered and has only few remaining speakers. In such circumstances, opportunities for data collection are limited, and an optimal setup may not be attainable. In this study, this led to, e.g., an unavoidable imbalance between conversational and monological data, which must be taken into account when interpreting the results presented in this chapter.

3.1.2 Distribution of Chafeian IU types in the corpus

Chafe (1994: 63–64) proposes a categorization scheme of IUs into different types:

Substantive IUs are those which convey events, states or referents (e.g. examples (47b), (43a), (15), (17)).

Regulatory IUs regulate interaction and the flow of information (e.g. examples (42a), (43b), (46a), (44b)).

Fragmentary IUs are unsuccessful IUs that are truncated or abandoned (e.g. (25c)).

Regulatory IUs are further grouped into textual (e.g., *and, then, well*), interactional (e.g., *mhm, you know*), cognitive (e.g., *let me see, oh*) and validational (e.g., *maybe, I thin*k) (Chafe 1994: 65). In Totoli, frequent items of this category are discourse particles and interjections such as *mh*, *io*, *aih*. Other items included are connectors such as *bali* 'so/then' and the filler element *anu*.

Table 3.1 shows the distribution of the three types in the corpus of Totoli. The category "Rest" includes uncodable or unintelligible utterances.

The frequency distribution indicates that 71.9% of the 3226 CIUs in the corpus are of the substantive type, while 15.9% are regulatory and 6.6% are fragmentary CIUs. The majority of CIUs in both settings are of the substantive type. However, the distribution of CIU types within each setting differs: the proportion of regulatory CIUs is 12.8 percentage points higher in conversational data than in monological data, and the proportion of substantive CIUs is 12.7 percentage points lower.

The frequency distribution of Chafeian IU types is rarely reported. However, in a study on Japanese two-party conversations, Matsumoto (2003: 50) found that

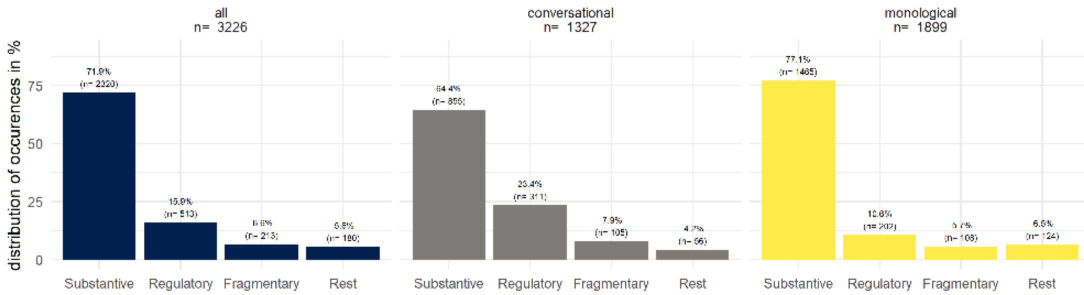

Figure 3.1: Frequency distribution of CIU types

81% of the IUs were substantive, 17% were regulatory, 0.9% were fragmentary, and 1.1% were uncodable.

Chafe (1994: 63) has suggested that this categorization "is useful because certain aspects of an analysis can be directed at one of these types to the exclusion of the others." However, the distinction between regulatory and substantive IUs is not always clear-cut. While Croft (1995, 2007) agrees that the IU is the basic discourse unit, he does not adopt the Chafeian categorization and provides his own set of criteria. Similarly, Tao (1996: 59) has stated that "in order to avoid any arbitrary decisions, I have chosen not to discriminate between the two types of IUs but instead provide a detailed grammatical taxonomy of all IUs." For the Totoli corpus, I have utilized the Chafeian classification. Overall, I found that the vast majority of IUs were easily classifiable.

3.1.3 The length of Intonation Units of the corpus

The length of IUs has received a lot of attention in the literature and has been used as a central argument for the IU as cognitive unit (see §3.3). Yet, little effort has been made to measure the length of an IU, as it is not a straightforward task. Several ways of measurement are conceivable. Chafe (1994: 64) discusses the number of words an IU contains as the "simplest and most obvious measure."

With regard to English, one is left with several figures:

- In an early publication, Chafe (1980: 14) states the average length of all IUs taken together to be about 6 words.

- In Chafe (1988: 42), he suggests that "(t)he intonation units of ordinary spoken language show a relatively constant length [...]. In English the mean length of an intonation unit is between 5 and 6 words."

- In Chafe (1993: 39), he states that "one finds the modal length of regulatory units to be one word and that of substantive units to be five words."

- In his account of the English Pear Story corpus, Chafe (1994: 64) finds Regulatory IUs to have a mean length of 1.36 words, and Substantive IUs have a mean of 4.84 words with a modal length of 4.

- Croft (1995), referencing Altenberg (1987: 282) and Crystal (1976: 256), reports counts of 4 and 6 words, respectively.

Disregarding the counts for regulatory IUs, we can summarize that an IU in English roughly contains about four to six words. Importantly, these counts hold for English only and languages vary with regard to the average number of words an IU contains.

With regard to typologically diverse languages, mean numbers of words per IU vary from two to five.

- Tao (1996: 52–54) reports an average of 3.5 and a modal length of 3 words for Mandarin Chinese.

- Chafe (1994: 148) reports a modal and an average length of 2 words for English.

- Himmelmann et al. (2018: 224) report average numbers of words per IU in 4 typologically unrelated languages: 5.13 for German, 3.73 for Papuan Malay, 3.37 for Wooi, and 3.44 for Yali.[1]

Commenting on the difference in length of IUs in the four investigated languages, Himmelmann et al. (2018: 222) note that the main difference lies in the grammatical structure, but that the orthographic conventions of the languages also play a role.

Figure 3.2 shows their distribution of substantive CIUs in the Totoli corpus and gives information on their length, measured in number of words. Both conversational and monological CIUs are included. The distribution of CIUs shows that the modal length is 2 words in conversational data and 3 words in monological data. Monological data is less skewed and CIUs with more than one word are much more common.

Measuring the length of IUs in words has its drawbacks. As the data for Totoli show, the length of CIUs varies substantially according to the data type used,

[1]Only the consensus values of expert annotators are given here.

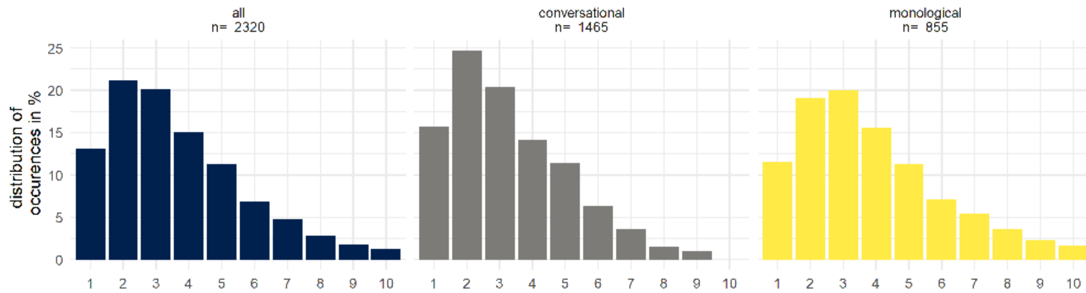

Figure 3.2: Distribution of substantive CIUs of the corpus according to lengths in words

even within the same language. Furthermore, the grammatical structure and orthographic conventions of any given language will result in different counts that may render a comparison difficult.

As Fenk-Oczlon & Fenk (2002: 222) put it:

> In languages with a pronounced tendency to synthetic (agglutinative or fusional) morphology we have to expect a lower number of words per intonation unit (and in polysynthetic and incorporating languages even one long word that we would encode in a sentence comprising 5 or 6 words.)

This touches on topics of wordhood (Tallman 2020) and alternative measures of the length of an IU have been suggested. Research on language acquisition has long focused on the length of "utterances" and on the question of what to measure (for an overview of different measures, see Allen & Dench 2015).

A possible alternative is to measure the number of syllables an IU contains (Schuetze-Coburn 1994: 161). Himmelmann et al. (2018) propose another alternative for measuring the length of IUs in terms of content words. Probably the most straightforward way, yet also the one most susceptible to speaking style and individual speaker differences, is to measure the length of IUs in terms of duration in seconds. Fenk-Oczlon & Fenk (2002: 221) cite Määttä (1993), who reports an average length of a "breath group" of about 2.1–2.2 seconds. Chafe (1980) reports a mean length of 2 seconds (incl. pauses).

I include descriptive statistics of the CIU in Totoli in terms of the different measurements – number of words, syllables and total duration – for each of the Chafeian (1994) IU types; see Table 3.2.

In this section, I have described some of the fundamental properties of the IU in general and the CIU in Totoli specifically. In the next section, I analyze the

Table 3.2: Description of CIUs in the corpus in terms of number of words, syllables and total duration in seconds, divided over Chafeian (1994) IU-types and data types

	all			conversational			monological		
	n words	*n* syll	dur	*n* word	*n* syll	dur	*n* words	*n* syll	dur
substantive									
mean	3.81	8.90	1.48	3.43	7.80	1.26	4.04	9.54	1.61
median	3	8	1.24	3	7	1.07	3	8	1.35
sd	2.35	5.31	0.92	2.35	4.39	0.77	1.99	5.68	0.97
regulatory									
mean	1.24	1.86	0.52	1.2	1.85	0.50	1.3	1.87	0.57
median	1	1	0.40	1	2	0.37	1	1	0.48
sd	0.77	1.68	0.61	0.55	1.19	0.70	1.05	2.31	0.43
all									
mean	3.24	7.31	1.26	2.75	5.91	1.02	3.60	8.35	1.45
median	3	6	1.04	2	5	0.84	3	7	1.21
sd	2.32	5.46	0.93	1.95	4.52	0.80	1.50	5.86	0.98

tonal specifications of CIUs in the corpus and propose an intonational model thereof.

In order to conduct a corpus-based analysis, I segmented the corpus into Compound Intonation Units, for which I described the criteria in §1.4. Whether native and naive listeners actually perceive the segmented Compound Intonation Units as such can be answered by revisiting the results from the RPT experiment (§2.1). Table 2.1 in §2.1.1 provides an overview of the stimuli used in the RPT experiment. In the experiment, participants rated 71 speech samples, of which 26 consisted of 2 CIUs and 4 consisted of 3 CIUs, according to the segmentation criteria applied to the corpus. Hence, I collected boundary judgments for 105 words in final position of a CIU. The 71 boundary judgments for words in final position of a speech sample were discarded, resulting in 34 boundary judgments for CIU-final but not stimulus-final words.

Figure 3.3 shows the b-scores (the proportion of speakers who marked the respective word as prominent or as preceding a boundary, cf. §2.1.4) for words in CIU-final position in comparison to b-scores for words in CIU-internal position.

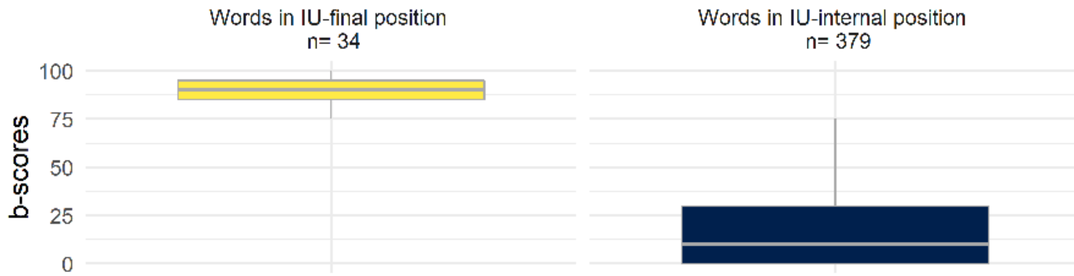

Figure 3.3: Comparison of b-scores of words in CIU-final and in CIU-internal position: median = 10, mean = 20.91 for words in CIU-internal position; median = 90, mean = 85.44 for words in CIU-final position; words in final position of a speech sample are excluded

The boundary scores for words that occur in CIU-final position are substantially higher (median = 90; mean = 85.44) than those for words in CIU-internal position (median = 10; mean = 20.91). This correlation shows that what I considered a Compound Intonation Unit in the segmentation of the corpus is actually perceived as a unit, and thus confirms the viability of units obtained from the corpus segmentation.

3.2 An intonational model of the Compound Intonation Unit in Totoli

Having described the corpus in detail, I will now turn to the study of Totoli intonation. This study focuses on the Compound Intonation Unit, which is – to repeat Chafe's (1987: 22) definition of an Intonation Unit – a "sequence of words combined under a single, coherent intonation contour, usually preceded by a pause."

In this section, I propose an intonational model of the CIU in Totoli. The model is couched in the autosegmental-metrical framework and the ToBI framework (Arvaniti & Fletcher 2020, Jun 2005c, 2014a, Ladd 2008) and assumes singleton Intonation Units or Compound Intonation Units. The model takes up Ladd's (2008: 297; chapter 8.2) notion of the Compound Prosodic Domain (CPD), i.e. strings of IUs which are recursively embedded and together form a CIU. Singleton IUs or CIUs are strings of words that are combined under a coherent intonation contour, usually preceded by a pause (Chafe 1987: 22). They are perceived as such by listeners due to a complex interplay of cues mainly pertaining to pitch, rhythm and voice quality (Schuetze-Coburn 1994: 93–155, Himmelmann 2006: 260–270, Du Bois et al. 1992, 1993, and Cruttenden 1997: 29–39). Tonal specifications are assigned at the level of the IU, and they are associated with their right-edge boundary and consist of a bitonal edge-tone complex. The right-edge of an IU – a singleton IU or an embedded IU of a CIU – is demarcated by one of the three proposed boundary tones (see §3.2.1 and §3.2.2). In a CIU, only the last embedded IU is followed by typical final cues, such as e.g. pause and pitch reset (Schuetze-Coburn et al. 1991: 217). Tonal specifications of singleton IUs and embedded IUs are equal and vary only with regard to their frequency distribution (see §3.2.5).

The intonational model for singleton IUs and CIUs in Totoli is shown in Figure 3.4. It brings together insights from the two experiments described in Chapter 2 and findings from a large scale investigation of tonal events and syntactic content of the prosodic units presented in the remainder of this chapter (§3.2.1–§3.2.4).

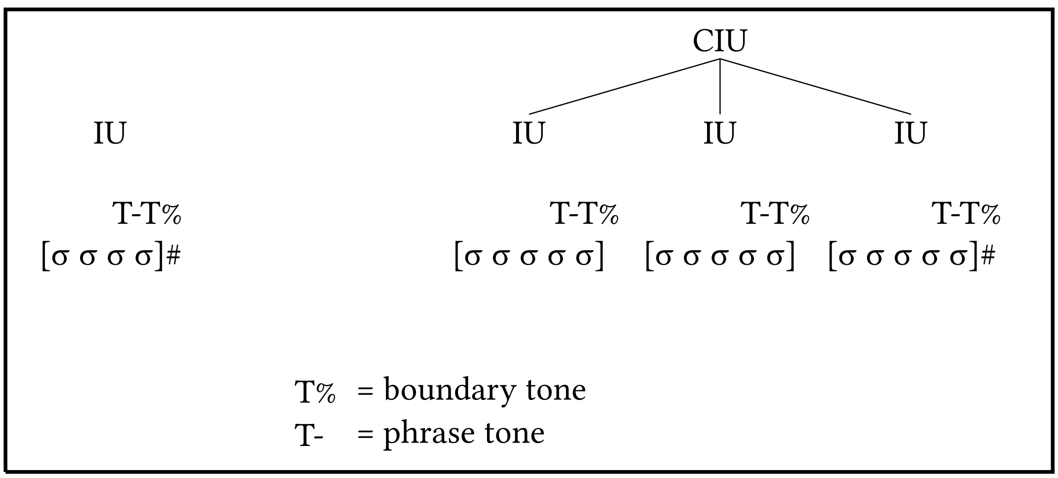

Figure 3.4: The CIU in Totoli: A singleton IU on the left-hand side, recursively embedded IUs into a Compound Intonation Unit (CIU) on the right-hand side.

The model analyzes IU-final pitch events as bitonal boundary-tone complexes, consisting of a phrase tone (T-) and a boundary tone (T%). In §3.2.2, I show that the right-edge boundary of an embedded IU of a CIU is not merely classified by a single boundary tone such as an H% or T%. Instead, I propose a set of three different final boundary-tone complexes that are essentially the same as those occurring at the right-edge boundary of the last IU of a CIU or a singleton IU respectively.

In the model, I use the more theory-neutral label *phrase tone* instead of *phrase accent* that is typically anchored to a metrically strong syllable (Grice et al. 2000).

The alignment of this tone is roughly the prefinal syllable but lacks near-constant timing (see Maskikit-Essed & Gussenhoven 2016: 356). This will become obvious from the discussion in the following sections §3.2.1 and §3.2.2.

The IU – singleton IUs and embedded IUs of CIUs – regularly maps onto syntactic or grammatical units, such as a subject or object NP, a verb or a VP, an adverbial phrase or a complement clause. This observation will be discussed in §3.3.

The only tonal event in singleton IUs is the obligatory final boundary-tone complex. In CIUs, each embedded IU is marked by a final boundary-tone complex. The syntactic content is decisive in whether a construction is uttered as a singleton IU or a CIU consisting of several embedded IUs (see §3.3). Hence, the difference between embedded IUs of CIUs and non-embedded IUs – singleton IUs or final IUs of CIUs – lies in their co-occurrence with other boundary phenomena such as pitch reset, final lengthening, pauses and glottalization. In the remainder of the work, I will use CIU as a cover term to refer to both singleton IUs and Compound IUs – hence, those which co-occur with other boundary phenomena – as opposed to embedded IUs of CIUs.

An important observation is that the tonal marking at the right-edge boundary of singleton IUs, embedded non-final IUs of CIUs, and final IUs of CIUs is essentially the same, as demonstrated in the following sections §3.2.1 and §3.2.2.

3.2.1 Tonal events at the right-edge boundaries of CIUs

In this section, I discuss tonal events occurring at the right-edge boundary of CIUs, i.e., the final IUs of CIUs and singleton IUs respectively, as exemplified in Figure 3.5.

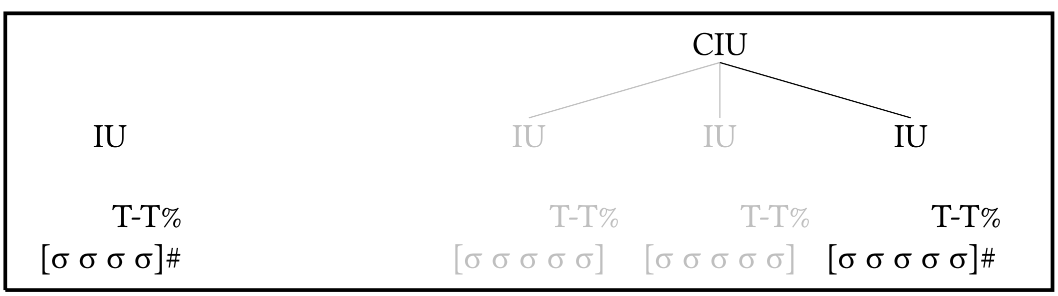

Figure 3.5: Visualization of final IUs of CIUs and singleton IUs discussed in this section

Tonal events at the right-edge boundary of CIUs and singleton IUs can be classified into a set of three different tonal contours. To classify and annotate their pitch contours, I visually inspected all IUs. The summarizing plot of these is shown in Figure 3.6. The three contours are explained in the following.

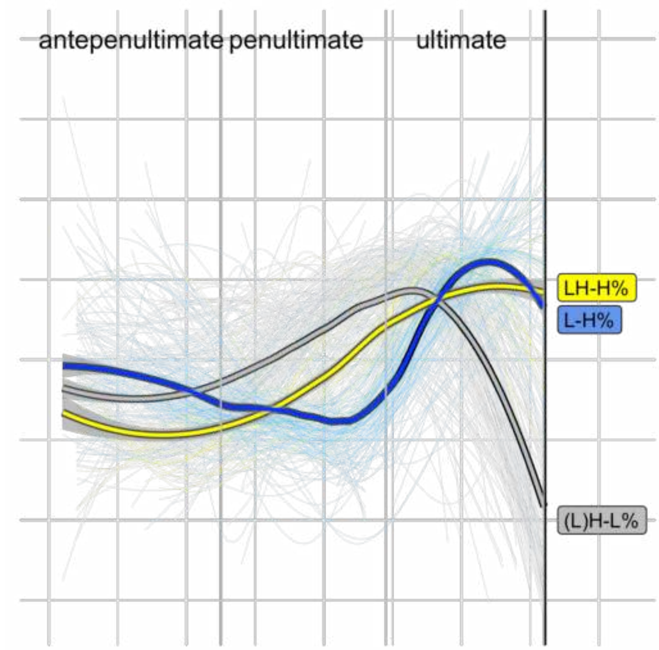

Figure 3.6: Spaghetti plot showing the intonation contours of the final three syllables of each of the three final boundary-tone complexes with the items superimposed on each other and with an average contour produced by *Loess* smoothing in R (Team 2017). Vertical bars indicate syllable boundaries; only CIUs with final $CV.CV.CV]_{CIU}\#$ syllable structure are displayed; values are z-transformed for CIU.

In the model proposed here, the final boundary-tone complex is analyzed as consisting of a phrase tone (T-) and a boundary tone (T%). The phrase tone can be a low tone L-, a high tone H- or a rising tone LH-. The boundary tone can be either a low tone L% or a high tone H%. Figure 3.7 depicts schematic versions of the three possible combinations of a phrase tone and a boundary tone, including two rising patterns and one falling pattern.

The two rising patterns are rather similar and their main difference is the domain of the pitch rise, or the low tone L- respectively. However, they are clearly distinct in their function, as I show below (cf. Table 3.3).

In the following section, I will provide a brief summary of each contour and illustrate them with examples from the corpus. After that, I will discuss their functions and distribution in the corpus. For the purpose of exemplification, only singleton IUs will be displayed in the figures below.

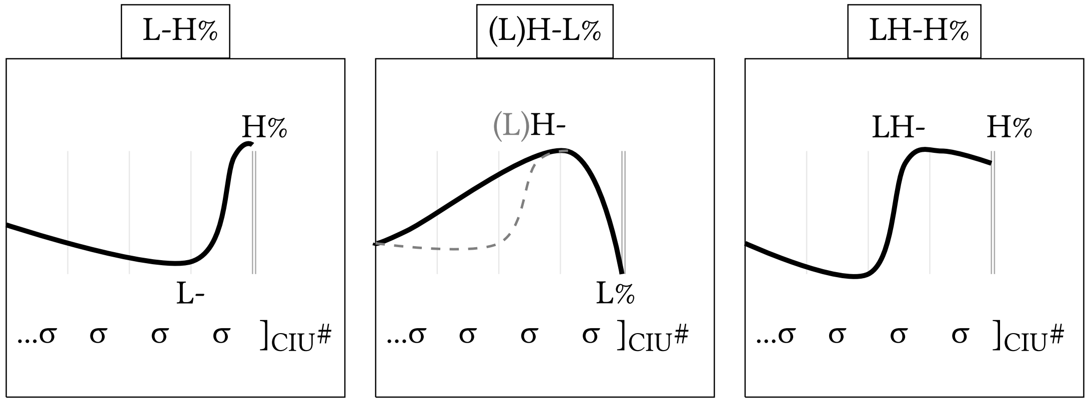

Figure 3.7: Schematic representation of CIU-final boundary-tone complexes: vertical bars indicate syllable boundaries.

3.2.1.1 The L-H% boundary-tone complex

An instance of the L-H% boundary-tone complex is given in the periogram of example (1) in Figure 3.8.

Pitch starts around the middle of the speaker's current range. Over the initial 15 syllables, pitch remains near level and then drops 4 st towards the low target of the L- phrase tone located at the boundary between the penultimate (*.ni.*) and the ultimate syllable (*.pu#*) of the IU. On the last syllable, pitch rises almost 19 st to the high target of the H% boundary tone.

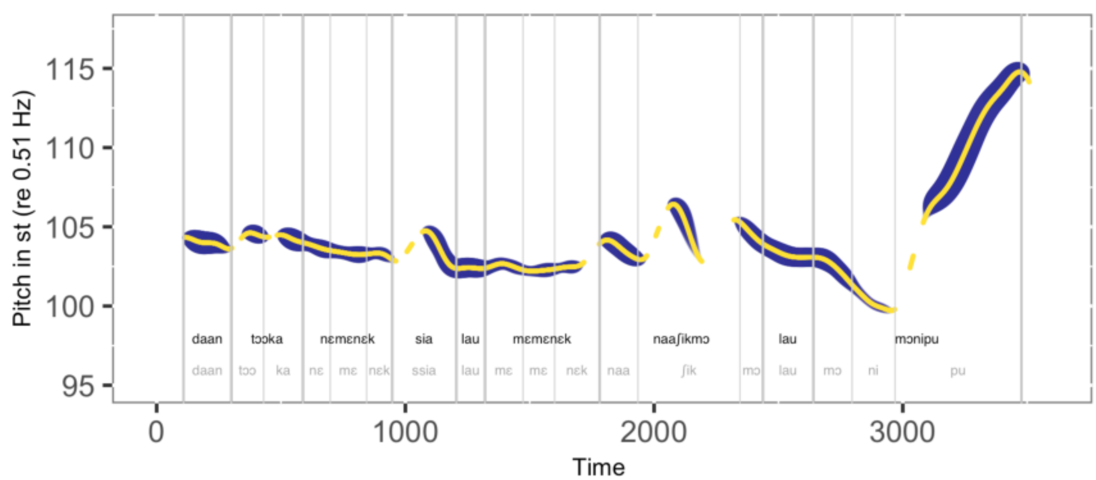

Figure 3.8: Periogram with pitch track (in st) for example (1), with final boundary-tone complex L-H%, speaker SELP

(1) *daan tɔɔka nɛmɛnɛk isia lau mɛmɛnɛk naaʃik lau mɔnipu*
daan tɔɔka nɔN-pɛnɛk isia lau mɔN-pɛnɛk nɔ-aʃik
later finished AV.RLS-climb 3S presently AV-climb ST.RLS-busy
lau mɔN-tipu
presently AV-pick
'after he climbed; eagerly picking (pears)' (pearstory_36_SELP.015) ▷

Speaker SELP shows a very high pitch range, especially on the final syllable (*.pu#*). This is not uncommon in the corpus and is frequently observable when speakers are very engaged in their conversation or narration (cf. example (15)) / Figure 3.29, example (18) / Figure 3.32, example (17) / Figure 3.31).

3.2.1.2 The (L)H-L% boundary-tone complex

I analyze the combinations of an LH- or H- phrase tone with the boundary tone L% as variations of the same tonal complex, referred to here as (L)H-L%. The difference resides in the domain of the pitch rise, as indicated by the dashed line in Figure 3.7. In the LH-L% variant, the main domain for the pitch rise is the penultimate syllable. In the H-L% variant, on the other hand, the rise in pitch may extend over several syllables.

Consider first the LH-L% variant, which is exemplified by the pitch contour of example (2) in Figure 3.9. Pitch begins around the middle of the speaker's current range and remains near flat with a slight downtrend over the first 7 syllables of the IU. Starting at the beginning of the penultimate syllable (*.du.*), pitch rises 10 st to the high target of the LH- phrase tone at the beginning of the vowel of the last syllable (*.na#*). Pitch then drops 12 st to the low target of the L% of the IU boundary tone.

(2) *tauna dɛi anu baduna*
tau-0=na dɛi anu badu=na
put-UV=3S.GEN LOC FILL shirt=3S.GEN
'it is fixed at his whatchamacallit shirt' (pearstory_12_RSTM.055) ▷

The H-L% boundary-tone complex is illustrated by (3), for which the periogram is shown in Figure 3.10. The rise to the high target does not occur on the penultimate syllable exclusively, but instead happens gradually. Hence, the contour is labeled H-L%.

The pitch contour in Figure 3.10 shows that pitch begins mid-level of the speaker's current range and gradually rises about 4 st over the initial 6 syllables of the IU. The pitch rise reaches the high target of the H- phrase tone at the

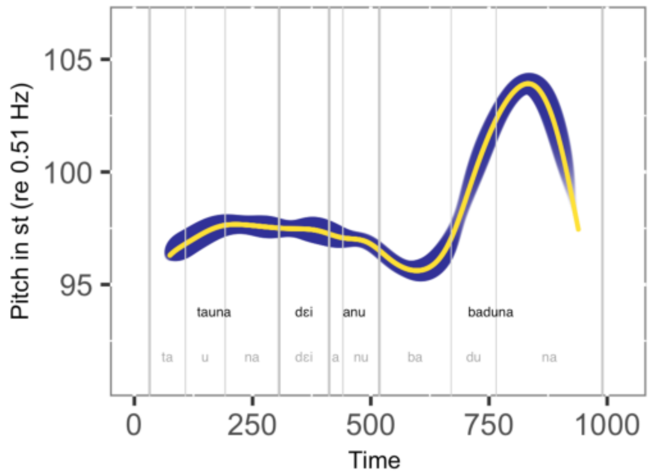

Figure 3.9: Periogram with pitch track (in st) for example (2), with final boundary-tone complex LH-L%, speaker RSTM

beginning of the vowel of the ultimate syllable (*.kɔ#*) and then drops about 15 st to the low target of the L% boundary tone.

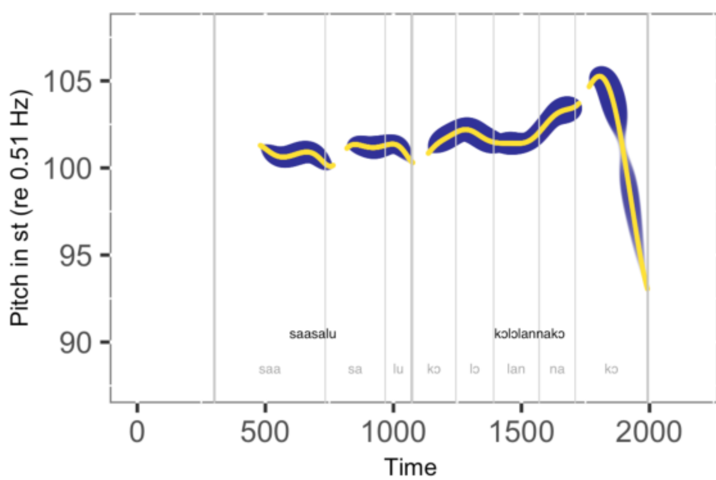

Figure 3.10: Periogram with pitch track (in st) for example (3), with final boundary-tone complex H-L%, speaker SP

(3) *saasalu kɔlɔannakɔ*
RDP-salu kɔlɔan=na=kɔ
RDP-to.face right=3S.GEN=AND
'it is facing to the right side' (spacegames_sequence4_KSR-SP.035) ▷

3.2.1.3 The LH-H% boundary-tone complex

An instance of the LH-H% boundary-tone complex is found in example (4), for which the periogram shown in Figure 3.11. Speaker ZBR enumerates several cultural events and festivities. Similar to example (3), the domain for the pitch rise is the penultimate syllable. On the last syllable, pitch remains (near) high. Similar to the L-H% above, the pitch contour shows a slight dip towards the end of the IU. This even visible in the summarizing spaghetti plot in Figure 3.6 and can be interpreted as an anticipation of the following low tone.

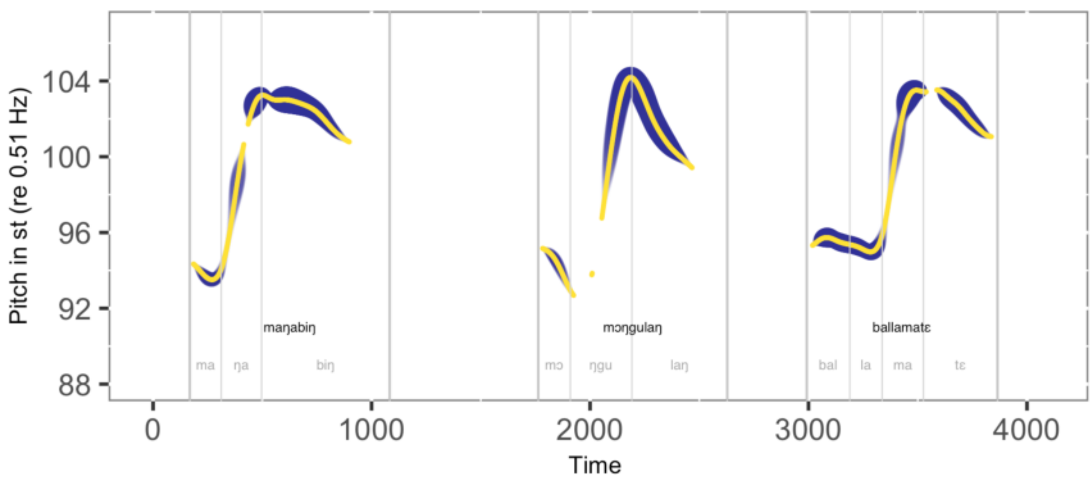

Figure 3.11: Periogram with pitch track (in st) for example (4), with final boundary-tone complexes LH-H%, speaker ZBR

(4) a. *maŋabing*
mɔN-kabing
ᴀᴠ-marry
'to marry'

b. *mɔŋgulaŋ*
mɔN-gulaŋ
ᴀᴠ-to.cradle
'to cradle'

c. *ballamatɛ*
ballamatɛ
funeral.ceremony
'the funeral ceremony'

(explanation-wedding-tradition_ZBR.265-267) ▹

While the LH-H% boundary-tone complex and the L-H% boundary-tone complex share some similarities, the main difference lies in the pitch rise domain. The two patterns can be easily differentiated from each other.

3.2.1.4 Distribution

The three boundary-tone complexes are the main tonal events occurring at the right-edge boundary of CIUs, i.e. singleton IUs and final IUs of CIUs. In addition to final boundary-tone complexes, there are rarely occurring discourse particles which attach to one of the boundary-tone complexes. They are not included in Figure 3.12 but are described in §3.2.4.

The frequency distribution of the different final boundary-tone complexes over the different data types is shown in Figure 3.12. Clearly evident is the fact that the (L)H-L% and the L-H% boundary-tone complexes are the two major patterns. The LH-H% pattern is a minor pattern but is very distinct in its function, as I will show in the next section (§3.2.1.5).

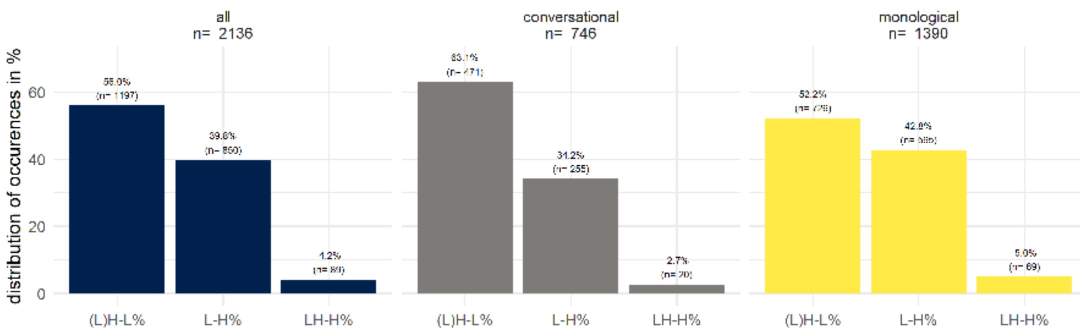

Figure 3.12: Frequency distribution of tonal events at the right edge of substantive CIUs within conversational and monological recordings, numbers are rounded to one decimal place.

The frequency distribution within the two data types is similar, with both types attesting more (L)H-L% boundary tones than L-H%. The distribution is more heavily skewed in conversational data, in which 60.1% of CIUs occur with the (L)H-L% boundary tone and 29.3% with the L-H%. This trend is less pronounced in the monological data, where there is a difference of only 10 percentage points. Furthermore, monological data show more final LH-H%.

3.2.1.5 Function

The difference in distribution can be explained by the different functions of CIU-final boundary-tone complexes. These are summarized in Table 3.3.

Table 3.3: Summary of functions of CIU-final boundary-tone complexes

CIU-final boundary-tone complex	function
L-H%	continuation
(L)H-L%	finality
LH-H%	non-final elements of lists

Monologues yield a higher proportion of CIUs with the LH-H% pattern because, in descriptive texts in the corpus, speakers frequently enumerate referents, events or entities; see example (4) above. In lists, especially for non-final elements of lists, the LH-H% is the preferred intonation pattern. However, the pattern is not exclusively reserved for such CIUs but can also occur in other environments where the speaker wants to express a high degree of continuation. Similarly, (non-final) elements of lists can also be uttered with the L-H% when signaling less strong continuation.

The distributional differences between the boundary-tone complexes observed in Figure 3.12 support the proposed functions in the following way: The higher proportion of CIUs with final (L)H-L% in conversations as compared to monologues reflects the fact that in conversations a paragraph may consist of a question and an answer only, while in monologues narrations are organized into longer paragraphs, containing several CIUs.

The function of (L)H-L% as signaling finality and L-H% as signaling continuation is nicely illustrated by *Tail-Head Linkage* (THL) constructions in narrations. de Vries (2005: 262) defines THL as

> […] a way to connect clause chains in which the last clause of a chain is partially or completely repeated in the first clause of the next chain.

In Totoli, instances of THL occur mostly in unplanned narratives and contribute to what de Vries (2005: 378) calls *processing ease*, as it links paragraphs and maintains event coherence. At the same time, they serve as a planning device, allowing speakers more time to plan the next paragraph. In the corpus, recordings of retellings of the Pear Story yield a considerable number of instances of THL constructions, as speakers are given the task to extemporize a coherent account of the story-line of a previously unknown story. In the Pear Story retellings, CIUs are usually grouped into higher-level units above the CIU which may be termed "paragraphs" (see Himmelmann & Ladd 2008: 251). A paragraph consists of a series of CIUs ending on L-H% and concludes with a final CIU marked by the (L)H-L% pattern. In a THL construction, the final CIU – the Tail of a paragraph – is repeated in full or in part as Head of the subsequent paragraph. The excerpt of a Pear Film retelling in example (5a)–(5n) provides an illustration.

(5) a. *bali tau pagauan* L-H%
‘So a gardener …’

b. *<na> kɔnɔnipu alpukaatna* L-H%
‘… is picking the avocados,’

c. *nipɛnɛkanna dɛi bataŋna* L-H%
‘he is climbing up the trunk,’

d. *niambinnamai uliai babɔ* L-H%
‘he is getting (the avocados down) from the top,’

e. *sagaat nadabumai dɛi buta* L-H%
‘half (of the avocados) fell to the ground,’

f. *bai indʒan nakaalamai* L-H%
‘and then after (he) picked (them) up,’

g. *ninauna pɔniai <mɔi> nitauna dɛi karandʒaŋ* L-H%
‘he brought them there and put them in the basket,’

h. *dɛi llɛŋgɛt* L-H%
‘in the hamper,’

i. *kaddaan tau* L-H%
‘(then) there was a person,’

j. *nɔtumalibkɔ* L-H%
‘(he) passed by,’

k. *biibindas tɔɔlaŋ* **H-L%**
‘(and he) pulled a goat.’

l. *biibindas tɔɔlaŋ* **L-H%**
‘(Though the person) pulled a goat,’

m. *tapi ganɛga tumalibkɔ* H-L%
‘he only passed by.’

n. *iŋga daan noosa kaddaanmai maŋŋana saasapɛda* L-H%
‘Not long after that, there came a child, cycling.’

(pearstory_11_SP.001-014) ▷

Examples (5a)–(5k) form one paragraph and are each uttered with the L-H% boundary-tone complex, marking non-finality. The final CIU of the paragraph – the Tail of the paragraph – is (5k) *biibindas tɔɔlaŋ* 'pulling a goat', which is uttered with the H-L% boundary-tone complex. It is repeated as the Head of the subsequent paragraph, this time bearing the L-H% boundary-tone complex. The realization of the two CIUs of the THL construction in (5k)–(5l) above are displayed in Figure 3.13.

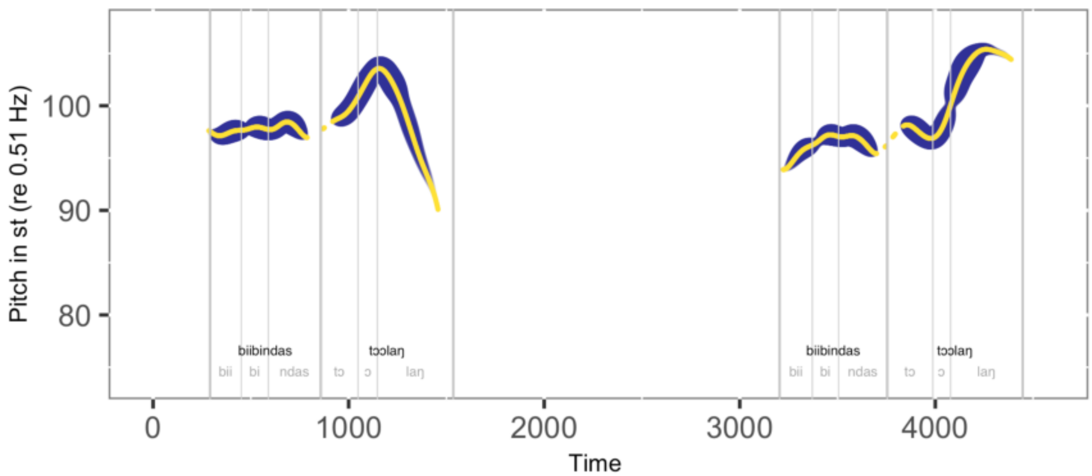

Figure 3.13: Periogram with pitch track (in st) for example (5k)–(5l), realization of the two segmentally identical CIUs of the THL construction; speaker SP

These examples support the hypothesis that (L)H-L% signals finality, while L-H% serves as "continuer", signaling non-finality of a CIU with regard to a higher-level discourse unit. Note that both contours occur in interrogative as well as declarative sentences.

3.2.2 Tonal events at the boundaries of non-final, embedded IUs of CIUs

This section deals with tonal events at the right-edge boundaries of non-final IUs of CIUs. This is exemplified in Figure 3.14 below.

Tonal events at the right-edge boundary of non-final, embedded IUs of CIUs can be equally classified into three different tonal contours. The summarizing plot of the three boundary-tone complexes is shown in Figure 3.15.

The final boundary-tone complex consists of a phrase tone (T-) and a boundary tone (T%): the phrase tone is a low tone L-, a high tone H- or a rising tone

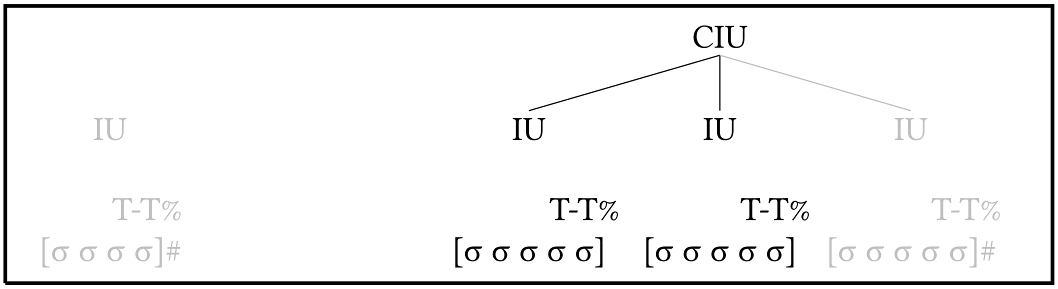

Figure 3.14: Visualization of non-final IUs of CIUs being discussed in this section

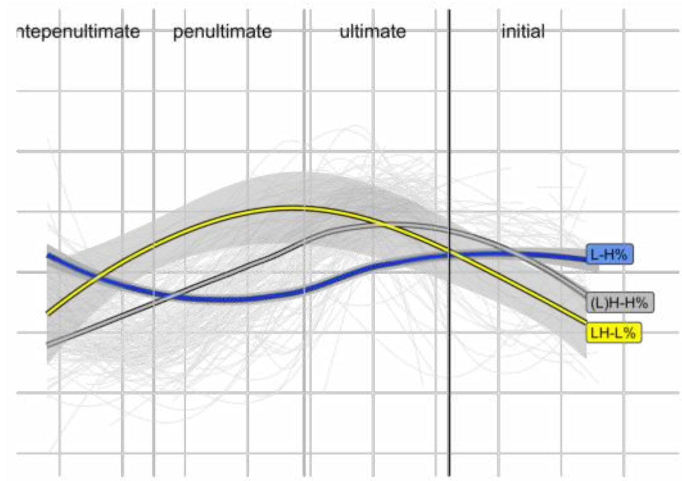

Figure 3.15: Spaghetti plot showing the intonation contours of the final three syllables of each of the three final boundary-tone complexes of embedded and non-final IUs with the items superimposed on each other and with an average contour produced by *Loess* smoothing in R (Team 2017): vertical bars indicate syllable boundaries; only final $CV.CV.CV]_{IU}[CV$... syllable structure are displayed; values are z-transformed for IU.

LH-, while the boundary tone can be either a low tone L% or a high tone H%. Figure 3.16 depicts a schematic version of the three final boundary-tone complexes of non-final, embedded IUs of CIUs. In the following, I briefly summarize each boundary-tone complex. Subsequently, I illustrate them with examples and then describe their distribution in the corpus.

The three final boundary-tone complexes of embedded IUs of CIUs are illustrated with examples from the corpus. For the purpose of exemplification, only CIUs consisting of two embedded IUs are used. Hence, they only have a single CIU-internal tonal event, i.e., that of the non-final, embedded IU.

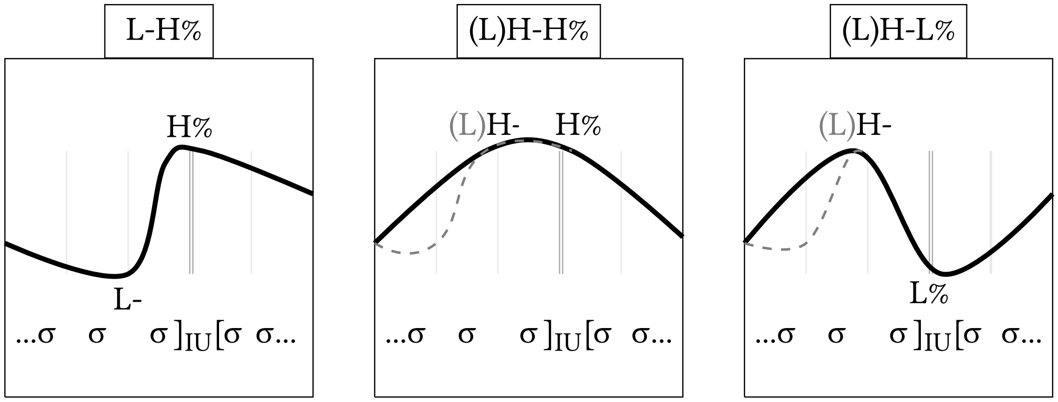

Figure 3.16: Schematic representation of final boundary-tone complexes of embedded IUs: vertical bars indicate syllable boundaries

3.2.2.1 The L-H% boundary-tone complex

An instance of the final boundary-tone complex L-H% is found in example (6), for which the periogram is shown in Figure 3.17.

The initial word of the CIU, *sapɛda* 'bicycle', forms its own embedded IU, demarcated by the final boundary-tone complex L-H%. Pitch starts around the middle of the speaker's current range and it reaches the low target of the L- phrase tone located at the boundary between the penultimate (*.pɛ.*) and the ultimate syllable (*.da*) of the IU. On the ultimate syllable, pitch rises 6 st to the high target of the H% boundary tone.

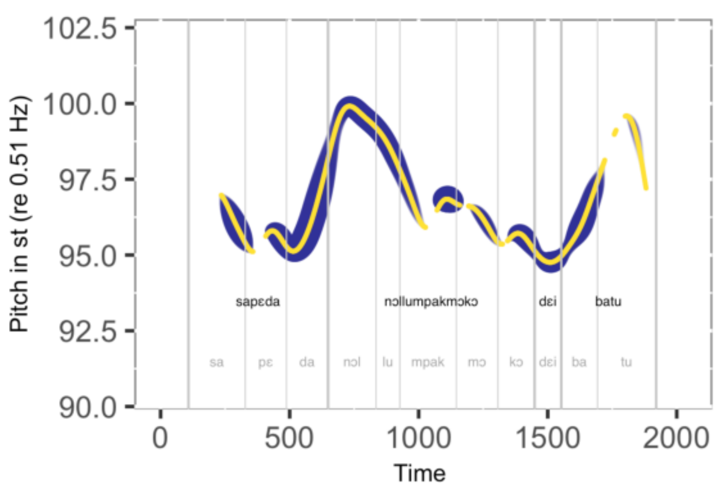

Figure 3.17: Periogram with pitch track (in st) for example (6), example of final boundary-tone complex L-H% on embedded IU *sapɛda*, speaker SP

(6) *sapɛda | nɔllumpakmɔkɔ dɛi batu*
sapɛda nɔ-RDP-lumpak=mɔ=kɔ dɛi batu
bicycle ST.RLS-RDP-hit.against=CPL=AND LOC stone
'the bicycle hit the stone' (pearstory_14_SP.028) ▹

3.2.2.2 The (L)H-H% boundary-tone complex

The (L)H-H% is a summarizing label designating the two variants LH-H% and H-H%. The difference is that in the LH-H% variant the main domain for the pitch rise is the penultimate syllable. In the H-H% variant, on the other hand, the rise in pitch may extend over several syllables. The two variants are illustrated below.

An instance of the LH-H% boundary-tone complex is found in example (7), for which the periogram is given in Figure 3.18. The initial 5 words form a separate IU, the final word of which is *kakaita* 'your grandfather'.

Pitch starts around the middle of the speaker's current range and gradually drops about 3 st over the first 8 syllables. On the prefinal syllable (*.kai.*) of the final word *kakaita* 'your grandfather' of the first IU, pitch rises 5 st towards the high target of the LH- phrase tone at the beginning of the final syllable (*.ta*). Pitch then remains high over the final syllable of the IU as the IU ends on an H% boundary tone.

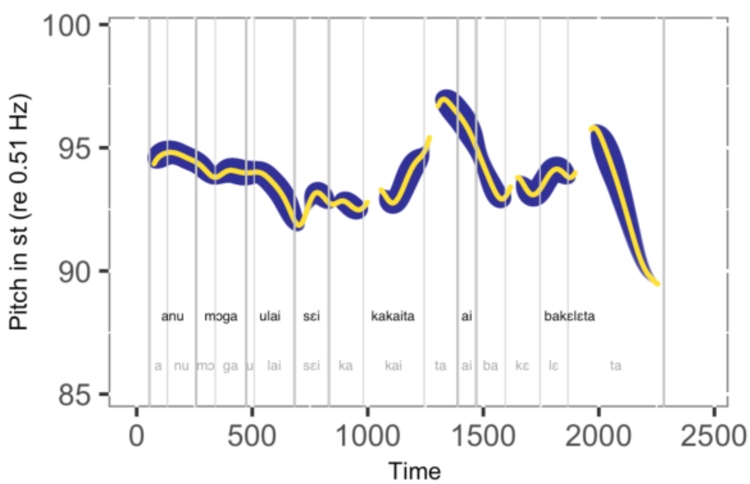

Figure 3.18: Periogram with Pitch track (in st) for example (7), example of LH-H% boundary-tone complex on *kakaita*, speaker SYNO

(7) *anu mɔga ulai sɛi kakaita | ai bakɛlɛta*

anu	mɔga	uli=ai	sɛi	kakai=ta	ai
REL	only	from=VEN	HON	grandfather=1PI.GEN	and

bakɛlɛ=ta
grandmother=1PI.GEN

'which (only) comes from our grandfathers and grandmothers'

(explanation-lelegesan_SYNO.002) ▷

The H-H% variant of the boundary-tone complex is exemplified in (8), with its corresponding periogram depicted in Figure 3.19. The first three words form a separate IU, the final word of which is *kami* '1pe'. Here, the domain of the pitch rise to the high target of the H% boundary tone is not the prefinal syllable (*ka.*) exclusively, but extends over several syllables (*mɔ.i.ta.ka.*).

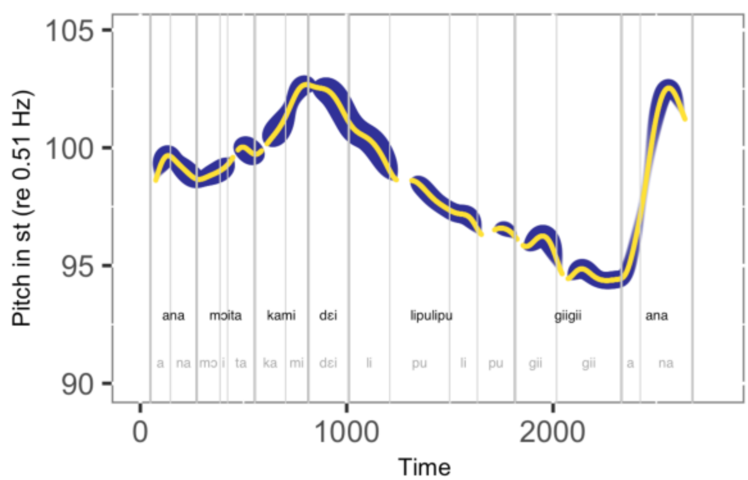

Figure 3.19: Periogram with pitch track (in st) for example (8), example of H-H% boundary-tone complex on *kami*, speaker ZBR

(8) *ana mɔita kami | dɛi lipulipu giigii ana*

ana	mɔ-ita	kami	dɛi	RDP-lipu	RDP-gii	ana
if	POT-see	1PE	LOC	RDP-country	RDP-different	MED

'when we look at that in other countries'

(explanation-wedding-tradition_ZBR.258) ▷

3.2.2.3 The (L)H-L% boundary-tone complex

Lastly, the (L)H-L% boundary-tone complex and its two realizations H-L% and LH-L% are exemplified. Similar to the above, in the LH-L% variant the main domain for the pitch rise is the penultimate syllable. In the H-L% variant, on the other hand, the rise in pitch extends over several syllables. The two variants are illustrated below.

An instance of the final boundary-tone complex H-L% is displayed in the periogram of example (9) in Figure 3.20. The initial word *moanɛ* 'man' forms a separate IU, demarcated by the IU-final boundary-tone complex H-L%.

Pitch starts around the middle of the speaker's current range. Pitch then rises 10 st over the first two syllables until it reaches the high target of the H- phrase tone located between the penultimate (*.a.*) and the ultimate syllable (*.nɛ*) of the IU. On the ultimate syllable, pitch drops 7 st to the low target of the IU-final L% boundary tone.

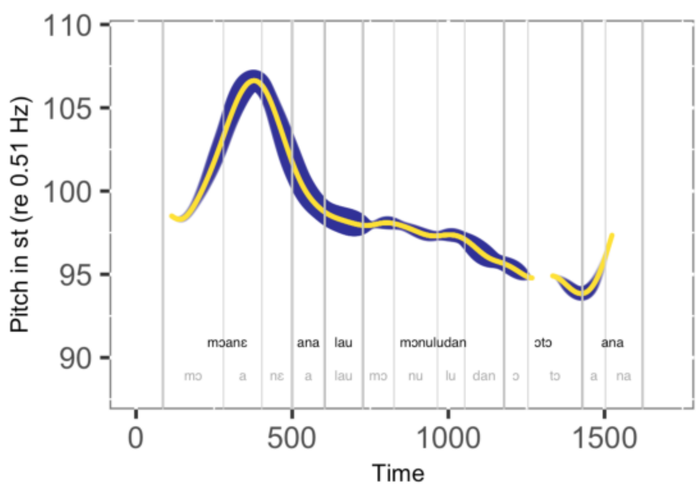

Figure 3.20: Periogram with pitch track (in st) for example (9), example of final boundary-tone complex H-L% on *mɔanɛ*, speaker SP

(9) *mɔanɛ | ana lau mɔnuludan ɔtɔ ana*
mɔanɛ ana lau mɔN-sulud-an ɔtɔ ana
man MED presently AV-push-APPL car MED
'it is a man who is pushing that car' (QUIS-focus_SP.010) ▷

Example (9) is particularly interesting. In this example, the word *mɔanɛ* 'man' is informationally important and therefore constructed as the first element of an cleft construction. The construction marks the focus on *mɔanɛ* 'man', which is then followed by *ana lau mɔnuludan ɔtɔ ana* 'this (one) is pushing the car'. The pitch contour on *mɔanɛ* 'man' could potentially be interpreted as a prominence-lending pitch movement on a word that is in focus. However, in Chapter 2, and in particular in §2.2, I showed that Totoli does not make use of such means to mark focus. Focus is expressed by means of a cleft construction in this case. The focus constituent is then prosodically uttered in its own IU with a H-L% boundary tone that signals finality. The prosodic realization of syntactic constructions is further discussed in §3.3.

Finally, the LH-L% variant is illustrated by example (10) and its periogram in Figure 3.21. The first three words form a separate IU, the last word of which is *pɔni* 'again'.

Pitch begins around the middle and initially drops about 4 st. On the prefinal syllable (*pɔ.*), pitch rises 9 st towards the high target of the H- phrase tone, located at the beginning of the penultimate syllable of the IU-final word *pɔni*. On the final syllable, pitch drops about 7 st.

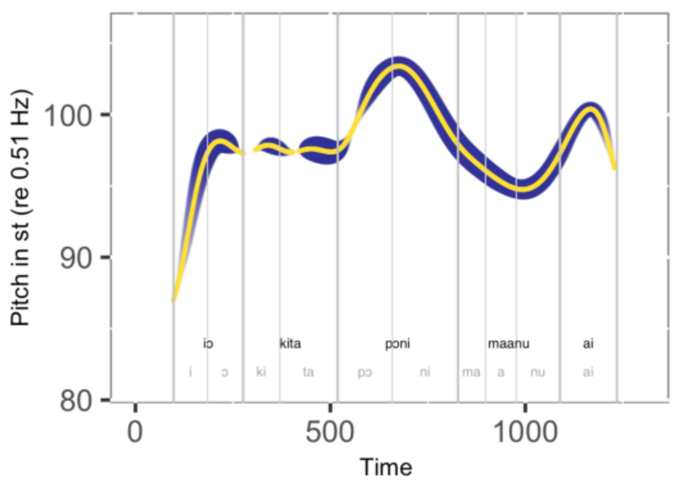

Figure 3.21: Periogram with pitch track (in st) for example (10), LH-L% boundary-tone complex on *pɔni*, speaker SP

(10) *io kita pɔni | maanuai*
io kita pɔni mɔ-anu=ai
OK 2S again AV-FILL=VEN
'ok, now you do one again' (spacegames_sequence1_KSR-SP.252) ▹

3.2.2.4 Distribution

The distribution of the different final boundary-tone complexes of non-final, embedded IUs is shown in Figure 3.22.

The distribution shows that the (L)H-H% and the L-H% are the major final tonal patterns of non-final, embedded IUs of CIUs. The LH-H% pattern is only marginally attested. The distribution is similar over the different data types.

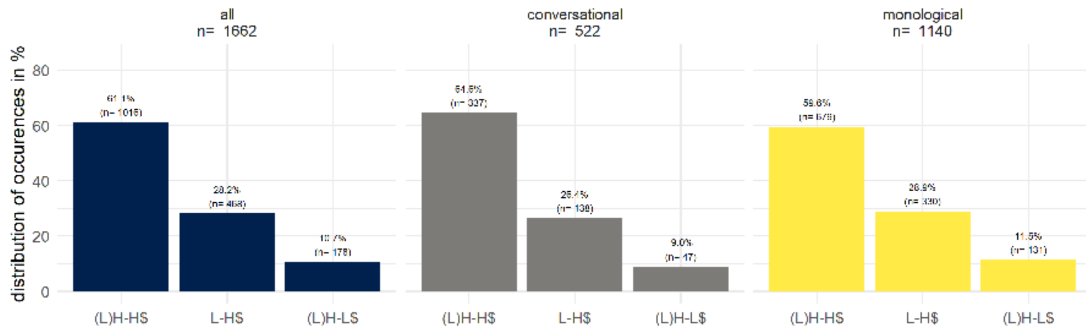

Figure 3.22: Frequency distribution of tonal events at the right-edge boundary of embedded, non-final IUs, within conversational and monological recordings

3.2.3 Phonetic variability

The segmental material influences the realization of the pitch contours in that a lack of sonorant material in the final and prefinal syllable results in the pitch rises or falls to be only partly realized.

For the purpose of illustration, I use examples of singleton IUs or final IUs of CIUs with final boundary tones (L)H-L% and L-H% respectively, as they are the major final boundary tones. The effects are the same for the final LH-H% boundary-tone complex and also for final boundary-tone complexes of embedded, non-final IUs of CIUs.

Consider Figure 3.23 and Figure 3.24, which are the periograms of another example of a THL given in example (11). As is specified for all THL constructions, the first CIU bears the final (L)H-L% boundary-tone complex and the second one the L-H%. These are only partly realized, due to the voiceless plosives [k] and [t] in the onset of the final (*ki.*) and the prefinal syllable (*.ta#*).

In Figure 3.23, the rise in pitch on the penultimate syllable of the CIU (*ki.*) is interrupted and is only visible on the short vowel of the penultimate syllable. The drop in pitch of about 11 st of towards the low boundary tone L% is realized in full on the final syllable. In Figure 3.24, the rise to the high target of the H% boundary tone is interrupted and hence results in a jump in pitch of about 6 st on *ki.ta#*.

(11) a. *anu ampi kɔlɔanan | saasaluai kita*
anu ampi kɔlɔanan RDP-salu=ai kita
REL side right RDP-to.face=VEN 2S
'So the (one) on the right-hand side is facing you'

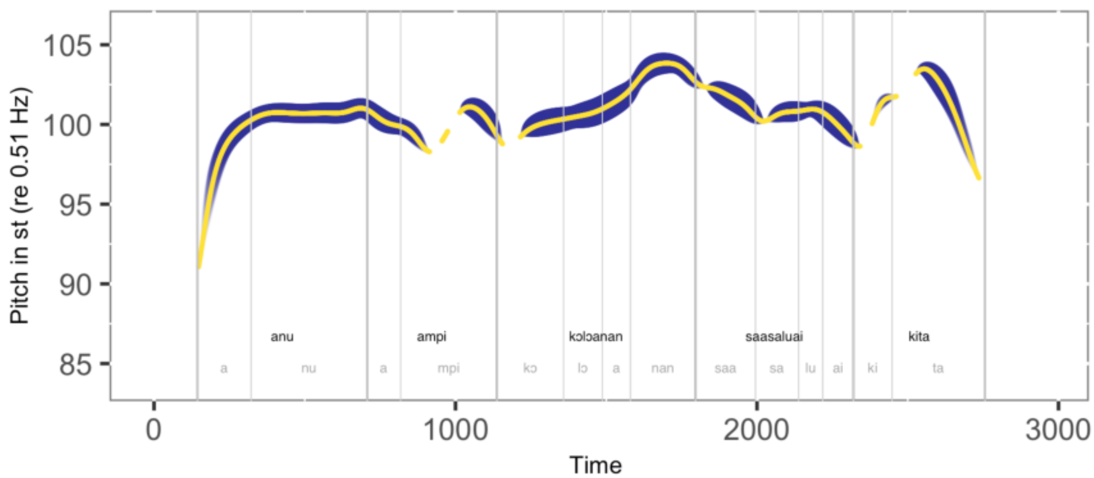

Figure 3.23: Periogram with pitch track (in st) for example (11a) with CIU-final boundary-tone complex LH-L%, speaker SP

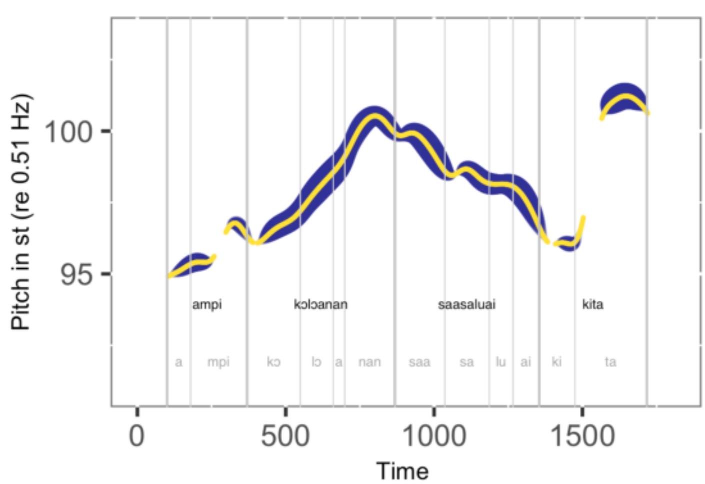

Figure 3.24: Periogram with pitch track (in st) for example (11b) with CIU-final boundary-tone complex L-H%, speaker SP

b. *ampi kɔlɔanan | saasaluai kita*
ampi kɔlɔanan RDP-salu=ai kita
side right RDP-to.face=VEN 2S
'the (one) on the right-hand side is facing you'
(spacegames_sequence4_KSR-SP.231 & 233) ▷

If the segments of the final and the prefinal syllable are fully sonorant, the boundary-tone complexes are fully realized, as can be seen in the two realizations of example (12) in Figure 3.25 and Figure 3.26.

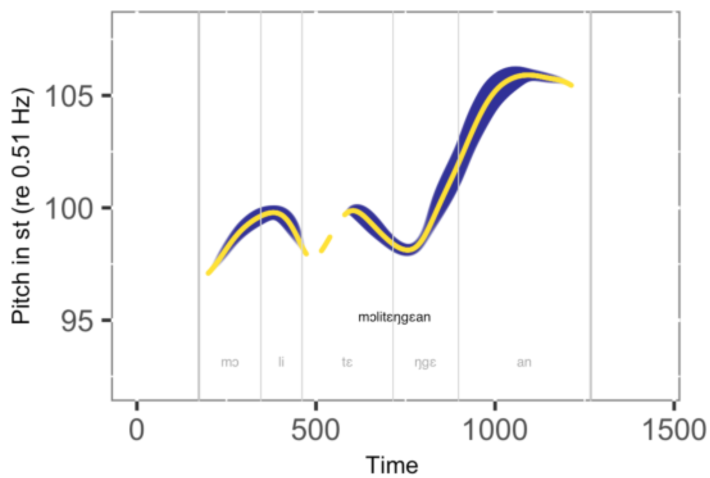

Figure 3.25: Periogram with pitch track (in st) for example (12) with IU-final boundary-tone complex L-H%, speaker SP

(12) *mɔlitɛŋgɛan*
mɔli-tɛŋgɛ-an
RCP-back-RCP
'(they are) back to back' (spacegames_sequence4_KSR-SP.071 & 105) ▷, ▷

For final syllables with a short vowel, the main domain of the rise or fall to the T- phrase tone is the penultimate syllable. Segmental material affects the shape of the tonal contours but not the location of tonal targets. However, if the IU-final syllable involves a long vowel, the tonal targets of both the phrase tone and the boundary tone are realized in full on that syllable.

For an illustration of the LH-L% and the L-H% on a final syllable with a long vowel, see examples (13) and (14), which both end on the word *pɔmɔɔ* 'back then/-first'. Two different realizations are given in Figure 3.27 and Figure 3.28.

The pitch contour of example (13) in Figure 3.27 shows that the rise to the high target of the LH- phrase tone and the subsequent fall to the low target of the L% boundary tone are both realized on the ultimate syllable (*.mɔɔ*).

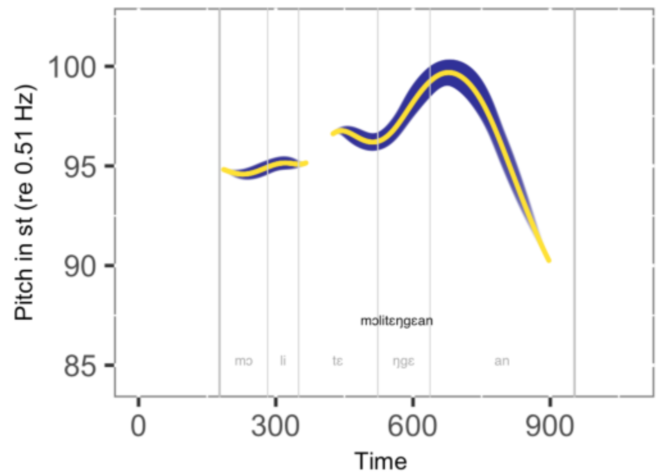

Figure 3.26: Periogram with pitch track (in st) for example (12) with IU-final boundary-tone complex LH-L%, speaker SP

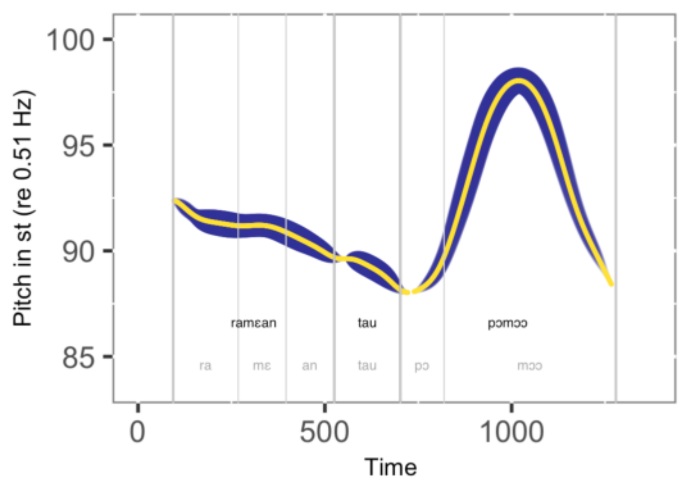

Figure 3.27: Periogram with pitch track (in st) for example (13) with IU-final boundary-tone complex LH-L% on a final syllable with a long vowel, speaker SYNO

(13) *ramɛan tau pɔmɔɔ*
ramɛ-an tau pɔmɔɔ
lively-NMLZ person back.then
'The crowd/amusement of the people back then'

(explanation-lelegesan_SYNO.085) ▹

Similarly, the realization of example (14) in Figure 3.28 shows that the drop in pitch towards the low target of the L- phrase tone and the subsequent rise towards the high target of the H% boundary tone are also both realized on the final long syllable (*.mɔɔ*).

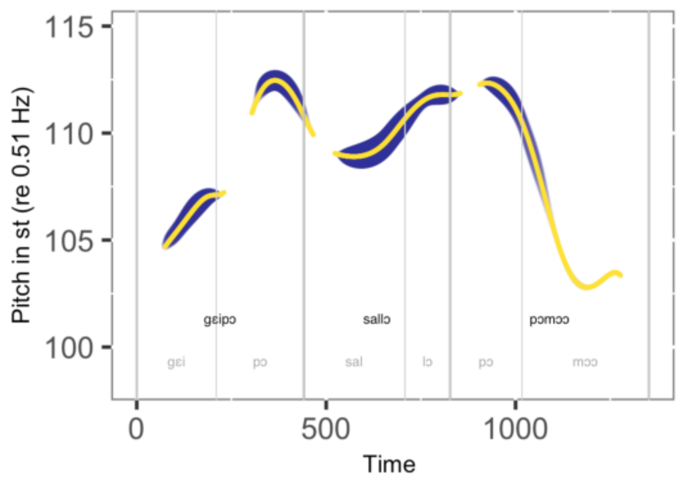

Figure 3.28: Periogram with pitch track (in st) for example (14) with CIU-final boundary-tone complex L-H% on a final syllable with a long vowel, speaker RD

(14) *geipɔ | sallɔ pɔmɔɔ*
geip=pɔ sallɔ pɔmɔɔ
neg=INCPL basket first
'no, but the basket first' (pearstory_13_RD.015) ▹

For the purpose of exemplification, I discussed the different boundary-tone complexes on rather short CIUs above. However, many CIUs contain several embedded IUs. Therefore, I discuss two examples of such rather long CIUs. Example (15) consists of a long CIU that begins with an embedded IU spanning 5 words and is demarcated by the final boundary-tone complex L-H% realized on the final word *itu* 'DIST'. The rise in pitch of about 5 st to the high target of the H%

boundary tone is realized merely as a jump, due to voiceless plosive [t] in the onset of the final syllable (*.tu*). Pitch then drops 8 st to the low target of the L-H% boundary-tone complex of the following IU containing the word *sapɛda* ‘bicycle’. Pitch then drops towards the low target of the IU-final boundary-tone complex L-H% of the final IU in the CIU. The final rise on the last syllable (*.mpak*) is again realized only as a jump in pitch due to the voiceless syllable onset.

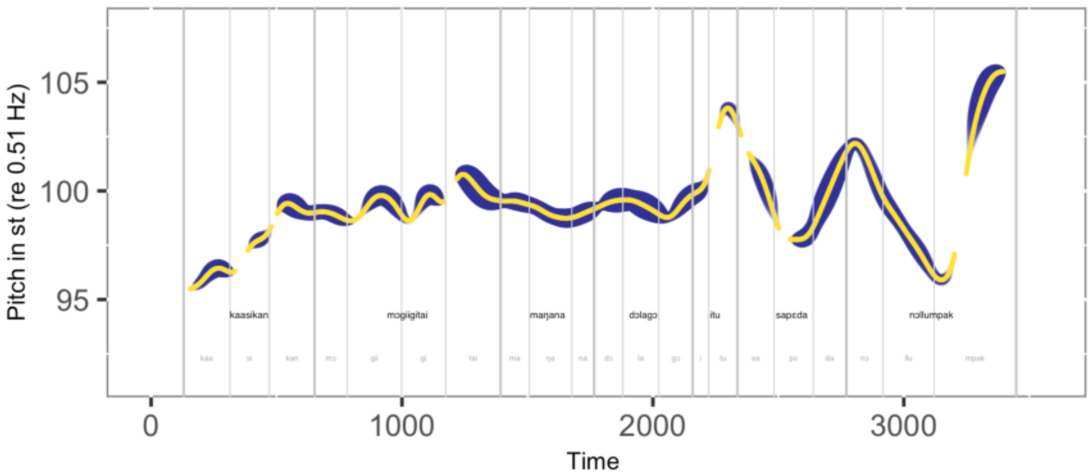

Figure 3.29: Periogram with pitch track (in st) for example (15) with CIU-final boundary-tone complex L-H%, speaker SP

(15) *kaasikan mɔgiigitai maŋana dɔlago itu | sapɛda | nɔollumpak*
kɛasikan mɔg-RDP-ita-i maŋana dɔlagɔ itu sapɛda
excitement AV.NRLS-RDP-watch-APPL child girl DIST bicycle
nɔ-RDP-lumpak
AV.RLS-RDP-hit.against
‘because of his excitement in looking at the girl, his bicycle crashed (against the stone)’ (pearstory_11_SP.025) ▹

Consider example (16) and its visualization in Figure 3.30. The CIU consists of several embedded IUs. The first two bear the final boundary-tone complexes LH-H%, realized on *siritana* ‘this story’ and *daan* ‘EXIST’. The word *maaling* ‘to get lost’ bears the IU-final LH-L% boundary-tone complex. The final IU of the CIU ends on an L-H% boundary-tone complex.

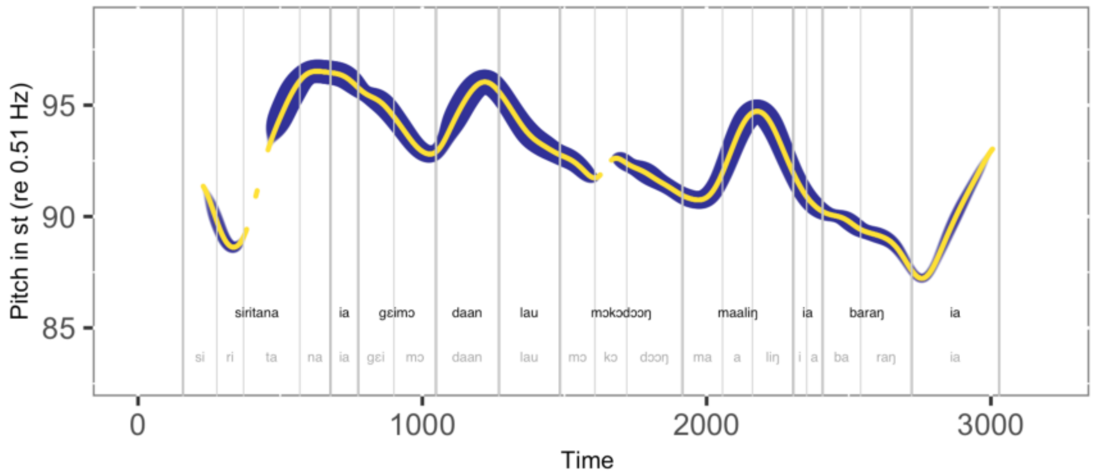

Figure 3.30: Periogram with pitch track (in st) for example (16) with IU-final boundary-tone complex L-H%, speaker SYNO

(16) *siritana | ia gɛimɔ daan | lau mɔkɔdɔɔng maaliŋ | ia baraŋ ia*

sirita=na	ia	gɛimɔ	daan	lau	mɔkɔ-dɔɔŋ	mo-aliŋ	ia
story=3S.GEN	PRX	not	EXIST	presently	ST.AV-want	ST-disappear	PRX

baraŋ	ia
goods	PRX

'This story will never again get lost; this thing'

(explanation-lelegesan_SYNO.007) ▷

The question is whether the tonal contours can be explained as involving IU-final H% boundary tones only, rather than the combination of a LH- or L- phrase tone with an H% boundary tone. Such an analysis would not capture the fact that the pitch rises occur on one syllable only. Analyzing the IU-final rise as an H% boundary tone alone would not explain why the pitch contour does not remain high after the initial high target towards the end of the first embedded IU and the high targets of the high boundary tones of the following IUs.

In considering example (16) and its visualization in Figure 3.30, it becomes evident that we must assume an L- or LH- phrase tone to account for the drop in pitch before the final rises on the ultimate syllables. The alternative would be to assume that embedded IUs begin with an IU-initial low tone. This, however, would not explain why pitch drops steadily over the entire IU towards the final or prefinal syllable.

To conclude the discussion of phrase-final pitch events, a note on discourse particles is due.

3.2.4 Discourse particles

In addition to one of the boundary-tone complexes, a discourse particle can optionally occur at the end of a singleton IU or CIU. The two prosodic clitics *wi* and *ɛɛ* are the most frequently attested in the corpus. Other discourse particles are not frequent enough to allow for any generalizations. The discourse particles *wi* and *ɛɛ* are uttered under a coherent pitch contour together with the host CIU – either singleton IU or a CIU – and no pause occurs between them. Impressionistically, most CIUs sound complete if the 'prosodic clitic' is cut off using any annotation software. These encliticized discourse markers are tonally specified as either rising or falling, independent of the boundary-tone complex of the host CIU. They are tonally specified for either H%, to signal continuation or L%, to signal finality.

A frequently occurring discourse particle in the corpus is *wi*. Similar to Indonesian *kan*, the Totoli discourse marker *wi* is used as "a request of verification or confirmation, or it may be a marker of conjoint knowledge" (Wouk 1998: 403). In the corpus, this discourse particle frequently occurs in recordings of the Space Game task (Levinson et al. 1992). In this task, two participants are each given an identical set of photos and must find matching photos in a memory game. As one participant has to identify the photo being described by the second participant without seeing the latter's stack of photos, the consultants frequently ask for verification or confirmation of whether the photo selected indeed matches the intended image. In these contexts, the discourse marker *wi* is used. It is tonally specified for H%.

Figure 3.31 shows the realization of example (17), with *wi* in CIU-final position.

(17) *dɛllɔ ɛŋgaɛŋgat | anu dɛi ulin | wi*
dɛllɔ RDP-ɛŋgat anu dɛi ulin wi
like RDP-lift.up REL LOC back INTJ
'like being lifted up, the one at the back, right?'

(spacegames_sequence3_KSR-SP.225) ▹

The L-H% boundary-tone complex is realized on the word *ulin* 'back'. After a high target on the last syllable *ɛŋgaɛŋgat* 'being lifted' of the first IU, pitch drops towards the low target of the L- phrase tone located at the boundary between the penultimate and ultimate syllables of the IU-final word *ulin* 'back'. On the final syllable, pitch rises 6 st towards the high target of the H% boundary-tone complex. The discourse marker *wi* occurs after the rise of the L-H% boundary-tone complex, extending it by another 15 st.

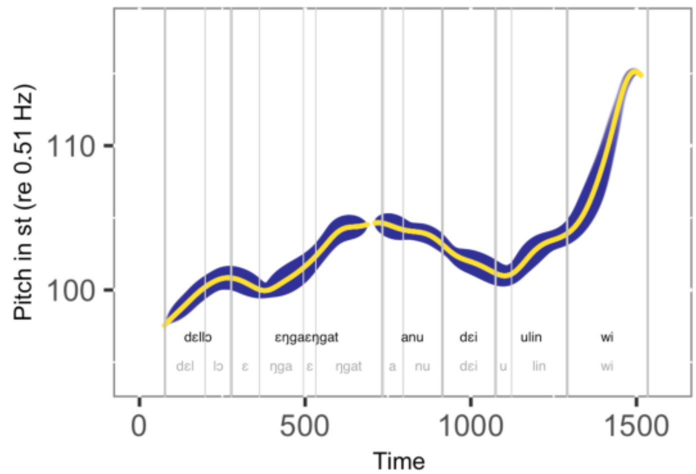

Figure 3.31: Periogram with pitch track (in st) for example (17) with CIU-final boundary-tone complex L-H%, followed by the discourse particle *wi* with H% tone, speaker SP

Taken from the same recording, example (18) is an instance of the same prosodic clitic realized after an IU-final LH-L% boundary-tone complex, as shown in Figure 3.32.

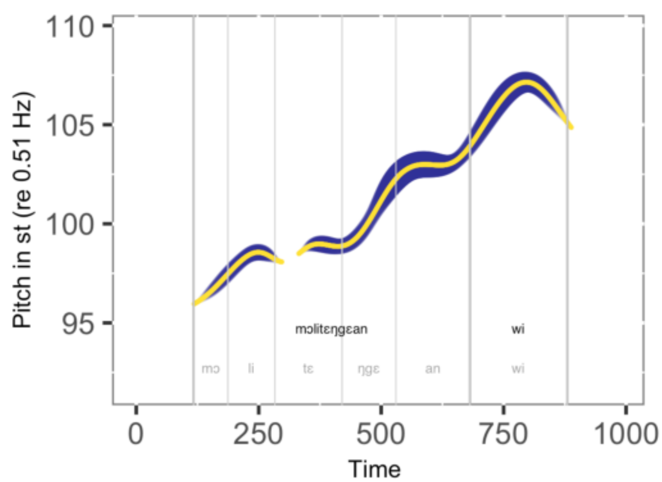

Figure 3.32: Periogram with pitch track (in st) for example (18) with final boundary-tone complex L-H%, followed by the discourse particle *wi* with H% tone, speaker SP

(18) *mɔlitɛŋgɛan | wi*
mɔli-tɛŋgɛ-an wi
RCP-back-RCP INTJ
'(they are) back to back, right?' (spacegames_sequence4_KSR-SP.124) ▹

Figure 3.32 shows the pitch rising 6 st towards the high target of the phrase tone, located at the beginning of the syllable *.an*. The subsequent final major drop in pitch is only partially realized, as the rise to the H% tone of the discourse particle extends into the coda of the preceding syllable.

Note the two examples with a slight dip on the final discourse particle *wi* in Figure 3.31 and even more pronounced so in Figure 3.32. This is parallel to the realization of the L-H% and the (L)H-H% pattern described above and appears to be a characteristic of the rising patterns in Totoli (see Dombrowski & Niebuhr 2010 for a discussion on convex vs. concave rising patterns).

Another frequently occurring final discourse marker is *ɛɛ*. The discourse marker is prosodically realized as a clitic, tonally specified for L%. It is used as an emphasizer, asserting the validity of the question or, as in example (19), reaffirming the correctness of the statement of the host CIU. Often it is a request for action. In example (19), the speaker urges the interlocutor to find the intended photo. Figure 3.33 shows the pitch contour of example (19).

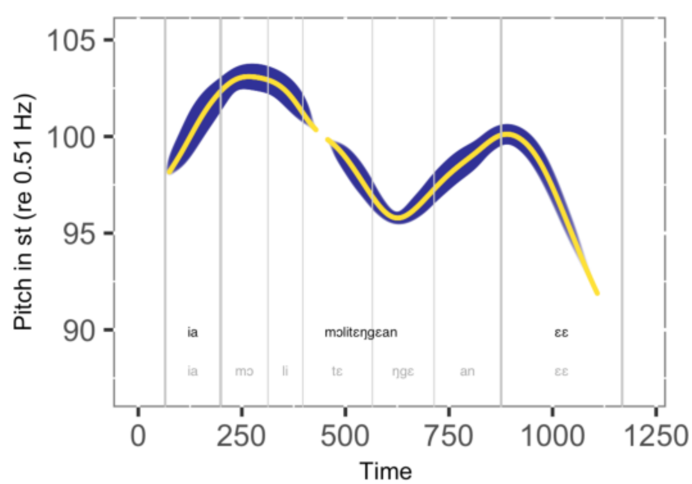

Figure 3.33: Periogram with pitch track (in st) for example (19) with final boundary-tone complex L-H%, followed by a discourse particle, speaker KSR

(19) *ia | mɔlitɛŋgɛan | ɛɛ*
ia mɔli-teŋgɛ-an ɛɛ
yes RCP-back-RCP INTJ
'yes, (they are) back to back!' (spacegames_sequence4_KSR-SP.278) ▷

The CIU ends on the L-H% boundary-tone complex followed by the L% boundary tone of the prosodic clitic. The boundary-tone complex is realized on the final two syllables of *mɔlitɛŋgɛan*, to which the prosodic clitic is added. On the vowel of the syllable *.ŋgɛ.*, the low target of the L- phrase tone is located somewhat earlier than in contexts without a prosodic clitic. Pitch then rises 5 st towards the high target of the H% boundary tone, reaching its peak at the boundary of the syllable *.an* and the following prosodic clitic. Pitch then drops 9 st towards the low target of the L% boundary tone of the prosodic clitic. The combination of an (L)H-L% boundary-tone complex followed by a prosodic clitic specified for L% is realized either as a sustained pitch plateau following the L% of the boundary-tone complex, or is integrated into the major final drop in pitch. This can be seen in Figure 3.34, which depicts the pitch contour of example (20). The major final fall to the L% of the H-L% boundary-tone complex is realized on the last syllable of *mɔliulunan* 'being in a row'. The prosodic clitic is then added, resulting in a further fall at the bottom of the speaker's range.

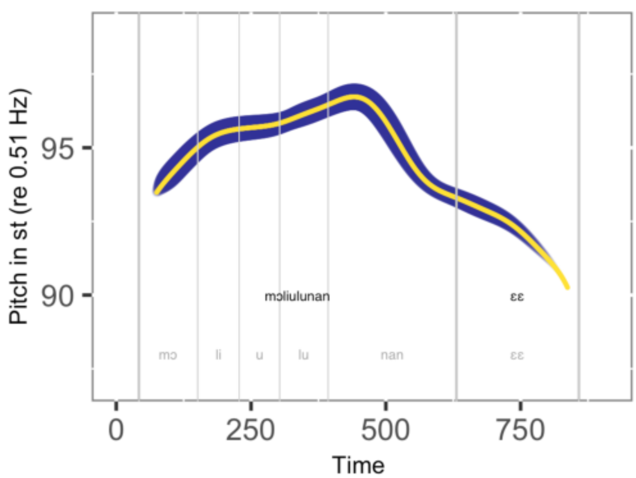

Figure 3.34: Periogram with pitch track (in st) for example (20) with final boundary-tone complex H-L%, followed by a discourse particle, speaker KSR

(20) *mɔliulunan | ɛɛ*
mɔli-ulun-an ɛɛ
RCP-ROW-RCP INTJ
'(they are) in a row!' (spacegames_sequence3_KSR-SP.017) ▹

If the preceding syllable ends on a vowel, the final prosodic clitic tends to be realized as part of the final major fall in pitch, as shown in Figure 3.35 from example (21).

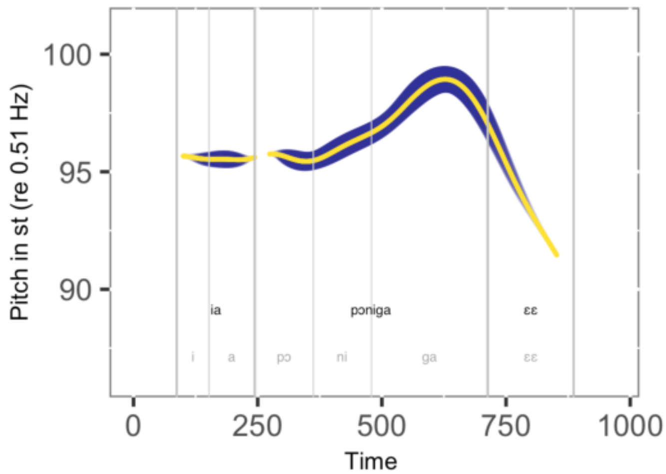

Figure 3.35: Periogram with pitch track (in st) for example (21) with final boundary-tone complex H-L% and a following discourse particle, speaker KSR

(21) *ia pɔniga | ɛɛ*
ia pɔni=ga ɛɛ
PRX still=? INTJ
'there is this one still!' (spacegames_sequence3_KSR-SP.136) ▷

3.2.5 Discussion

In this section, I developed a model of the intonation of Totoli (§3.2) on the basis of a corpus of (semi-)spontaneous speech (§3.2.2–§3.2.4) and informed by insights from the experiments described in Chapter 2.

In Totoli, a high proportion of singleton IUs are observed, where the final boundary-tone complex is the only pitch event. However, CIUs are also common. Table 3.36 shows the average number of IUs contained in a substantive CIU. The bins represent the number of IUs contained in a segmented CIU of the corpus, and the height of the bins represents their overall proportion. The proportion and absolute numbers are stated above the bins.

The distribution is skewed, and less than 10% of CIUs contain four or more embedded IUs. In the entire corpus, 43.4% of IUs are singletons, with 41.3% in monological data and 47.0% in conversational data. CIUs consisting of two embedded IUs occur at a proportion of 35.3% in the entire corpus, with 36.1% in conversational data and 34.8% in monological data.

In sections §3.2.1–§3.2.4, I analyzed pitch events at the right-edge boundaries of singleton IUs, non-final embedded IUs of CIUs, and final IUs of CIUs. I pro-

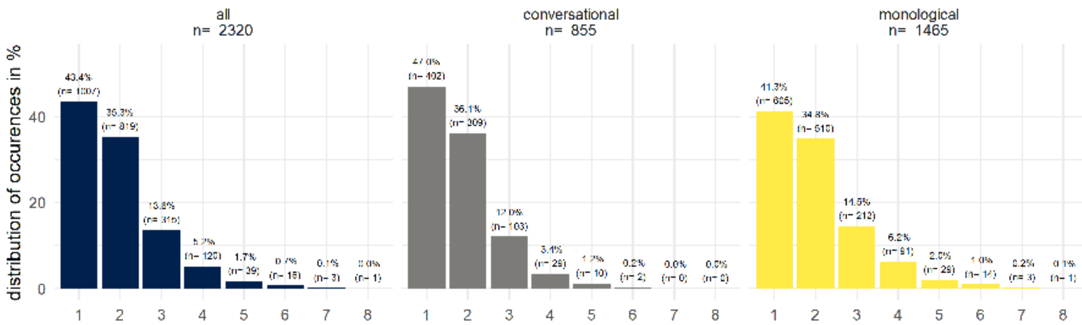

Figure 3.36: Frequency distribution of substantive CIUs containing n embedded IUs, 1 equals a singleton IU.

posed a classification of tonal events, which includes three boundary-tone complexes of each type. Table 3.4 provides a summary of the different tonal complexes.

Table 3.4: Summary of proposed IU-final boundary-tone complexes

IU-final boundary-tone complexes of non-final IUs of CIUs	IU-final boundary-tone complexes of singleton IUs or final IUs of CIUs.
(L)H-H%	LH-H%
L-H%	L-H%
(L)H-L%	(L)H-L%

Table 3.4 presents the different boundary-tone complexes, arranged such that those with similar tunes are in the same row. The only difference is between the (L)H-H% and the LH-H% boundary-tone complexes. I argue that in CIU-final position, the domain for the pitch rise is exclusively the penultimate syllable, expressed here by an LH- phrase tone. In final position of embedded, non-final IUs of CIUs, however, there is more variation with regard to the domain of the pitch rise, expressed here by the label (L)H-.

So far, the main difference between the final boundary-tone complex of embedded, non-final IUs of CIUs and singleton IUs or final IUs of CIUs is that, by definition, the latter co-occurs with other boundary phenomena which do not occur at the end of embedded, non-final IUs of CIUs or may not be as pronounced. For instance, this could be a reset in pitch and an interruption in the pitch contour, a following pause, final syllable lengthening, and vocal fry (cf. Schuetze-Coburn 1994: 93–155, Himmelmann 2006: 260–270, Du Bois et al. 1992, 1993, and Cruttenden 1997: 29–39). In the exposition above, I focused mainly on a discussion of pitch events. Further investigation of boundary strength may reveal interesting insights into the interplay of boundary type and boundary phenomena in Totoli (Schwiertz 2009, Cho 2005, Fougeron & Keating 1997).

Regarding the tonal patterns, the main difference pertains to the distribution of the types of boundary-tone complexes. While the (L)H-H % pattern is the main pattern at the right-edge boundary of non-final IUs of CIUs, it is the minor pattern in CIU-final position, i.e. of singleton IUs and final IUs of CIUs. Conversely, (L)H-L % is the main pattern occurring at the right-edge of CIUs but the minor pattern occurring at the right-edge of embedded, non-final IUs of CIUs. In both positions, the L-H % boundary-tone complexes occur in comparable proportions. This is displayed in Figure 3.37.

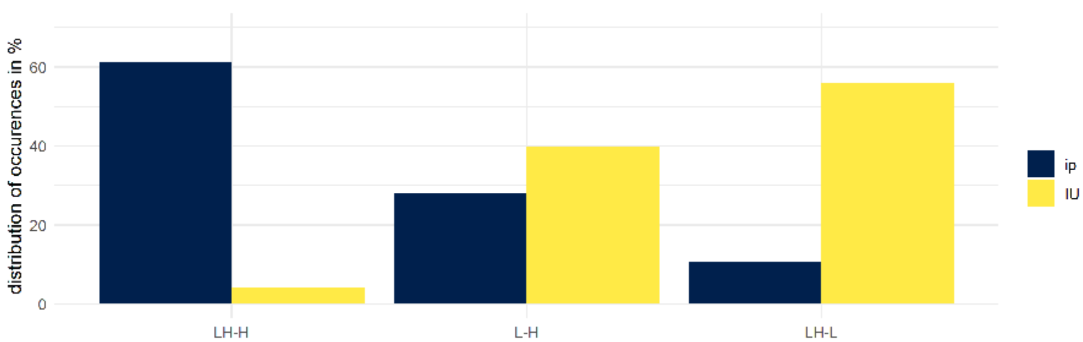

Figure 3.37: Frequency distribution of tonal events at the right-edge boundary of embedded IUs and non-final IUs of CIUs (e. IUs) in blue and singleton IUs or final IUs of CIUs in yellow

Considering the similarities between tonal events demarcating the right-edge boundaries of CIUs and non-final IUs of CIUs, the question arises as to how one can explain the differences in distribution.

In section §3.2.1.5 I described the different patterns as expressing varying degrees of finality or continuation. The LH-H% pattern expresses a high degree of non-finality or continuation, the L-H% pattern expresses a medium degree of continuation or non-finality and the LH-L% pattern expresses finality. Taking this as the main function of the patterns explains why the LH-H% pattern is the most frequent pattern for embedded, non-final IUs of CIUs. In a chunk of speech, here the CIU, speakers phrase various grammatical units into separate prosodic units, for which the non-finality-signaling pattern, the LH-H%, is used to signal full integration. CIU internally, the LH-L% pattern is infrequent and only used in non-canonical constructions such as cleft constructions to signal less integration. The opposite holds true for singleton IUs and final IUs of CIUs. Here, the non-finality-signaling pattern LH-H% is used very infrequently, and in fact almost exclusively in lists, to signal non-finality. The finality-signaling pattern LH-L%, however, is very frequent. In both instances, that is CIU internally or CIU final, the L-H% pattern, signaling medium finality/non-finality is equally frequent.

The intonational model of the CIU in Totoli is based mainly on the inspection of tonal events. I showed that, internally in CIUs as well as in final position of CIUs, the same tonal patterns occur. This led me to postulate a recursively embedded structure of a Compound IU, rather than a hierarchical structure where higher-level units (i.e., the IU) consist of lower-level units (i.e., Accentual Phrases or intermediate phrases) and where higher-level units are tonally specified differently than lower-level prosodic units (see e.g., Jun & Fougeron 2000 for such an analysis of French and Ipek & Jun 2013 for Turkish).

Himmelmann's (2018) model of an IP in languages of Western Austronesia describes IU-internal tonal events as boundary-marking devices of smaller units, called the intermediate phrase (ip). His model proposes that the right-edge boundary of an IU consists of a phrase tone and a boundary tone and that ip-final boundary-tone complexes consist of a high target only. The model is reprinted in Figure 3.38.

$$\begin{array}{ccccc} \text{L\$} & \text{H\$} & \text{L\$} & \text{H\$} & \text{T-T\%} \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow\downarrow \end{array}$$

$$[[\sigma\sigma\sigma\sigma\sigma]_{ip}\ [\sigma\sigma\sigma\sigma\sigma\sigma]_{ip}\ \sigma\sigma\sigma\sigma]_{IP}$$

Figure 3.38: Himmelmann's (2018: 360) model of the IP in Western Austronesian languages: T$ representing an ipboundary tone, T% an IP boundary tone, and T- and IP phrase accent.

Before discussing the theoretical implications of such an analysis, it is worth examining two further illustrative examples of Tail-Head Linkage (THL) constructions (see §3.2.1.5 for an explanation of THLs). These make an interesting case in point. The Tail IU is repeated in part or in full and serves as the Head IU of the subsequent paragraph. The Tail always bears the IU-final boundary-tone complex(L)H-L% and the Head always has the L-H% boundary-tone complex. Two examples are given below in (22) and (23). Consider first (22) and its periogram in Figure 3.39. The first IU is repeated after a long pause and the adverb *indʒan* ‘after’ is added, pointing to the adverbial status the Head of a THL construction holds for the subsequent paragraph.

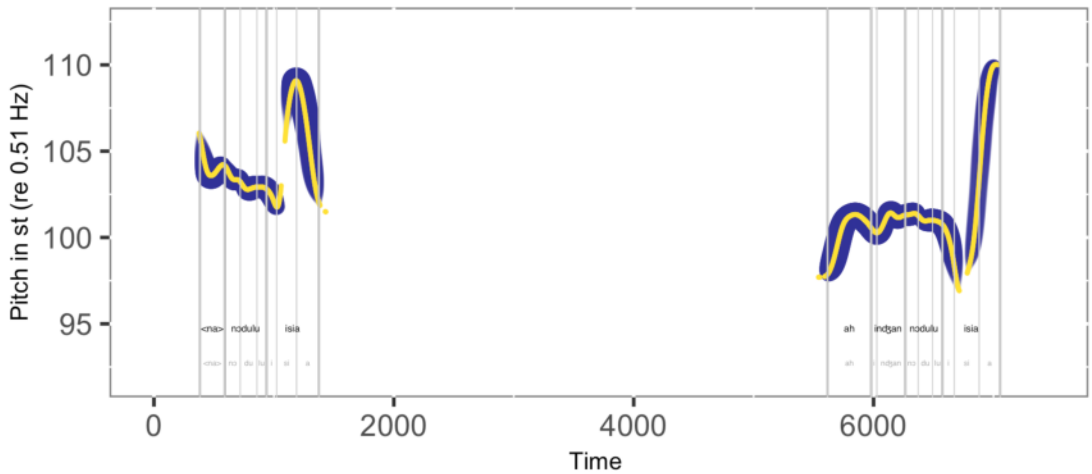

Figure 3.39: Periogram with pitch track (in st) for example (22); speaker SUD

(22) a. *<na> nɔdulu isia*
nɔ-dulu isia
AV.RLS-help 3s
‘helping him’

b. *ah indʒan nɔdulu isia*
ah indʒan nɔ-dulu isia
INTJ after AV.RLS-help 3s
‘after helping him’ (pearstory_38_SUD.056) ▹

The final boundary-tone complexes and the subsequent pauses mean that the status of examples (22a) and (22b) as separate CIUs is unambiguous. However, in some instances, the Head is directly followed by further syntactic material, as in the THL in example (23), for which the periogram given in Figure 3.40.

(23) a. *maŋana umbasan dɛdɛŋ*
maŋana umbasan dɛdɛŋ
child young.man small
‘(there was) a young boy’

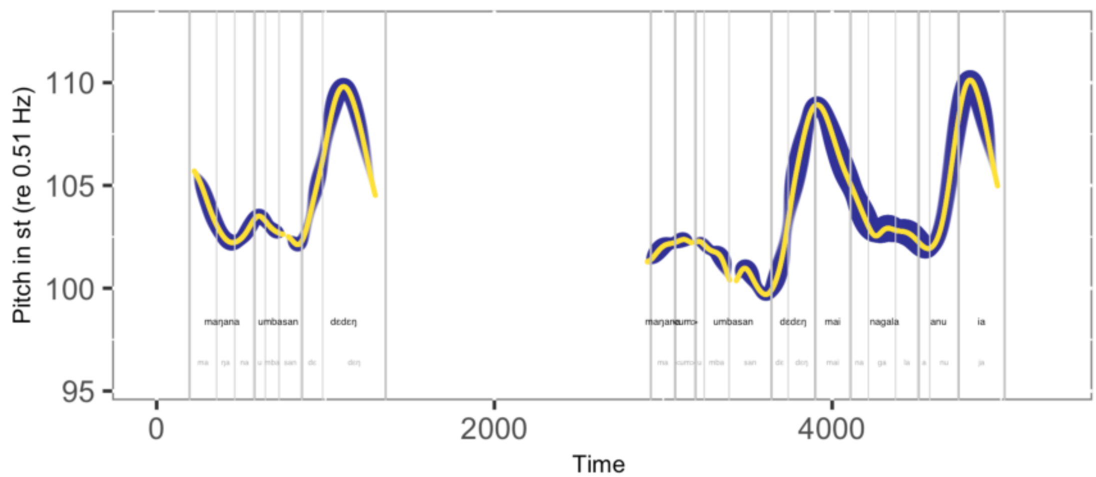

Figure 3.40: Periogram with pitch track (in st) for example (23), speaker AT

b. *ma <um> umbasan dɛdɛŋ | mai nagala anu ia*
maŋana umbasan dɛdɛŋ mai nɔg-ala anu ia
child young.man small come AV.RLS-fetch FILL PRX
'(there was) the boy; he comes to take the thingy'
(pearstory_23_AT.039-40) ▷

As expected, the Head *maŋana umbasan dɛdɛŋ* 'young boy' bears the final boundary-tone complex LH-L%, as it is the standard pattern for Tails of THLs. The difference pertains to the Heads of the two THLs in example (22b) and example (23a). In (23b), the Head – here in near exact repetition – is followed by further syntactic material, without any pause, pitch reset, or other boundary phenomena. In this case, the Head would have to be analyzed as the first ip of the IP, if one applies Himmelmann's (2018) model.

IP	IP
ip	ip
[[<na> nɔdulu isia]]	[[ah inʤan nɔdulu isia]]

Figure 3.41: Prosodic organization of examples (22a)–(22b), according to Himmelmann's 2018 model

However, the tonal events at the right edge of the Heads (22b) and (23b) are the same in both of the THL constructions. In example (22), we would label it L-H%

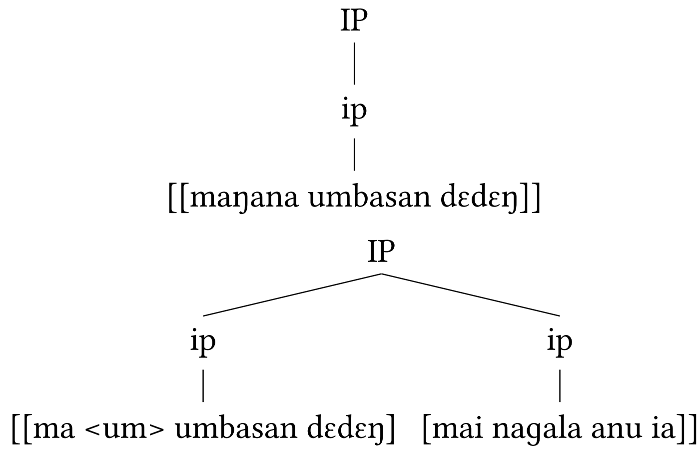

Figure 3.42: Prosodic organization of examples (23a)– (23b), according to Himmelmann's 2018 model

as it occurs in IP-final position. In the second THL construction in example (23), we would have to label the final tonal pattern L-H$, as it occurs at the edge of an ip, integrated into an IP ($ indicates an ip boundary tone in Himmelmann's 2018 model). However, the model assumes that (non-final) ips and IPs are tonally differentiated. We could do away with this seeming contradiction by assuming no further intonational level between the phonological word and the IU/IP, and by describing IUs such as those in example (23b) as recursively parsed into IUs:

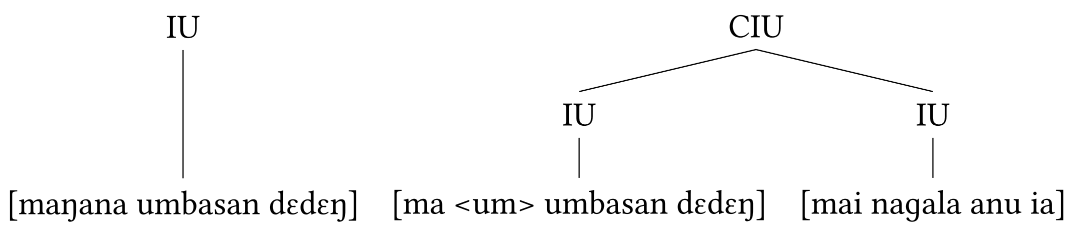

Figure 3.43: Alternative prosodic organization of example (23)

One reason for opposing the alternative analysis in Figure 3.43 is the Strict Layer Hypothesis (SLH; Selkirk 1986, Nespor & Vogel 1983, 1986, Vogel 2019), which predicts that any prosodic structure consists exhaustively of units of the next level down in the prosodic hierarchy, and allows no recursivity.

Though widely applied, the SLH causes empirical problems (see the discussion in Ladd 2008: chapter 8.2). With evidence from tone sandhi in Xiamen Chinese, Chen (1987) shows that tone groups and IUs regularly intersect and hence violate the SLH in that a tone group may be associated with two IUs. On the issue of overlapping domains in Luganda, Hyman et al. comment:

> The alternative in Luganda which we consider in work in progress is that the SLH and some of the claims of its advocates must be significantly weakened to allow cyclic assignments of postlexical domains. (1987: 107)

The model proposed by Himmelmann (2018: 369) analyzes IU-internal tonal events as boundary tones of ips and leaves open the status of IU-final material that follows the last ip-final boundary, see Figure 3.38.

One possibility is to analyze this as constituting an ip as well. Himmelmann argues that the Strict Layer Hypothesis (SLH; Selkirk 1986) would demand such an analysis, but points out that tonal targets are too different (his model assumes simple boundary tones at the right-hand edge of ips and boundary-tone complexes at the end of IUs) and that one would have to assume that the tune of IU-final ips are deleted or overwritten by the IU-final boundary tones. Himmelmann (2018) notes that IU-final boundary-tone complexes are of a different type and do not include ip-level tones, which in itself is again a violation of the SLH. However, I showed that final boundary tones are very similar, if not identical; hence, no overwriting rule would have to be postulated.

Abolishing the notion of an intermediate phrase level altogether and assuming recursive parsing of IUs into CIUs, we can avoid the difficulties in explaining that tunes are essentially the same except for the presence or absence of IU boundary phenomena. In THL constructions, one would avoid having to use different labels for essentially the same tonal pattern.

Also evident from the examples above are the obvious differences to the Accentual Phrase (AP), the postulated prosodic unit below the IU in many of the prosodic descriptions in the two volumes edited by Jun (2005c, 2014a). Not only does the AP differ from the IU in its tonal marking but the "AP has been typically defined as a tonally marked prosodic unit which contains one word" (Jun 2014b: 532). Example (15) in Figure 3.29 above is a particularly instructive instance of an adverbial phrase realized as embedded IU that spans 5 words / 15 syllables and has a near-flat level contour except for the right-edge boundary-tone complex (*kaasikan mɔgiigitai maŋana dɔlagɔ itu* 'because of his excitement in looking at the girl …'). The example adverbial phrase is uttered as one prosodic phrase and clearly larger than the AP and the prosodic word. Therefore, we cannot do away with recursive embedding of IUs by assuming simply the level of the non-recursively embedded IU and the phonological word or the AP or both.

The prosodic organization in Figure 3.43 equates to the model proposed by Ladd (2008: 297) which he calls the Compound Prosodic Domain (CPD): "A CPD is a prosodic domain of a given type X whose immediate constituents *are themselves of type X*."

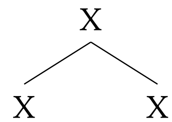

Figure 3.44: Exemplification of Compound Prosodic Domains, reproduced from Ladd (2008: 297)

In a more recent account by Selkirk (2011) termed *Match Theory*, recursivity is permitted and attributed to syntactic constituency-respecting *Match Constraints*.

Evidence obtained from an inspection of the IUs as they occur in the corpus of Totoli leave no doubt that tonal events at the edges of prosodic units are essentially the same. Therefor, I argue that complex IUs are best be described as CIUs that consist of a string of embedded IUs.

I started by assuming that speech is chunked into prosodic units, which are demarcated by a set of boundary phenomena and which is perceived as such by listeners (§1.4 and §3.1). I then described and categorized the tonal events at the end of such units (§3.2.1). In a further step, I looked at tonal events within such units and found that they are essentially the same as those that occur at the end (§3.2.2). I concluded that they are also right-edge boundary tones of prosodic units which regularly match syntactic units. Based on the observation that tonal events of all kinds of prosodic units are essentially the same, I argue for assuming recursive embedding of IUs into Compound IUs. The results here show that tonal contours are engaged at the level of the IU but not the CIU. It is crucial to point to the fact that the argument for recursion is only based on the tonal realization of prosodic units alone.

Further evidence for recursion come from syntax, as briefly shown above. In section §3.3, I discuss this aspect more thoroughly by comparing the syntactic content of singleton IUs with that of embedded, non-final IUs of CIUs.

3.3 Intonation Units and grammatical units in Totoli

In the previous Chapter 3.2, I presented an in-depth analysis of the tonal patterns of prosodic units in Totoli. In this section, I will investigate the syntactic content of prosodic units in the Totoli corpus (see §3.1.1) and analyze the grammatical units they typically contain. Specifically, I will first investigate which structures are usually found in CIUs, whether they are singleton IUs or Compound IUs. Secondly, I will investigate the syntactic structures embedded in CIUs and compare them to those found in singleton IUs.

In this present work, I adopt a discourse-oriented approach based on a corpus of natural, (semi-)spontaneous, and unscripted speech. Working with such data highlights the flexibility in the syntactic content of prosodic units. The question arises of the type of syntactic content that can exist within a prosodic unit and whether there are any regularities in the relationship between them.

A confounding factor pertains to the concept of CIUs, which I have introduced in this study, referring to either singleton IUs or CIUs, as distinct from embedded IUs of CIUs. In my analysis, the term CIU denotes those prosodic units that are demarcated by typical boundary cues such as pitch reset and/or pause, as well as other criteria mainly related to pitch, rhythm, and voice quality, as mentioned in §3.1 above. Therefore, CIU is equivalent to IU as reported in the literature below, as compound intonational units are not posited for these languages.

According to Ladd (2008: 288), explicit phonetic definitions are necessary for determining the criteria of IU and prosodic domain types in general. One of the confounding factors he identifies in the segmentation of spontaneous speech is the presumption that the division of syntactic units into prosodic ones reflects syntactic criteria, with many assuming that:

> [...] the various prosodic domains are defined by descriptions of how syntactic structure is mapped onto prosodic structure. (Ladd 2008: 289)

One of the significant achievements in prosodic phonology was the realization that prosodic boundaries systematically differ from syntactic boundaries. This was famously discussed by Chomsky & Halle (1968), who provided the frequently cited example of right-branching relative clauses. The syntactic boundaries are reprinted in (24a) and the prosodic boundaries in (24b).

(24) a. This is [the cat that caught [the rat that stole [the cheese]]]
  b. this is the cat - that caught the rat - that stole the cheese

(Chomsky & Halle 1968: 372)

Such systematic misalignment of syntactic and prosodic boundaries is usually interpreted as the result of mapping a complex syntactic structure onto an "intuitively 'flatter' or 'shallower' prosodic structure'" (Ladd 2008: 290). Within Prosodic Phonology (Nespor & Vogel 1983, Selkirk 1986, Nespor & Vogel 1986), mapping constraints which describe the relation between syntactic and prosodic units were formalized. As Féry (2016: 62–63) puts it:

> The basic idea of all models accounting for the syntax-prosody mapping is that the syntactic component is submitted to an algorithm – a set of rules

> or constraints – the aim of which is to map a prosodic structure to it. Theoretical issues relate to the way this correspondence is formulated as well as to the resulting prosodic constituency.

In recent years, new alignment constraints have been proposed, including Wrap Theory (Truckenbrodt 1999) and Match Theory (Selkirk 2011). The underlying assumption of such theories is that syntactic constituents correspond to prosodic units. In Match Theory, this assumption is expressed by the Match Clause, Match Phrase and Match Word constraints (Selkirk 2011: 5):

i. Match Clause: A clause in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ι [intonational phrase], in phonological representation.

ii. Match Phrase: A phrase in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ϕ [phonological phrase], in phonological representation.

iii. Match Word: A word in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ω [phonological word], in phonological representation.

Ladd (2008: 289) comments on these accounts:

> In my view, it makes no sense to treat accounts like Nespor and Vogel's or Selkirk's as definitions; rather, they are hypotheses, predictions about the correspondence between one type of independently definable structure and another. [...] Unless the syntactic and the phonological structures are defined in their own terms, the whole exercise becomes purely circular.

Focusing on natural data, works by Iwasaki & Tao (1993), Schuetze-Coburn (1994), Croft (1995), Tao (1996), Iwasaki (1996) and more recently Croft (2007), Park (2002), Matsumoto (2003) and Wouk (2008) have provided detailed descriptions of the syntactic content of IUs as they are found in corpora of spontaneous speech from a variety of typologically unrelated languages. These accounts have shown the flexibility of the syntactic content of IUs but have also revealed some regularities.

One tendency found in these studies is that approximately 50% of all IUs in a corpus consist of a simple clause, e.g. 47.8% in English (Croft 1995: 849), 50.5% in

Wardaman (Croft 2007: 12), 47.9% in Mandarin Chinese (Tao 1996: 72), and 51.7% in Sasak (Wouk 2008: 150).

Moreover, there seems to be a considerable number of IUs that consist of a single NP (referred to as 'lone NP' by Croft 2007: 12); for instance, 13.7% in English (Croft 1995: 849), 21.1% in Wardaman (Croft 2007: 11), 25.9% in Mandarin Chinese (Tao 1996: 72), and 21.0% in Sasak (Wouk 2008: 150).

However, genre appears to have a substantial influence on the proportions of IU types. This must be taken into account when comparing the results from different studies, as they vary with regard to the types of data used.

Tao (1996), Wouk (2008), Matsumoto (2000), Schuetze-Coburn (1994) and Park (2002) use conversations between two or more participants, whereas Croft (1995, 2007) bases his analysis on monological data. The cross-linguistic comparison in Croft (2007: 12) conflates both data types. Another reason why these results should be approached with caution is that different coding conventions have been applied. In two different studies on Japanese, the difference in coding conventions leads to an 11.6 percentage point difference in the proportion of clausal IUs in Japanese (57.0% in Matsumoto 2000: 58 and 45.4% in Iwasaki & Tao 1993: 3). These factors have to be taken into consideration when comparing the results on the reported data (see also the comments in Park 2002: 642, Croft 2007: 12, and Wouk 2008: 139–144).

Despite the differences, these studies have provided cross-linguistic evidence which confirms the centrality of the (simple) clause and the lone NP with regard to grammatical structures typically found in IUs. The sentence, on the other hand, appears to be rather difficult to identify in spoken speech:

> This is not a problem found with most other grammatical units, such as the clause or the phrase, which are generally clearly identifiable in spoken language. Yet the sentence is generally taken to be the basic unit of syntactic analysis. On the other hand, a sentence cannot be equated with an IU, the spoken-language analyst's most popular unit of choice for analysis. An IU does not grammatically correspond to a sentence, since it frequently is a unit smaller than a sentence and sometimes (though quite rarely) is not a full grammatical constituent at all. (Croft 1995: 841)

As described above, the majority of IUs contain a clause or a phrase. Other types include IUs consisting of a single connective or an interjection. Croft (2007: 11) specifically argues that these should similarly be thought of as constituting independent grammatical units. He calls these "lexical IUs". Another central observation is that the number of broken or interrupted IUs, such as (uncorrected) false starts, disjointed IUs and fragmentary IUs, is very small (2% in the corpus on Wardaman, see Croft 2007: 11).

In sum, speakers produce short stretches of speech which are rarely broken or fragmented and which mostly contain a full grammatical unit. Based on these observations, Croft (1995: 845) proposes the *full Grammatical Units condition*:

> The overwhelming preference for IUs to be in the form of full GUs [grammatical units], other things being equal, will be called the *full GU condition*.

In the same article, Croft (1995: 872) offers a possible explanation to account for both (a) the small number of broken IUs and (b) the high number of rather short and syntactically simple IUs. Croft calls it the *IU storage hypothesis*.

Intonation Units are explained as cognitive units and are considered "linguistic expressions of focuses of consciousness, whose properties apparently belong to our built-in information-processing capabilities" (Chafe 1980: 48). As there is no inherent constraint on the size of IUs per se, there must be some sort of cognitive limitation. Croft's (1995: 873) *IU storage hypothesis* suggests that Intonation Units consist of grammatical units that are stored or precompiled in the memory of the speaker. He argues that this accounts for the overwhelming frequency of IUs consisting of a single clause or single NP. More complex structures need to be computed based on the precompiled or stored grammatical units, which is why complex structures, such as multiply embedded NPs, rarely occur in spontaneous speech and are usually broken across several IUs. Croft (1995: 873) explains this in terms of the cognitive limitations of humans:

> Stored/precompiled constructions – and IUs themselves – may be the manifestation of the limitations of short-term memory in processing. The IU storage hypothesis suggests that grammatical structure and organization have evolved to conform to the limitations as well as the capacities of the human mind, specifically those embodied in IU structure.

Chafe (1994: 108) offers another cognitively grounded explanation to account for the types of structures typically found in IUs, the *One New Idea Constraint*:

> Conversational language appears subject to a constraint that limits an intonation unit to the expression of no more than one new idea.

Chafe argues that speakers can only activate one concept at a time and the IU is the basic unit used to express this cognitive process. If a simple clause with one predicate and one argument is the typical exponent of an IU, then only one of the two elements expresses a new idea. Chafe acknowledges that counterexamples from spontaneous speech are plentiful and offers a variety of explanations for such structures. Himmelmann et al. (2018) offer a way to measure the information content of an IU by computing the average number of content words an IU typically contains. Their study on four typologically unrelated languages found an average of 1.6-1.8 content words per IU (see also the discussion on various ways of measuring the length of IUs in §3.1.3).

The *One New Idea Constraint* is limited to those IUs that Chafe (1994: 63) refers to as *substantive IUs*, that is, those which express ideas of events, states or referents. The other major IU type is *regulatory IUs*, which regulate interaction and information flow. A third and minor IU type is *fragmentary IUs*.

On that matter, Chafe (1994: 119) comments:

> In any case, the finding that people can activate only one new idea at a time, as well as the insight that finding gives us into what it means to constitute "one idea," may be at least as important as the finding that short-term memory is limited to seven items plus or minus two (Miller 1994). The magical number one appears to be fundamental to the way the mind handles the flow of information through consciousness and language.

Yet, cognitive limitations are not the only constraints at work. Research by Park (2002: 674) has shown that the IU is a "resource that participants in an interaction may use and manipulate to achieve their interactional goals." He shows that *substantive IUs*, too, are subject not only to cognitive constraints but also to interactional constraints. In what follows, I contrast different aspects of the IU as they occur in conversation and in monological recordings. I show that genre has a substantial effect on the size of IUs, and detail the proportion of (Chafeian) IU types and the proportion of IUs with regard to their syntactic content, among other results.

In this section, I investigate the grammatical structures that prosodic units typically contain. It is organized into four parts: In the first part, §3.3.1, I explore the grammatical structures found in CIUs – either singleton IUs or CIUs. In the second section, §3.3.2, I explore the grammatical structures found in embedded IUs of CIUs. The following section, §3.3.3, aims to compare the two from a syntactic point of view. In section §3.3.4, I review the findings from the analysis of tonal patterns occurring at the edges of prosodic units and revisit the evidence for recursive embedding of IUs in Totoli with the evidence from syntax.

3.3.1 Syntactic structures of singleton IUs and CIUs

I briefly discussed above several studies that have addressed the question of what grammatical structures are typically found in IUs (Croft 1995, 2007, Tao 1996, Park 2002, Schuetze-Coburn et al. 1991, Schuetze-Coburn 1994, Matsumoto 2003, Iwasaki 1996, Iwasaki & Tao 1993, Wouk 2008). The studies vary substantially with regard to the data they are based on. Yet, it is to be expected that genre has a substantial influence on the proportions of IU types. The influence of genre has been anticipated by Du Bois (1987: 836):

> It is worth emphasizing that, while conversation may well be the more frequent genre, narrative is especially likely to display conditions of relatively high information pressure (...) The heavy information pressure demands in narrative may well give it significance beyond what it otherwise would have for the adaptive shaping of grammar in response to discourse needs.

Despite this obvious fact, Croft (2007: 12) makes cross-linguistic claims about the syntactic nature of IUs by comparing the proportions of grammatical structures reported in different studies. The present work is based on a corpus of conversational and monological recordings, enabling a comparison of the proportions of different types of CIUs within these two data types. The analysis presented here systematically investigates the influence of genre – monological versus conversational – and demonstrates its strong impact on the proportions of grammatical structures found in CIUs. Comparing other, more subtle subtypes of genre is also conceivable and is likely to yield slightly different results concerning the distribution of the syntactic nature of prosodic units. In this chapter, I focus on two broad categories—conversations versus monologues—only, as per Du Bois's 1987 indication that the difference in information pressure is most pronounced in these categories.

3.3.1.1 Methodology and coding conventions

The analysis presented here examines the CIU, which refers to either a singleton IU or a Compound IU, as the primary unit of investigation. The study aims to determine the typical grammatical structures found within these units. As a result, the CIU is the sole domain for coding. It is a prosodic unit that consists of either a singleton IU or a CIU made up of a sequence of IUs, which is delimited by boundary marking cues such as a pause, a break in pitch contour, and a pitch reset.

This is illustrated in the three CIUs presented in examples (25a–25b). The second singleton IU in (25b) contains the noun *sɛllɛŋgɛt* 'one basket', which can be analyzed either as an argument to the verb in the preceding singleton IU in (25a) or as part of the following CIU in (25c). However, because it constitutes its own singleton IU, (25b) is analyzed as a nominal IU, while both (25a) and (25c) are considered clausal CIUs.

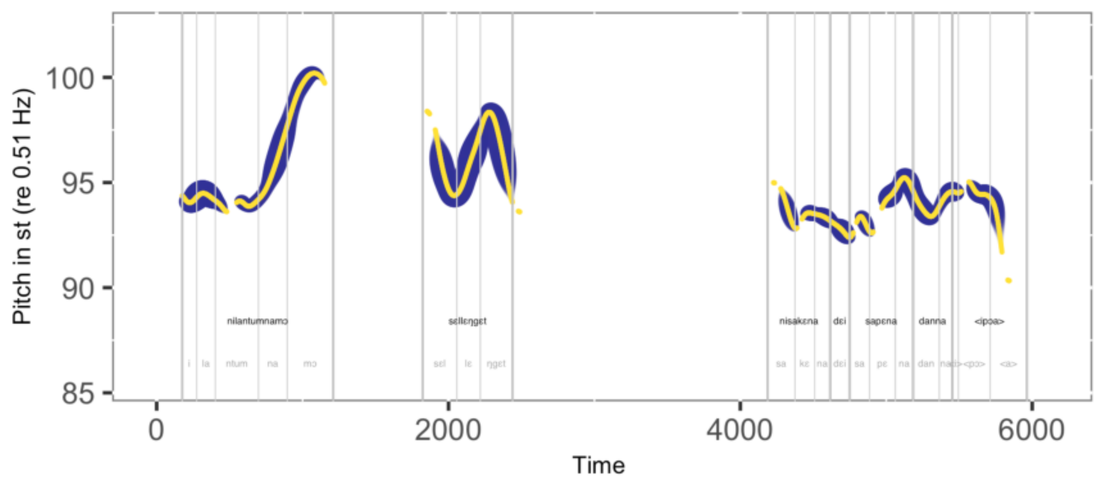

Figure 3.45: Periogram with pitch track (in st) for example (25), speaker SP

(25) a. *ilantumnamɔ*
ni-lantum-0=na=mɔ
RLS-bring.along-UV=3S.GEN=CPL
'he brought (it)'

b. *sɛllɛŋgɛt*
so-RDP-lɛŋgɛt
ONE-RDP-basket
'a basket'

c. *sakɛna dɛi sapɛda danna <ipoa>*
sakɛ-0=na dɛi sapɛda danna
get.on-APPL=3S.GEN LOC cycle then
'he put (it) on the bicycle and then' (pearstory_14_SP.019-21) ▷

The structural relations between CIUs are only examined in relation to specific aspects of their internal distinctions, which will be discussed in the following sections. In order to provide adequate context and facilitate understanding for the reader, I include the CIUs adjacent to the examples being discussed, although they may not always be explicitly elaborated upon.

3.3.1.2 Discussion of CIU types

Following Tao (1996), I differentiate four categories: (a) clausal CIUs, (b) nominal CIUs, (c) interactional CIUs and (d) other minor types. I will discuss these in the following sections.

3.3.1.2.1 Clausal CIUs

In this study, a clausal CIU is defined as one that contains at least one predicate. Two types of clausal CIUs are distinguished: independent and dependent. Moreover, independent CIUs are further classified based on the number of overtly expressed arguments. The definitions of these categories are detailed in the following sections.

Independent clausal CIUs The simplest form of an independent clausal CIU comprises a verbal predicate and a single overtly expressed argument, which may be either a lexical NP or a pronoun. An example of this is provided in (26), which illustrates a basic independent clause containing a preverbal pronominal argument *isia* 'he'.

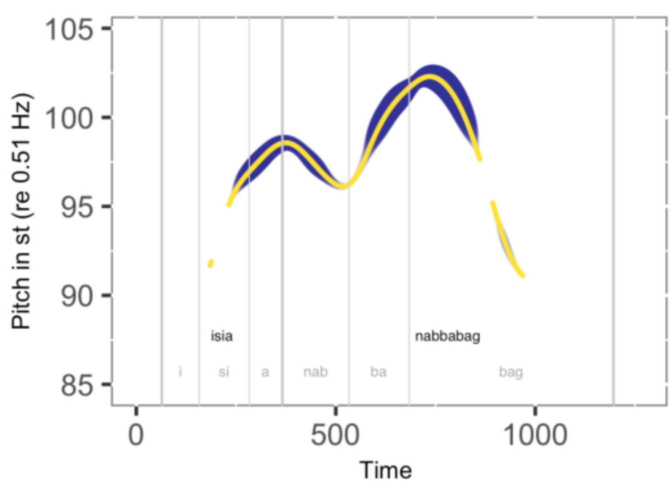

Figure 3.46: Periogram with pitch track (in st) for example (26), speaker IRN

(26) *isia nabbabag*
isia nɔ-RDP-babag
3S ST.RLS-RDP-crash.into
'he crashed into (it)' (pearstory_15_IRN.009) ▷

In Totoli, the agent argument of the undergoer voice is often realized as a clitic pronoun on the verb (for an explanation of the voice system, see Riesberg et al. 2019, Riesberg 2014). In the conventions of this study, such constructions are categorized as simple clauses with one overtly expressed argument. For instance, consider the three IUs in (27). The initial singleton IU (27a) comprises a verb with the agent argument expressed as the enclitic *=na* '3S.GEN'. The second CIU (27b) features the same verb, but the undergoer argument is expressed as the lexical NP *sapɛɔ itu* 'this hat', while the agent argument is unexpressed. The third singleton IU (27c) also contains a verb with the agent argument realized as an enclitic on the verb. According to the coding conventions employed in this study, all three CIUs are classified as simple independent clausal CIUs with one overtly expressed argument.

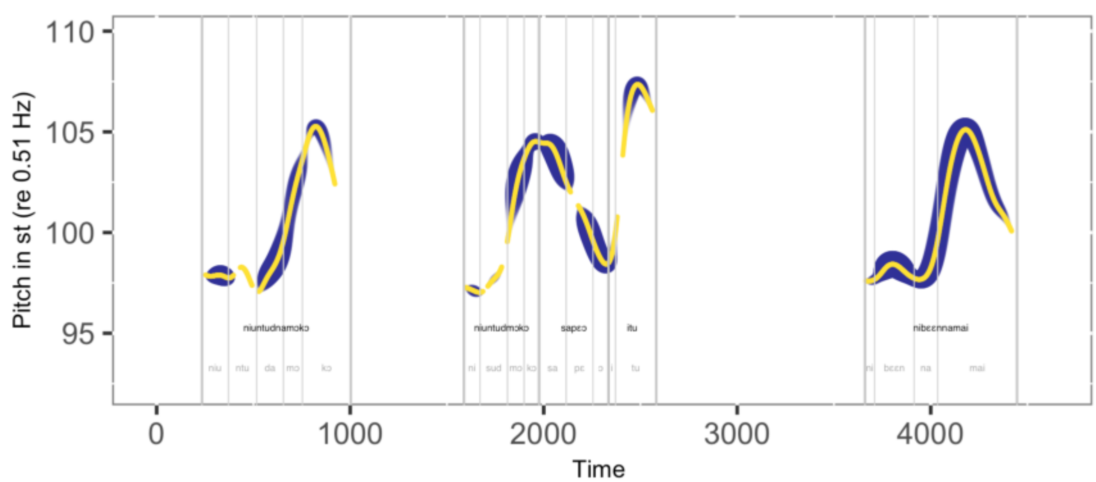

Figure 3.47: Periogram with pitch track (in st) for example (27), speaker SNG

(27) a. *niuntudnamɔkɔ*
ni-untud-0=na=mɔ=kɔ
RLS-bring-UV=3S.GEN=CPL=AND
'he brought (it)'

b. *niuntudmɔkɔ sapɛɔ itu*
ni-untud-0=mɔ=kɔ sapɛɔ itu
RLS-bring-UV=CPL=AND hat DIST
'(he) brought this hat'

c. *nibɛɛnnamai*
ni-bɛɛn-0=na=mɔ=ai
RLS-give-UV=3S.GEN=CPL=VEN
'he gave (it)' (pearstory_17_Sng.101-103) ▷

Oblique and core arguments are not distinguished in the analysis. Example (28) illustrates a clause where the negated verb *nɔliitaan* 'meet' is followed by an oblique argument introduced by the preposition *takin* 'with'. Despite being an oblique argument, this CIU is still coded as a simple independent clausal CIU with one overtly expressed argument.

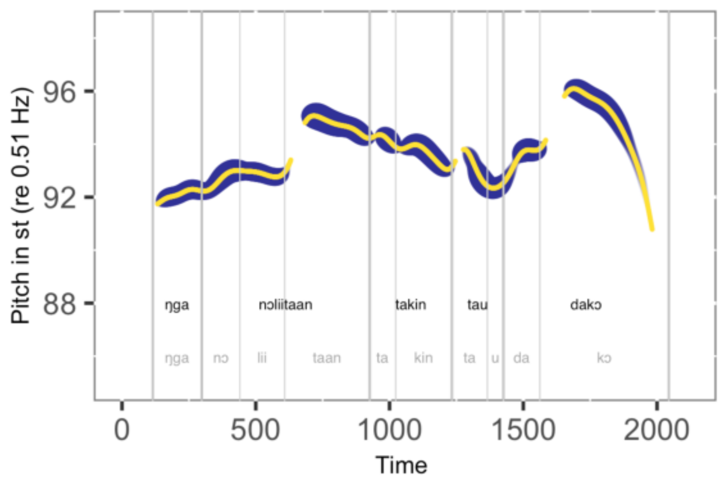

Figure 3.48: Periogram with pitch track (in st) for example (28), speaker RDA

(28) *ŋga nɔliitaan takin tau dakɔ*
iŋga nɔli-ita-an takin tau dakɔ
NEG RCP.RLS-see-RCP.RLS with person big
'(I) haven't met my parents' (lifestory_RDA_1.024) ▷

Equational predications are analyzed as simple clauses with one nominal predicate and one argument. In example (29), the CIU consists of an equational clause with two elements, *siritaku ia* 'my story' and *sirita tau pɔmɔɔ* 'the story of a person from the old times'. It is worth noting that Totoli does not use a copula. The CIU in example (29) is also considered a simple independent clausal CIU with one overt argument.

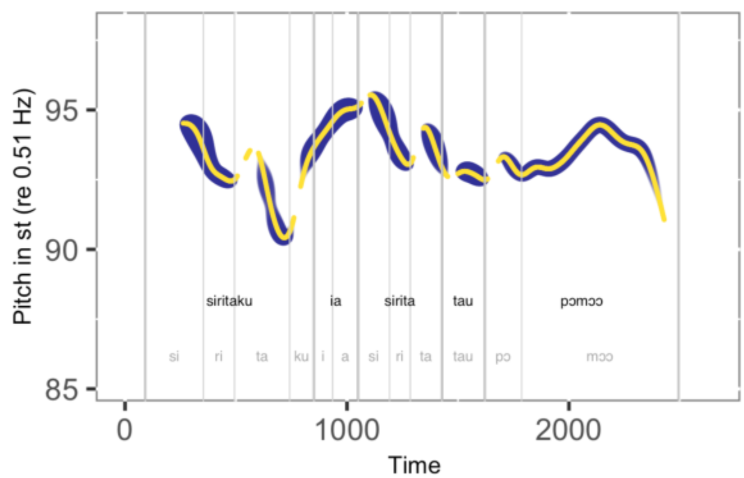

Figure 3.49: Periogram with pitch track (in st) for example (29), speaker RDA

(29) *sirita aku ia sirita tau pɔmɔɔ*
sirita aku ia sirita tau pɔmɔɔ
story 1S.GEN PRX story person first
'my story is the story of a person from the old times'
(lifestory_RDA_1.014) ▹

Totoli has two existential constructions. One construction involves a form of the existential predicate *daan*/*kaddaan*/*dadaan* 'EXIST'. The other construction involves an existential prefix *kɔ=*.

Examples of the existential prefix are given in the singleton IUs in (30a) and (30b). The bases *badu* 'shirt' and *sampaŋ* 'pants' occur with *kɔ=* and in this case with the negator *ŋga*. Each constitutes a full clause.

(30) a. *ŋga kabadu*
iŋga kɔ=badu
NEG EXIST-shirt
'there were no shirts'

b. *ŋga kasampaŋ*
iŋga kɔ=sampaŋ
NEG EXIST-pants
'there were no pants' (lifestory_RDA_1.032-034) ▹

An example of a singleton IU containing a construction with an existential predicate is given in (31a).

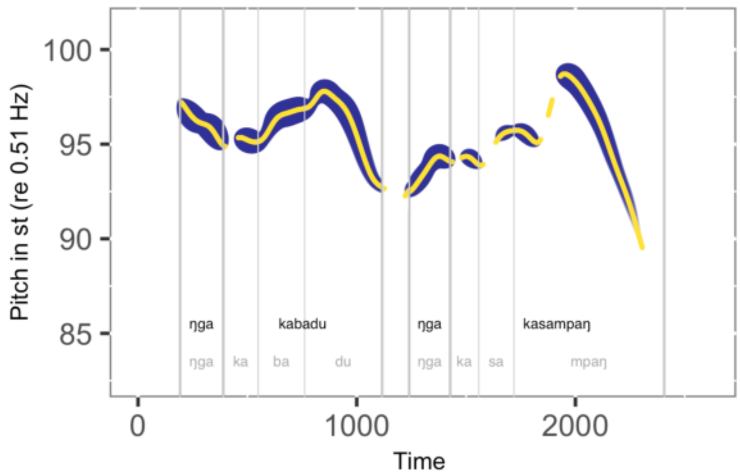

Figure 3.50: Periogram with pitch track (in st) for example (30), speaker RDA

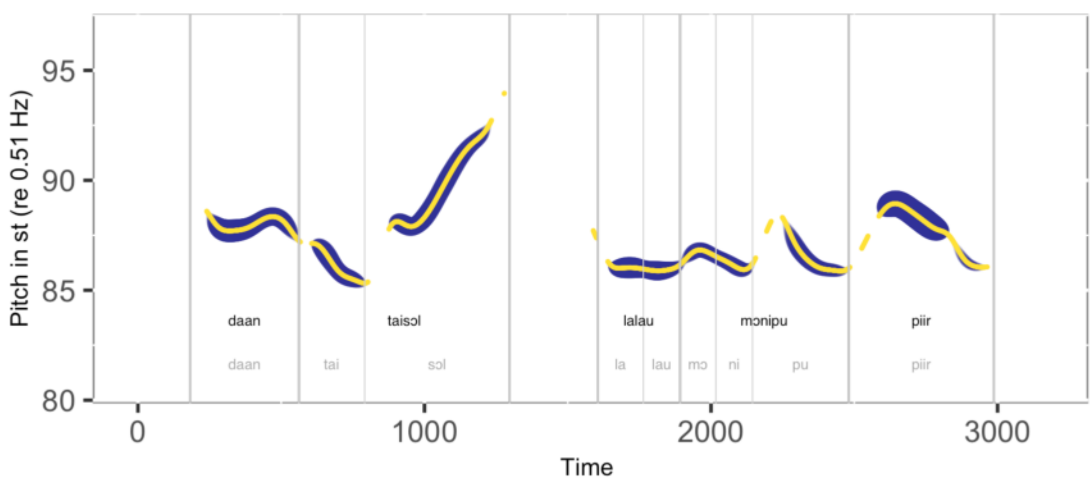

Figure 3.51: Periogram with pitch track (in st) for example (31), speaker FAH

(31) a. *daan taisɔl*
daan taisɔl
EXIST old.man
'there is an old man'

b. *laalau mɔnipu piir*
RDP-lau mɔN-tipu piir
RDP-presently AV-pick pear
'(he) currently picks pears' (pearstory_9_FAH.002-4) ▷

The constructions presented in (30a), (30b) and (31a) are full clauses. The three singleton IUs are considered simple clausal CIUs that consist of one (existential) predicate and one overtly expressed argument.

Clausal CIUs which contain more than one predicate and/or more than two overtly expressed arguments are referred to here as complex clausal CIUs. This includes CIUs containing a simple clause and a subordinate clause, e.g. a simple clause with a modifying relative clause or with an adverbial clause. Other cases are two coordinated clauses or a main clause with a complement clause parsed into one CIU. It is important to note that in Totoli, as well as in the local (Manado) Malay variety, a negated existential predicate is often used to negate entire clauses. In the counts used in this study, such constructions will appear as complex clausal CIUs since they involve two predicates. Such a construction is given in example (32). The clause *parhatikanna tau ipanau ia* 'he notices the people below' is negated with the negated existential predicate *daan*.

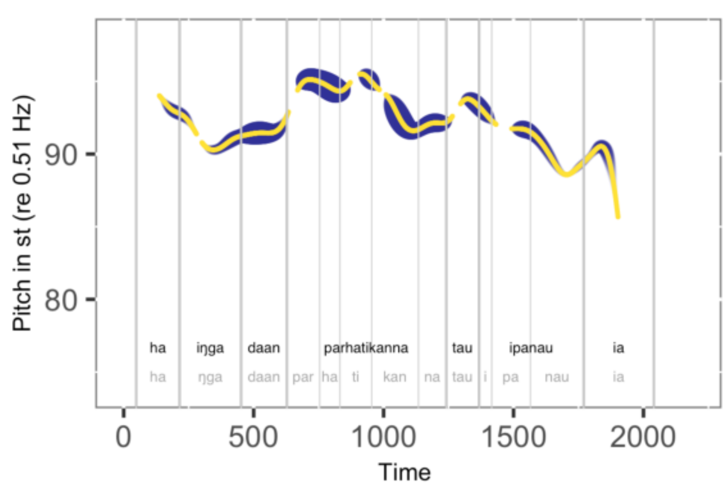

Figure 3.52: Periogram with pitch track (in st) for example (32), speaker RSTM

(32) *ha iŋga daan parhatikanna tau ipanau ia*
ha iŋga daan parhatikan-0=na tau i-panau ia
INTJ NEG EXIST pay.attention-UV=3S.GEN person LOC-under PRX
'He didn't notice the people below there' (pearstory_12_RSTM.090-92) ▷

An instance of two coordinate clauses parsed into one CIU is given in example (33). Note that no coordinating conjunction occurs.

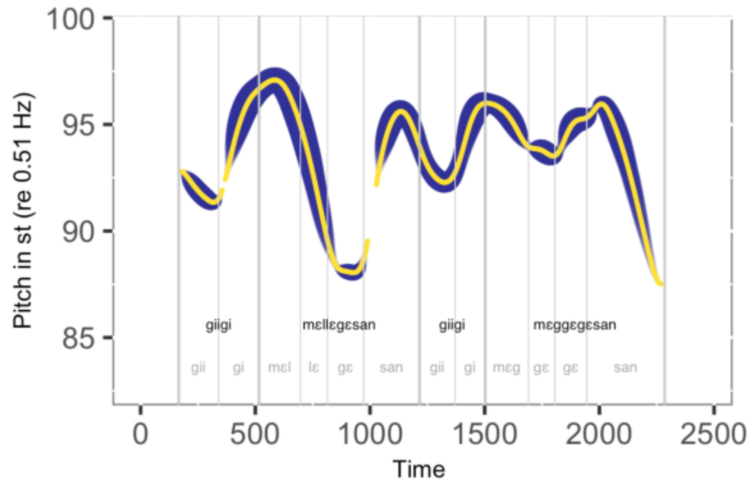

Figure 3.53: Periogram with pitch track (in st) for example (33), speaker SYNO

(33) *giigii mɛllɛgɛsan giigii mɛggɛgɛsan*
RDP-gii mo-lɛlɛgɛsan RDP-gii mo-RDP-gɛgɛs-an
RDP-different AV-Lelegesasn RDP-different AV-RDP-rub-APPL
'Singing Lelegesan is different from rubbing (your body).'
(lit. 'Singing Lelegesan is different (and) rubbing is different.')

(explanation-lelegesan_SYNO.032) ▷

Example (34b) is another instance of a complex clausal CIU which involves an adverbial clause and its matrix clause parsed in a single CIU.

(34) a. *nadabu sapɛɔna*
nɔ-dabu sapɛɔ=na
ST.RLS-fall hat=3S.GEN
'his hat fell'

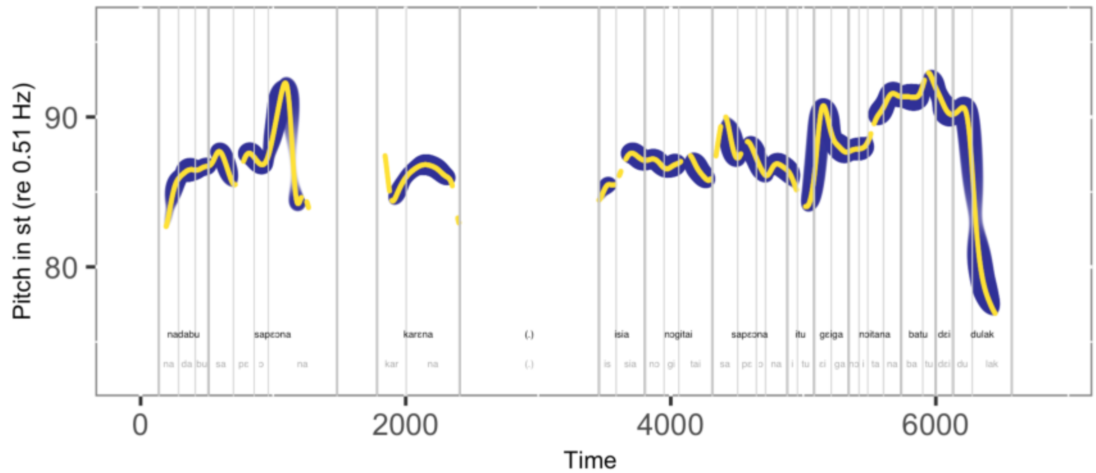

Figure 3.54: Periogram with pitch track (in st) for example (34), speaker FAH

b. *karɛna isia nɔgitai sapɛɔna itu gɛiga nɔitana batu dɛi dulak*
karɛna isia nɔg-ita-i sapɛɔ=na itu gɛiga
because 3S AV.RLS-see-APPL hat=3S.GEN DIST NEG
nɔ-ita-0=na batu dɛi dulak
POT-see-UV=3S.GEN stone LOC front
'Because he looks at the hat, he doesn't see the stone in front.'
(pearstory_9_FAH.026-27) ▹

Dependent clausal CIUs These include various adverbial clauses that occur in separate CIUs. An example is provided in (35a), where the initial element is a subordinating conjunction *indʒan* 'then' that unambiguously indicates its dependent status. Unlike (34b) above, the adverbial clause and its matrix clause are in two separate CIUs.

(35) a. *indʒan nɔpuliŋmɔ dɔua llɛŋgɛt itumɔkɔ*
indʒan nɔ-puliŋ=mɔ dɔua RDP-lɛŋgɛt itu=mɔ=kɔ
after ST.RLS-full=CPL two RDP-basket DIST=CPL=AND
'after the two baskets were full'

b. *nɔtumalibmɔkɔ tau gɔgɔɔt tɔalaŋ itu*
nɔ-t<um>alib=mɔ=kɔ tau RDP-gɔɔt tɔalaŋ itu
AV.RLS-<AUTO.MOT>pass.by=CPL=AND person RDP-hold goat DIST
'a person passed by holding a goat' (pearstory_12_RSTM.064-65) ▹

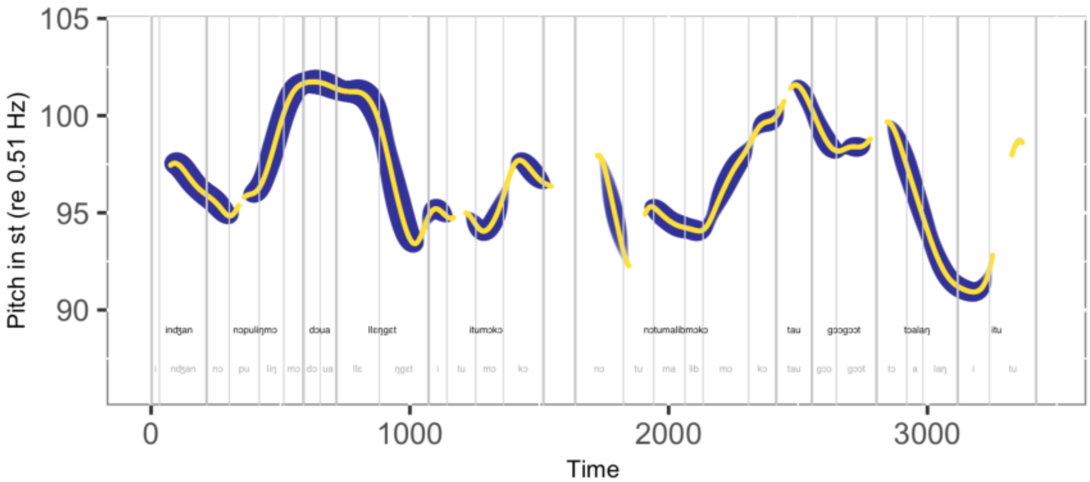

Figure 3.55: Periogram with pitch track (in st) for example (35), speaker RSTM

In several instances, the subordinated status of a clause is not indicated by a subordinating conjunction. Nonetheless, the intonation and the context clearly indicate its status. This will be discussed further in section §3.3.2. An example is provided in the IU in (36a), which includes an adverbial clause of either temporal or, more likely, causal status. However, no subordinating conjunction specifies one interpretation over the other.

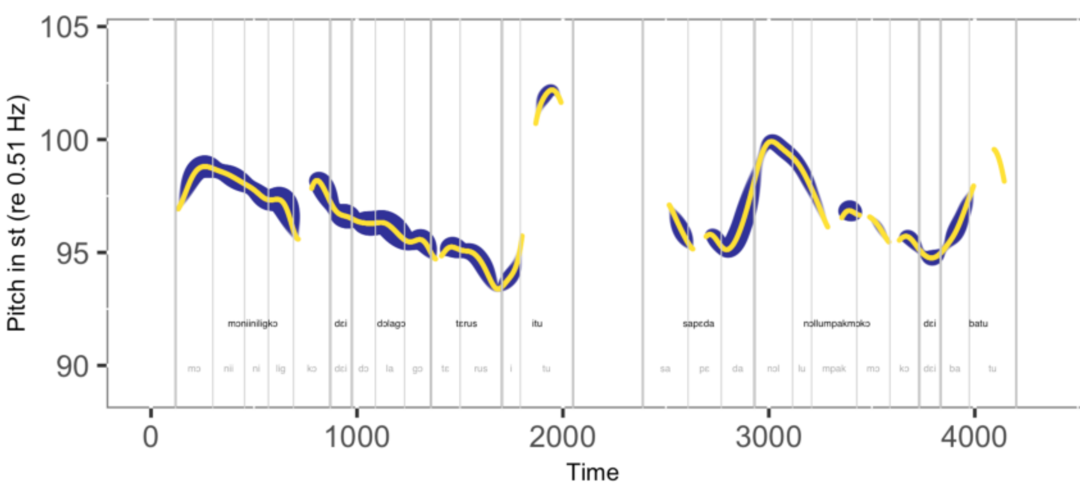

Figure 3.56: Periogram with pitch track (in st) for example (36), speaker SP

(36) a. *mɔniiniligkɔ dɛi dɔlagɔ terus itu*
mɔN-RDP-silig=kɔ dɛi dɔlagɔ tɛrus itu
AV-RDP-glance=AND LOC girl then DIST
'he looked at the girl constantly'

b. *sapɛda nollumpakmɔkɔ dɛi batu*
sapɛda nɔ-RDP-lumpak=mɔ=kɔ dɛi batu
bicycle ST.RLS-RDP-hit.against=CPL=AND LOC stone
'the bicycle crashes against the stone' (pearstory_14_SP.027-28) ▷

3.3.1.2.2 Nominal CIUs

Nominal CIUs are composed of either a single NP or a relative clause. The latter is included here because it is equally referential, hence the label "nominal CIUs" rather than "NP-CIUs" (cf. Croft 2007: 13 and Tao 1996: 79).

Croft (2007: 13) presents a basic categorization of nominal CIUs into three types:

Independent: are those nominal CIUs which have no structural relation with any of the adjacent intonation units.

Parallel: are separate CIUs containing "conjoined or appositive NPs" (Croft 2007: 13).

Arguments: are nominal CIUs that have a structural relationship with a neighboring CIU; i.e. they can be analyzed as an argument to a predicate of an adjacent clausal CIU.

The immediately adjacent CIUs are taken as the domain for category-internal classification of nominal IUs. A nominal CIU is considered an argument CIU if it can be analyzed as an argument of a clausal CIU immediately preceding or following it. Relative clauses with a head noun in the immediately preceding CIU are classified as parallel. Most free relative clauses that constitute their own CIUs can be analyzed as arguments and are classified as such; otherwise, they are classified as independent. Examples of all of these subtypes of nominal CIUs are provided below.

Nominal argument CIUs Example (37a) is an instance of a nominal argument CIU. It serves as an argument to the existential predicate in the subsequent CIU in (37b).

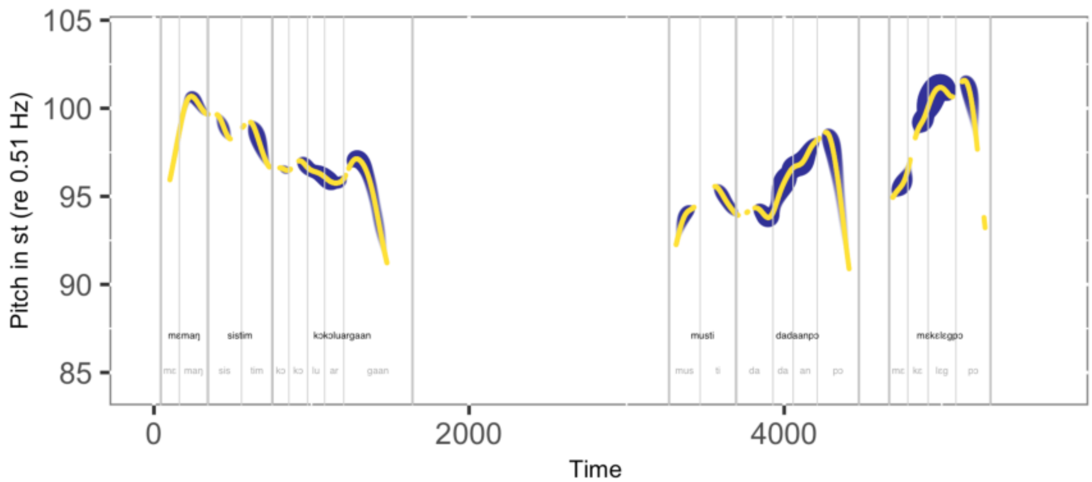

Figure 3.57: Periogram with pitch track (in st) for example 37, speaker ZBR

(37) a. *mɛmaŋ sistim kɔkɔluargaan*
mɛmaŋ sistim kɔ-kɔluarga-an
in.fact system NR-family-NR
'in fact the family system'

b. *musti dadaanpɔ*
musti RDP-daan=pɔ
have.to RDP-EXIST=INCPL
'has to remain'

c. *mɛkɛlɛgpɔ*
mɔ-kɛlɛg=pɔ
ST-strong=INCPL
'stay strong' (explanation-wedding-tradition_ZBR.249-251) ▷

A nominal CIU may also contain an NP with a modifier. For instance, in (38b), the CIU contains the NP *adfɔkaat* 'avocado' and the modifying relative clause *anu nilantumnakɔ ia* 'which he brought'. This CIU is a nominal argument CIU as it serves as the argument of the verb in the following clausal, singleton IU in (38c).

(38) a. *nɔllumpak dɛi batu*
nɔ-RDP-lumpak dɛi batu
ST-RDP-hit.against LOC stone
'(he) crashed against the stone'

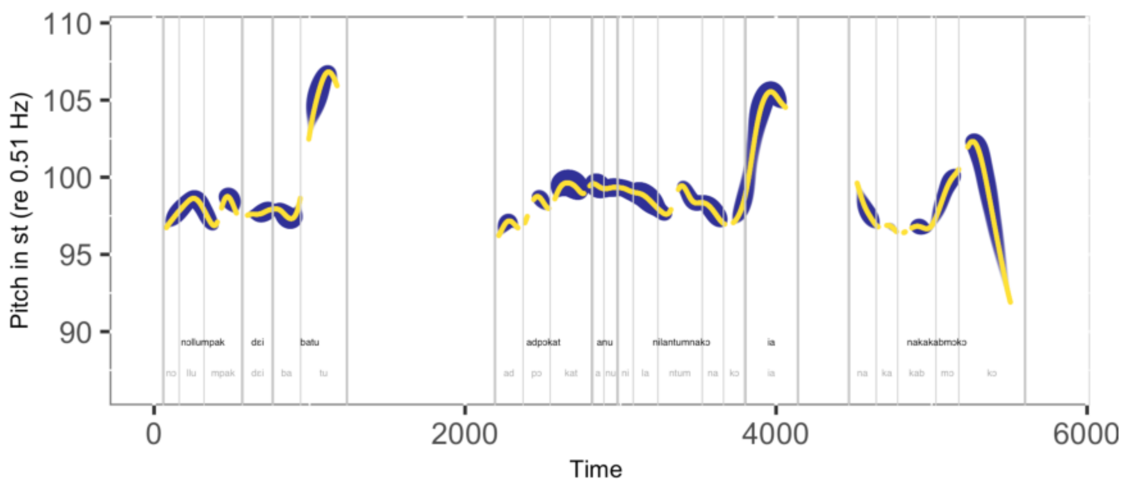

Figure 3.58: Periogram with pitch track (in st) for example (38), speaker SP

b. *adfɔkaat anu nilantumnakɔ ia*
alpukaat anu ni-lantum-0=na=kɔ ia
avocado REL RLS-bring.along-UV=3S.GEN=AND PRX
'the avocados which (he) brought'

c. *nakakabmɔkɔ*
no-kakab=mɔ=kɔ
ST.RLS-pour=CPL=AND
'scattered/poured' (pearstory_11_SP.027-28) ▷

CIUs consisting of a headless relative clause are argument CIUs if they serve as an argument to an adjacent CIU. Example (39b) is an instance of a headless relative clause phrased as a separate CIU. It serves as the argument to the verb in the preceding clausal CIU (39a).

(39) a. *sukati itaita*
sukat-i ita-i=ta
try-UV see-APPL=2S.GEN
'try to look for'

b. *anu saasalu dɛi puun kaju*
anu RDP-salu dɛi puun kaju
REL RDP-facing LOC tree wood
'the one facing the tree' (spacegames_sequence2_KSR-SP.198-199) ▷

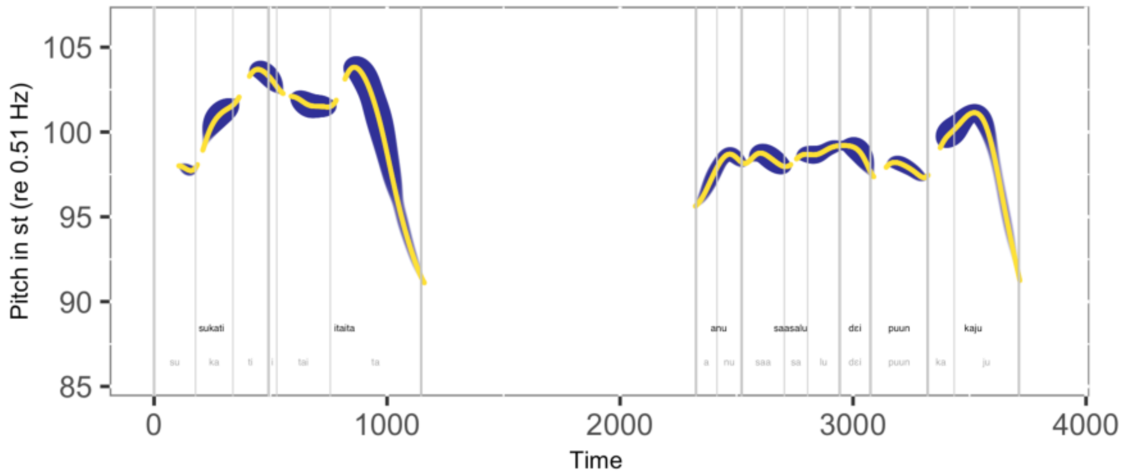

Figure 3.59: Periogram with pitch track (in st) for example (39), speaker SP

Parallel nominal CIUs The CIUs in (40b) and (40c) are examples of the parallel type. They are appositives of the argument *sasaakan* 'everybody' in the first CIU.

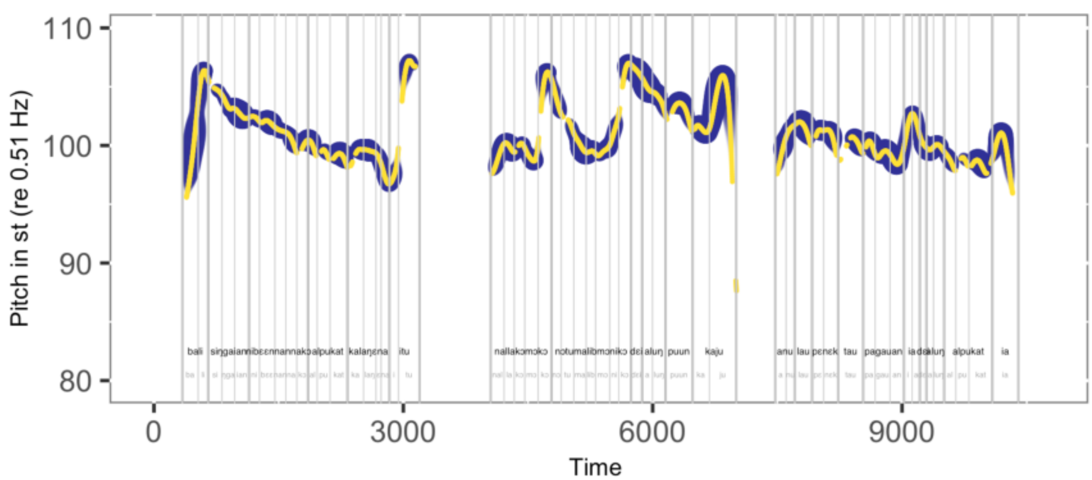

Figure 3.60: Periogram with pitch track (in st) for example (40), speaker SP

(40) a. *bali nnɛa ia mɔllinʤɔn sasaakan*
bali nɛnɛa ia mɔ-RDP-linʤɔn sasaakan
so today *prx* AV-RDP-gather everybody
'so today everybody gathers'

b. *ssaakan tau mɔntɔliusat*
sasaakan tau mɔntɔli-usat
all person BE.RELATED.AS-related
'all the relatives'

c. *mɔntɔliamaŋ*
mɔntɔli-amaŋ
BE.RELATED.AS-father
'the relatives of the father'
(explanation-wedding-tradition_ZBR.023-25) ▷

The singleton IU in (41c) is an example of a relative clause with its head in the preceding CIU, shown in (41b). It is analyzed here as a nominal CIU of the parallel type, as it modifies the head NP *puun kaju* 'the tree' in the preceding CIU.

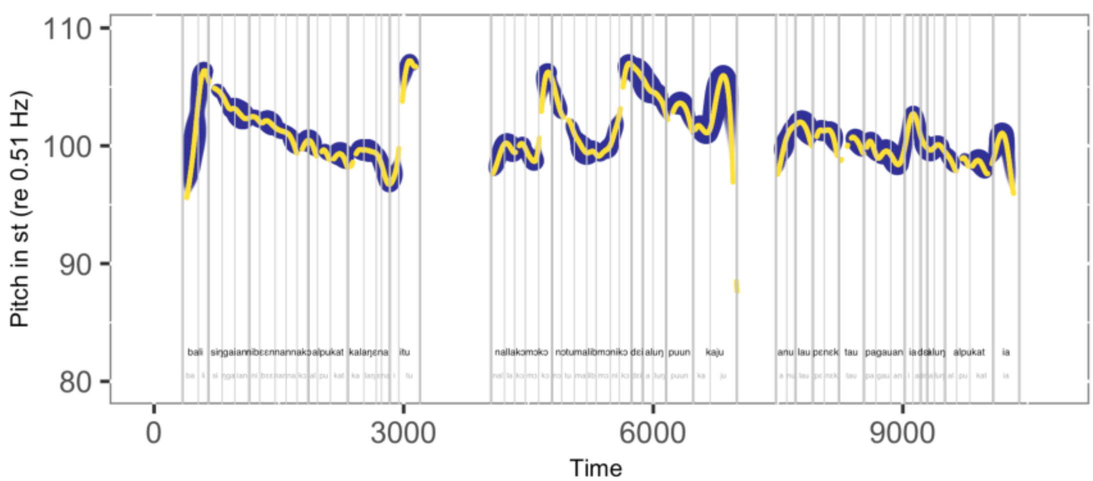

Figure 3.61: Periogram with pitch track (in st) for example (41), speaker SP

(41) a. *bali siŋgaian nibɛɛnannakɔ alpukaat kalaŋɛna itu*
bali siŋgaian ni-bɛɛn-an=na=kɔ alpukaat kalangɛna
so friend RLS-give-APPL=3S.GEN=AND avocado a.moment.ago
itu
DIST
'so the friends who were given the avocado earlier'

b. *nallakɔmɔkɔ nɔtumalibmɔ nikɔ dɛi aluŋ puun kaju*
nɔ-RDP-lakɔ=mɔ=kɔ nɔ-t<um>alib=mɔ
AV.RLS-RDP-walk=cpl=and AV.RLS-<AUTO.MOT>pass.by=CPL
pɔni=kɔ dɛi aluŋ puun kaju
again=AND LOC under tree wood
'they walked past, below the tree'

c. *anu lau pɛnɛk tau pagauan ia dɛi alung alpukaat ia*
anu lau pɛnɛk-0 tau pɔ-gauan ia dɛi alung alpukaat
REL presently climb-UV person GER-garden PRX LOC under avocado
ia
PRX
'which was just climbed up by the farmer under the avocados'
(pearstory_11_SP.043-45) ▷

Independent nominal CIUs These are nominal CIUs – often singleton IUs – which cannot be analyzed as bearing a structural relation with any adjacent CIU. They often perform a topic-introducing function (Croft 2007: 13), as in example (42b).

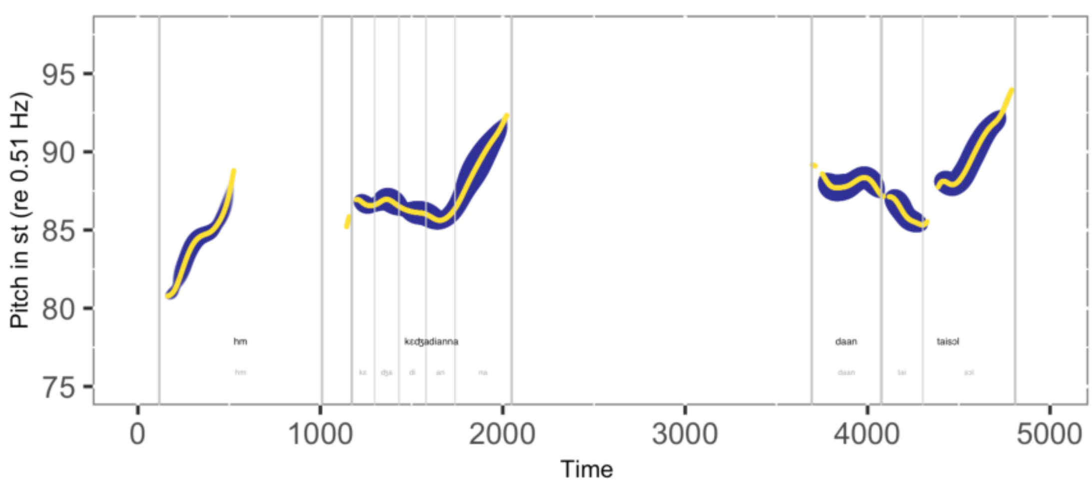

Figure 3.62: Periogram with pitch track (in st) for example (42), speaker FAH

(42) a. *hm*
'hm'

b. *kɛʤadianna*
kɛʤadian=na
event=3s.GEN
‘the situation’

c. *daan taisol*
daan taisɔl
EXIST old.man
‘there is an old man’ (pearstory_9_FAH.001-3) ▹

3.3.1.2.3 Interactional CIUs

This category includes CIUs – often singleton IUs – that consist of an interjection such as *eh, mm, io* and other discourse markers; see example (43b).

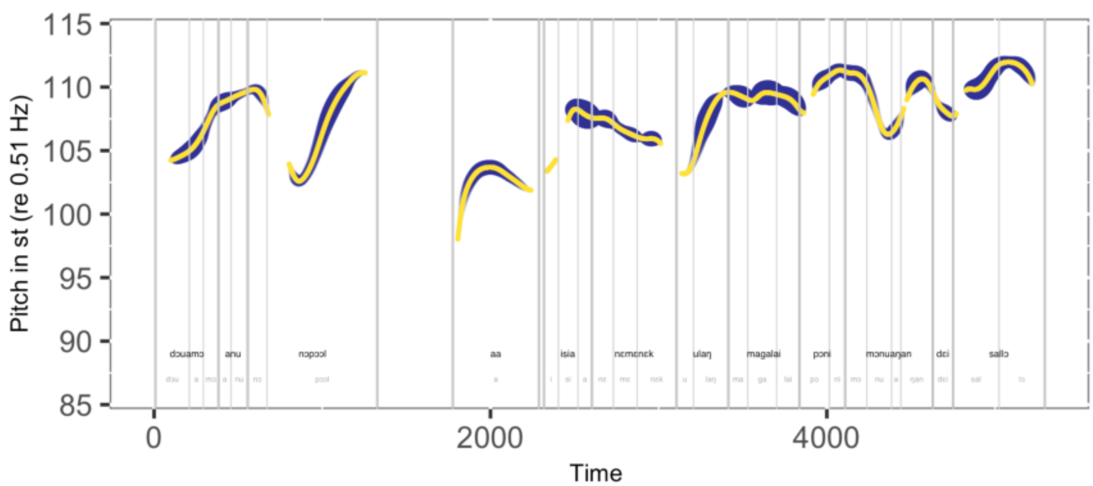

Figure 3.63: Periogram with pitch track (in st) for example (43), speaker RD

(43) a. *dɔuamɔ anu nɔpɔɔl*
dɔua=mɔ anu nɔ-pɔɔl
two=CPL REL ST.RLS-full
‘two are already full’

b. *aa*
aa
INTJ
‘aa’

c. *sia nɛmɛnɛk ulaŋ magalai pɔni mɔnuaŋan dɛi sallɔ*
isia nɔN-pɛnɛk ulaŋ mɔg-ala=ai pɔni mɔN-suaŋ-an dɛi
3S AV.RLS-climb repeat AV-fetch=VEN again AV-fill-APPL LOC
sallɔ
basket
'he climbs again, to take again and put it in the basket'
(pearstory_13_RD.025-27) ▷

The filler element *anu* is considered an interactional singleton IU only if it occurs as a bare root in a separate singleton IU such as in (44b). In the presence of verbal morphology, it is coded as clausal, as it is usually smoothly integrated into the clause structure. See the two CIUs in examples (45a)–(45b).

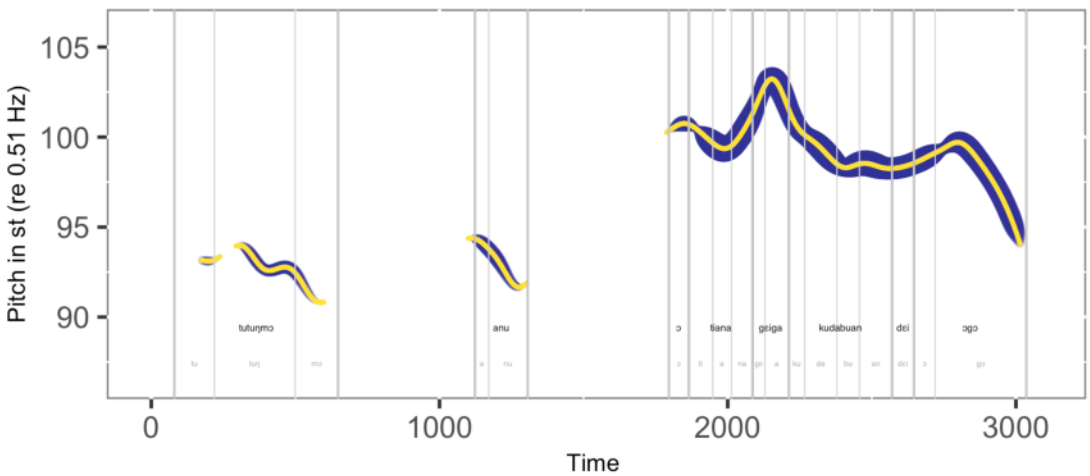

Figure 3.64: Periogram with pitch track (in st) for example (44), speaker RSM

(44) a. *tutuŋmɔ*
tutuŋ=mɔ
burn=COMPL
'Burn me!'

b. *anu*
anu
FILL
'...thinggy...'

c. *o, tiana, gɛiga, kudabuan dɛi ɔgɔ*
o tiana gɛiga ku=dabu-an dɛi ɔgɔ
INTJ QUOT NEG 1SG=throw-UV LOC water
'Oh no, she says, I will throw you in water!' ▹

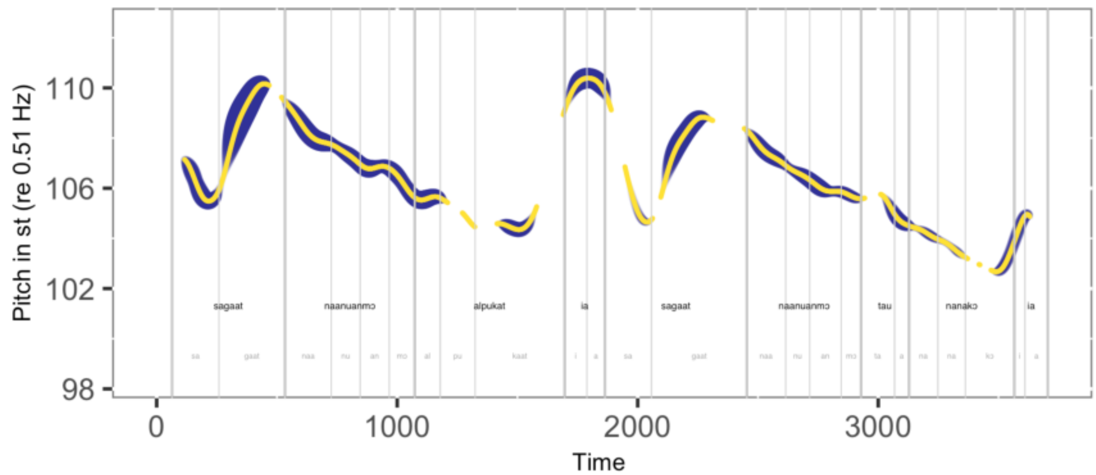

Figure 3.65: Periogram with pitch track (in st) for example (45), speaker SELP

(45) a. *a sagaat naanuanmɔ alpukaat ia*
a sɔ-gaat nɔ-anu-an=mɔ alpukaat ia
a ONE-part AV.RLS-FILL-APPL=CPL avocado PRX
'half of the avocados thingied'

b. *sagaat naanuanmɔ tau nanakɔ ia*
sɔ-gaat nɔ-anu-an=mo tau nɔn-takɔ ia
ONE-part AV.RLS-FILL-APPL=CPL person AV.RLS-steal PRX
'half of it was thingied by the thief' (pearstory_36_SELP.287) ▹

3.3.1.2.4 Others

This category includes adverbs and connectives that occur as single CIUs, as well as prepositional phrases. Additionally, it encompasses fragmentary CIUs and instances of code-switching.

Frequently, an adverb or connective itself forms a separate CIU—often a singleton IU. The most commonly used items in Totoli are *bali* 'so', *indʒan* 'then', *antuknakɔ* 'that is', *tapi* 'but', and *danna* 'then'. An example is provided in (46a).

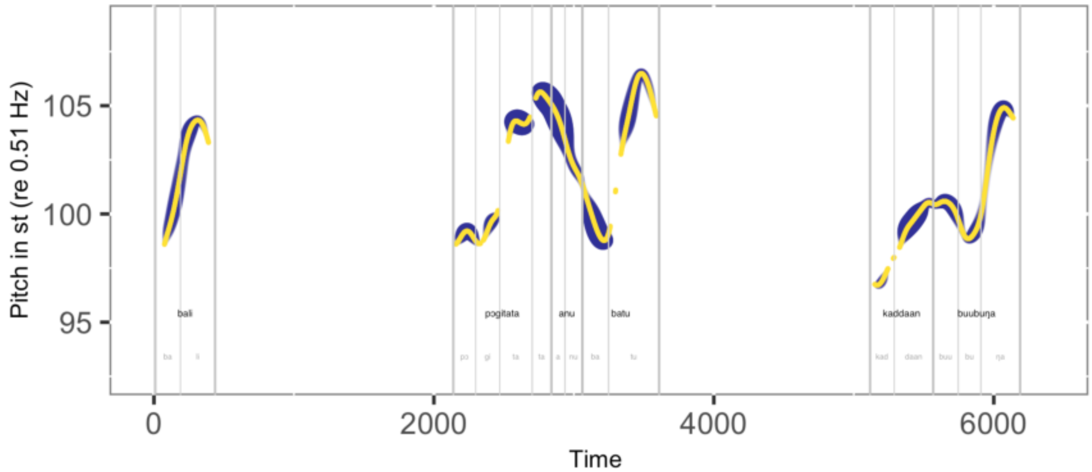

Figure 3.66: Periogram with pitch track (in st) for example (46), speaker SP

(46) a. *bali*
bali
so
'so'

b. *pɔgitata anu batu*
pɔg-ita-0=ta anu batu
SF-look.for-UV=2S.GEN REL stone
'look for a stone'

c. *kaddaan buubuŋa*
kɔ=RDP-daan RDP-buŋa
EXIST-RDP-EXIST RDP-flower
'(which) has flowers' ▷

Prepositional phrases involve a preposition and a nominal element in adverbial function forming a single CIU, as in example (47b).

(47) a. *namɔ nallakɔan baki tuku*
namɔ nɔ-RDP-lakɔ-an baki tuku
only AV.RLS-RDP-walk-APPL head knee
'only walking on knees'

b. *dɛi dalan babi*
dɛi dalan babi
LOC road pig
'on a secret path' (lifestory_RDA_1.072-74) ▷

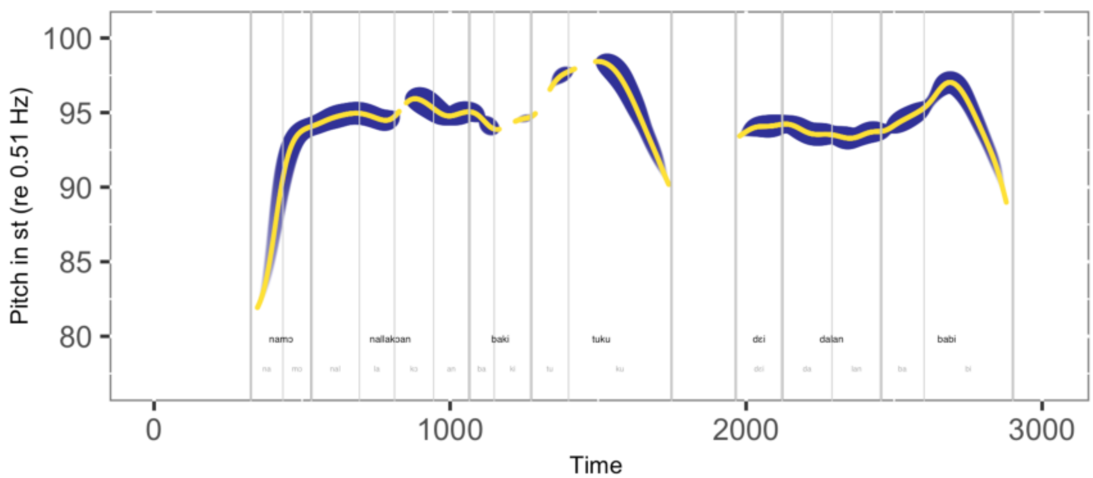

Figure 3.67: Periogram with pitch track (in st) for example (47a), speaker RDA

As per Tao (1996: 72) and Wouk (2008: 150), the CIUs grouped under "Other" include oblique arguments and adverbial adjuncts. Differentiating between these two types of CIUs in the corpus is relatively straightforward, and only a few ambiguous cases were encountered. One simple distinguishing characteristic is optionality. Quirk et al. (1985: 50) argues that while oblique arguments are usually obligatory, adverbial adjuncts

> "may be regarded, from a structural point of view, largely as 'optional extras', which may be added at will, so that it is not possible to give an exact limit to the number of adverbials a clause may contain."

Various other elements occur as separate CIUs and cannot be classified under any of the categories mentioned earlier. These are also included in the category "Other". One example is negatives, as shown in example (48a).

(48) a. *ɔ iŋga iŋga*
ɔ iŋga iŋga
O NEG NEG
'oh no no'

b. *iŋga daan kan mɔkɔdɔɔng maggalimɔ ia iŋga*
iŋga daan kan mɔkɔ-dɔɔng mɔ-RDP-gali=mɔ ia iŋga
NEG EXIST perhaps ST.AV-want AV-RDP-stop=CPL PRX NEG
'there is no longing to stop this, no'

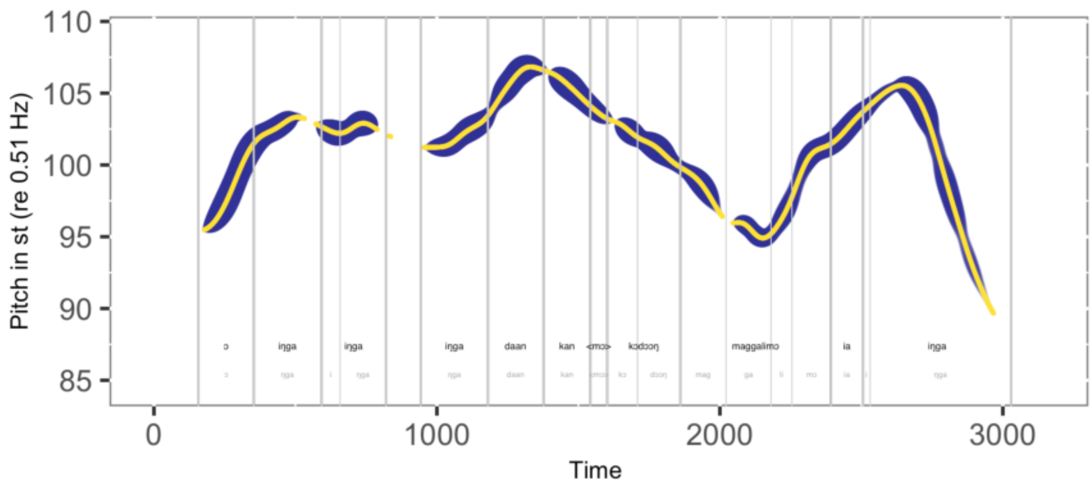

Figure 3.68: Periogram with pitch track (in st) for example (48), speaker RDA

(explanation-wedding-tradition_ZBR.341-342) ▷

Other units are numerals, shown in (49c), and quotative elements, such as in (49b).

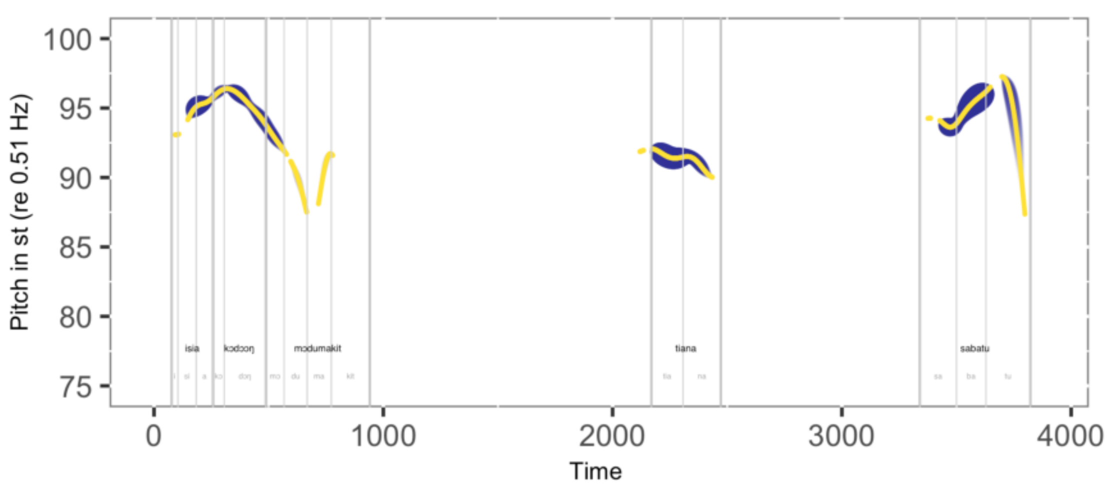

Figure 3.69: Periogram with pitch track (in st) for example (49), speaker RDA

(49) a. *isia kɔdɔɔŋ mɔdumakit*

isia	kɔ=dɔɔŋ	mɔ-d<um>akit
3s	EXIST-want	AV-<AUTO.MOT>across

‘he wants to cross’

b. *tiana*
tiŋana
QUOT
'he says'

c. *sabatu*
sabatu
sabatu
'one' (story-monkey-crocodile_RSM.030-32) ▷

3.3.1.3 Distribution and discussion

As an initial step, I examine the distribution of the four CIU types: (a) clausal CIUs, (b) nominal CIUs, (c) interactional CIUs, and (d) other minor CIU types. Figure 3.70 displays the frequency distribution of these CIU types in the corpus, showing both the overall distribution and the distribution within the conversational and monological data.

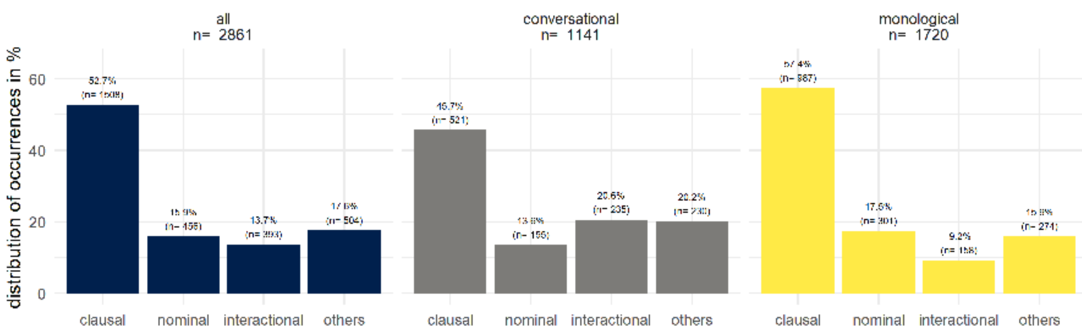

Figure 3.70: Frequency distributions of the four broad categories of CIUs within conversational and monological recordings

In the entire corpus, clausal CIUs account for 52.7%, nominal CIUs for 15.9%, and interactional CIUs for 13.7%. Other structures make up 17.6% of the corpus. The difference between conversational and monological data primarily concerns clausal and interactional CIUs. In conversational data, the proportion of clausal CIUs is 12.3 percentage points lower, while the proportion of interactional CIUs is 11.4 percentage points higher.

In Totoli, the clause is a major type of construction that constitutes a CIU. To further examine this type, I present the distribution of various types of clausal CIUs in Figure 3.71. Dependent clausal CIUs (cf. §3.3.1.2.1) are displayed in the right-hand columns. Independent clausal CIUs (cf. §3.3.1.2.1) are further subdivided into simple clausal CIUs (with zero, one, or two overtly expressed arguments), and complex clausal CIUs (involving more than one predicate and/or more than two arguments).

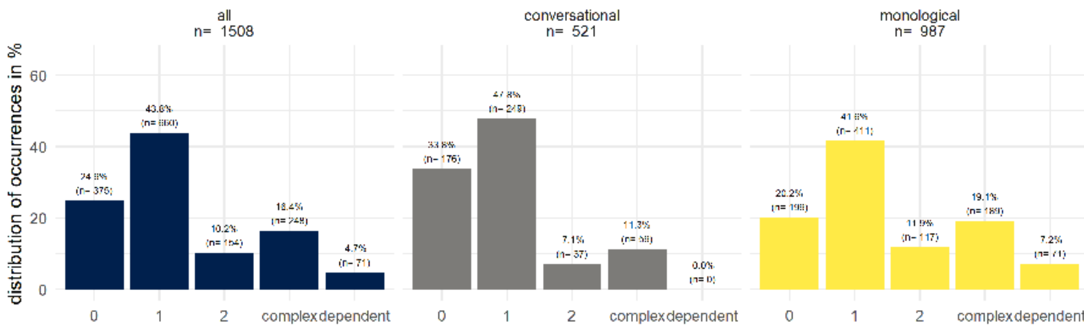

Figure 3.71: Frequency distributions of subcategories of clausal CIUs within conversational and monological recordings

In both conversational and monological data, there is a notable proportion of clausal CIUs with one overtly expressed argument (47.8% in conversational data, 41.6% in monological data), as well as a high proportion of elliptical CIUs, with no overtly expressed verb (33.8% in conversational data, 20.2% in monological data). It is also noteworthy that there are no CIUs containing a dependent clause in conversational data.

Figure 3.72 offers a detailed breakdown of the various types of nominal CIUs. This includes nominal argument CIUs (cf. §3.3.1.2.2), parallel nominal CIUs (cf. §3.3.1.2.2), and independent nominal CIUs (cf. §3.3.1.2.2).

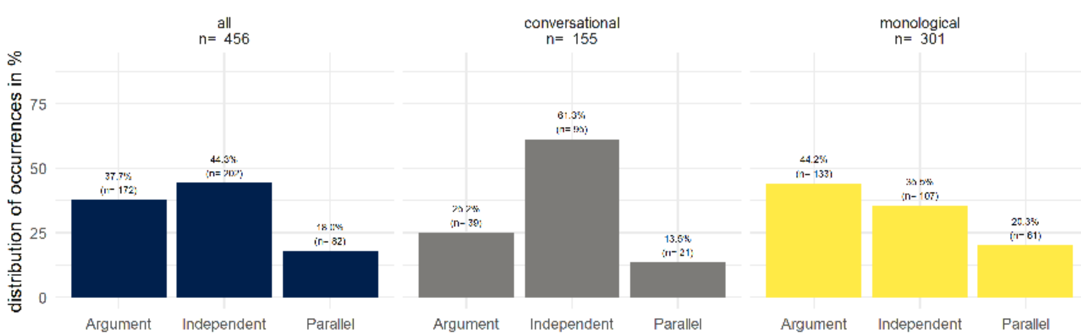

Figure 3.72: Distributions of argument, independent and parallel nominal CIUs within conversational and monological recordings; numbers are rounded to one decimal place.

The data show a high number of independent nominal CIUs and substantial differences between conversational and monological data: 61% independent nominal CIUs in conversations and 35.5% in monological data.

Table 3.5 summarizes the results.

Table 3.5: Summary of distributions of different CIU types and subcategories: total proportions are given in the left-hand columns, total numbers in the right-hand columns.

	all		conversational		monological	
clausal	52.7%	1508	45.7%	520	57.4%	987
0	13.1%	375	15.4%	176	11.6%	199
1	23.1%	660	21.8%	249	243.9%	411
2	5.4%	154	3.2%	37	6.8%	117
complex	8.7%	248	5.2%	59	11.0%	189
dependent	2.5%	71	0.0%	0	4.31%	71
nominal	15.9%	456	13.6%	155	17.5%	301
argument	6.0%	172	3.4%	39	7.7%	133
independent	7.1%	202	8.3%	95	6.2%	107
parallel	2.9%	82	1.8%	21	3.5%	61
interactional	13.7%	393	20.6%	235	9.2%	158
others	17.6%	4504	20.2%	230	15.9%	274
adv. & con.	3.4%	96	2.4%	27	4.0%	69
code-switching	1.8%	52	1.5%	17	2.0%	35
fragments	2.1%	61	3.6%	41	1.2%	20
further	3.8%	108	6.4%	73	2.0%	35
PP	3.1%	89	3.2%	37	3.0%	52
uncodable	3.4%	98	3.1%	35	3.7%	63
TOTAL	100%	2861	100%	1141	100%	1720

It is important to consider that Totoli is an endangered, understudied language. While Riesberg (2014), Himmelmann & Riesberg (2013) and Riesberg et al. (2021) provide detailed discussions of the major aspects of the verbal morphology, other aspects of its grammar are still not fully worked out. In some cases, I had to make coding decisions that readers may or may not agree with. One example relevant to the current discussion is the coding of negation with a form of the negated existential predicate *ko=/daan/kaddaan*. This form of negation occurs frequently in both Totoli and the local (Manado) Malay variety. A simple clause that is negated with an existential predicate will appear as a complex clause in the count, as it involves two predicates: the existential predicate and the predicate of the negated clause. Such decisions must be made for each language based on its unique grammar, and need to be considered when comparing reported data.

The investigation presented above aimed to answer the question of what grammatical structures a CIU typically contains. I will now turn to a discussion of grammatical structures found in embedded IUs of CIUs.

3.3.2 Syntactic structures of embedded IUs of CIUs

In this section, I will be discussing several aspects related to the grammatical units found in embedded IUs of CIUs. The analysis is straightforward and is not couched in the framework and coding conventions used above in §3.3.1. Here, I use a limited range of grammatical unit types which are sufficient to describe the majority of structures found, such as noun phrases, verbs, prepositional phrases, adverbial clauses, and relative clauses. To briefly illustrate some of the main aspects, I have provided illustrative examples below.

Noun phrases typically constitute their own (embedded) IUs, and for NPs in preverbal position, this is observed with consistency. An example is provided in (50), with the periogram shown in Figure 3.73. The simple one-word argument NP ‘*sagaat* meaning ‘the half’ is parsed into a separate embedded IU, as indicated by the pitch rise on the final syllable. It is followed by an IU containing the verb and a following prepositional phrase. The “|” symbol indicates a boundary of an embedded IU.

(50) *sagaat | madabumai dɛi buta*
sɔ-gaat mɔ-dabu=mɔ=ai dɛi buta
ONE-part ST-fall=CPL=VEN LOC earth
‘half of it fell to the ground’ (pearstory_14_SP.007) ▹

Headless relative clauses are consistently parsed as their own IUs when in preverbal position. For example, consider the CIU in example (51) and its corresponding pitch contour shown in Figure 3.74. The initial IU of the CIU comprises the headless relative clause *anu ampi kɔlɔanan* meaning ‘the one on the right side’. This is then followed by an IU containing both the verb *saasalu* ‘facing’ and the pronominal NP *kita* ‘us’.

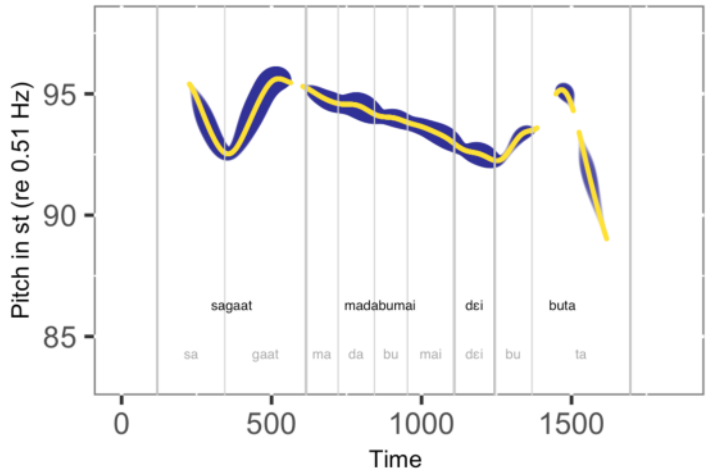

Figure 3.73: Periogram with pitch track (in st) for example (50), a CIU consisting of two embedded IUs, speaker SP

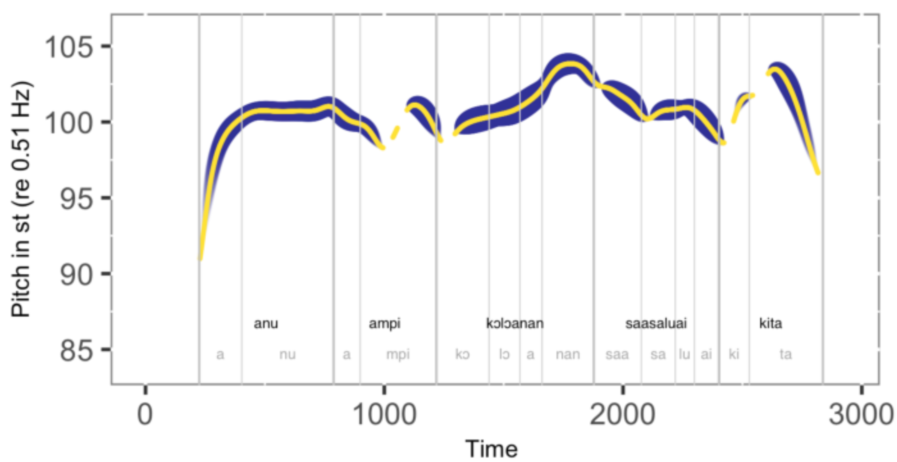

Figure 3.74: Periogram with pitch track (in st) for example (51), a CIU consisting of two embedded IUs, speaker SP

(51) *anu ampil kɔlɔanan | saasaluai kita*
anu ampi kɔlɔanan RDP-salu=ai kita
REL part right RDP-facing=VEN 2S
'the one on the right-hand side is facing you'

(spacegames_sequence4_KSR-SP.231) ▹

In examples (50) and (51) above, the verb is grouped together in one IU with the following constituent, such as the prepositional phrase in (50) and pronominal argument NP in (51). Such cases are common. However, in many instances, the verb and possible following adverbs are grouped as separate IUs, as seen in example (52), where the verb *bagulna* meaning 'he was beating' and the following adverb *pɔni* 'again' constitute the first IU of the CIU and are grouped separately from the following argument NP *kalibɔmbaŋ* meaning 'butterfly'.

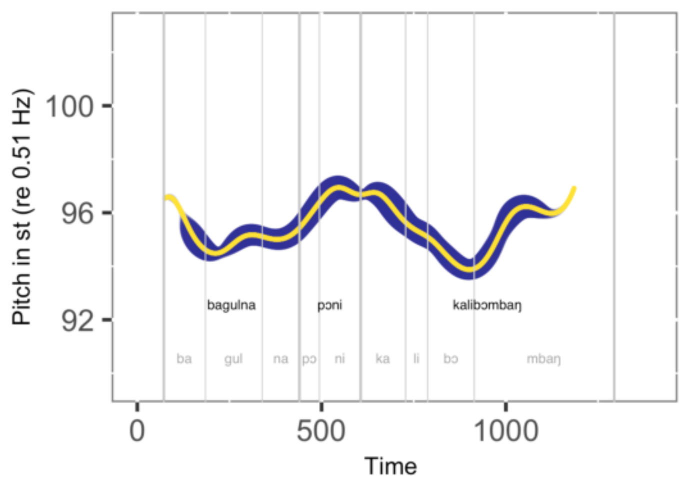

Figure 3.75: Periogram with pitch track (in st) for example (52), a CIU consisting of two embedded IUs, speaker RSM

(52) *bagulna pɔni | kalibɔmbaŋ*
bagul-0=na pɔni kalibɔmbaŋ
beat-UV=3S.GEN again butterfly
'he was beating the butterfly again' (story-monkey-butterfly_RSM.053) ▹

In fact, the same construction can be found with both realizations: with the verb and the following argument parsed together in one IU or separately in one IU each. Below are two instances of a nearly identical CIU which involves a verb and its oblique argument. In the first example (53) and its corresponding visualization in Figure 3.76, the verb and the oblique argument form separate, embedded IUs. This is clearly visible by the pitch rise on the last syllable of the verb *nɔlitaan* meaning ‘to meet’.

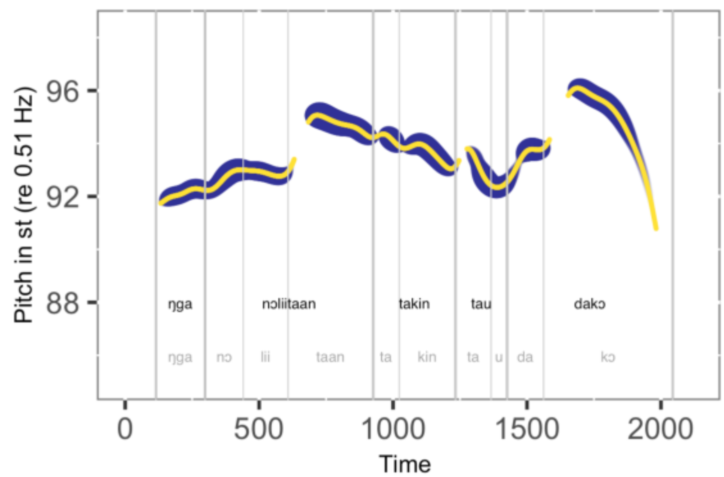

Figure 3.76: Periogram with pitch track (in st) for example (53), a CIU consisting of two embedded IUs, speaker RDA

(53) *iŋga nɔliitaan | takin tau dakɔ*
iŋga nɔli-ita-an takin tau dakɔ
NEG RCP.RLS-see-RCP.RLS with person big
‘(I) didn’t meet (my) parents’ (lifestory_RDA_1.124) ▷

Example (54) features an almost identical construction. However, in this case, the verb and its oblique argument constitute a single (embedded) IU together. There is no pitch rise observed on the last syllable of the verb *nɔlitaan* meaning ‘to meet’, which would typically indicate an IU boundary. The corresponding periogram with pitch track (in st) is provided in Figure 3.76.

(54) *danna | nɔliitaan takin tau dakɔ*
danna nɔli-ita-an takin tau dakɔ
then RCP.RLS-see-RCP.RLS with person big
‘Then, (I) met (my) parents’ (lifestory_RDA_1.115) ▷

Other instances involve an NP with a modifying relative clause that can either occur together in one IU or split into two IUs. The latter is more common when in postverbal position. For instance, in example (55) and its visualization

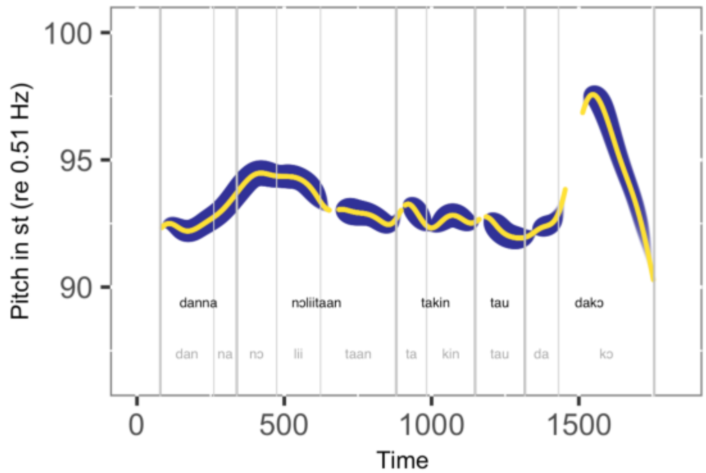

Figure 3.77: Periogram with pitch track (in st) for example (54), a CIU consisting of two embedded IUs, speaker RDA

in Figure 3.78, the first word *lau* 'currently' appears in a separate, embedded IU at the beginning of the CIU. It is followed by an IU containing the relative clause *anu lau suludan tau mɔanɛ ana* 'which is being pushed by the man', another IU with the prepositional phrase *dɛi dulak ɔgɔbbunna* 'in front of the well', and the final IU containing the verb *mɔitaku* 'I see'. The IU containing the relative clause spans six words or twelve syllables, and the IU with the prepositional phrase has four words or seven syllables. Such lengthy IUs are not uncommon in Totoli, as demonstrated by the examples in §3.2 (e.g. example (15) in Figure 3.29, and example (1) in Figure 3.8). This indicates that the (embedded) IU in Totoli differs significantly from the prosodic word and Accentual Phrase in Korean, as discussed in §3.2.5.

(55) *lau | anu lau suludan tau mɔanɛ ana | dɛi dulak ɔgɔbbunna | mɔitaku*
lau anu lau sulud-an tau ana dɛi dulak ɔgɔbbun=na
presently REL presently push-APPL person MED LOC front well=3S.GEN
mɔ-ita-0=ku
ST-see-UV=1S.GEN
'what is currently pushed by the man in front of the well, as I see it.'

(QUIS-focus_SP.026) ▹

Example (56) and its visualization in Figure 3.79 illustrate a long and complex CIU, where the various grammatical units are very regularly chunked into (embedded) IUs. The CIU commences with an IU containing the connective adverb

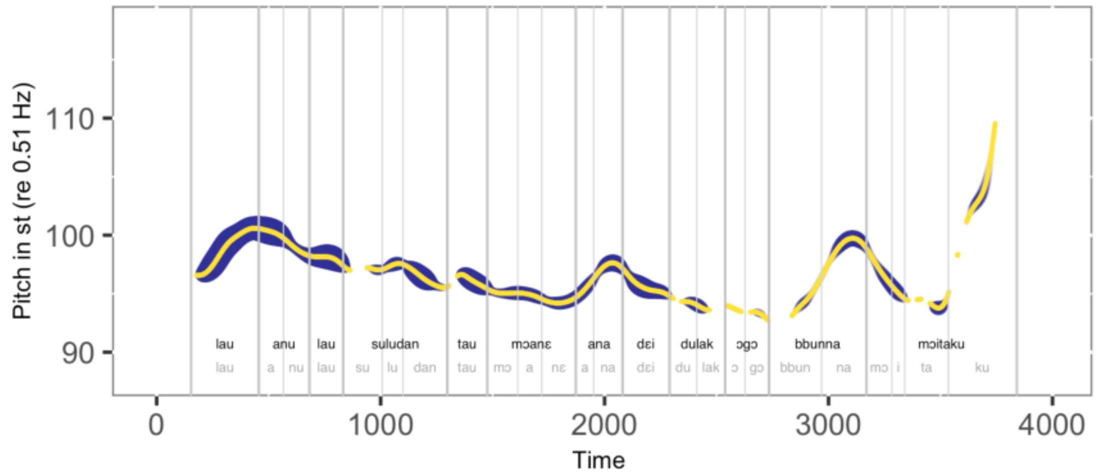

Figure 3.78: Periogram with pitch track (in st) for example (55), a CIU consisting of four embedded IUs, speaker SP

bali 'then/so', which typically constitutes its own IU. The following IU contains the subject NP *tau* 'person' along with its set of modifiers. This is succeeded by an IU containing the verb *nanaumai* 'to go down'. The subsequent IU contains a prepositional phrase that is repeated twice. The first instance involves the dummy/filler element *anuna*, and the second instance involves the intended/repaired prepositional phrase *ulai puun alpukaat* 'from the avocado tree'. The CIU concludes with an IU containing a relative clause that further modifies the noun in the prepositional phrase.

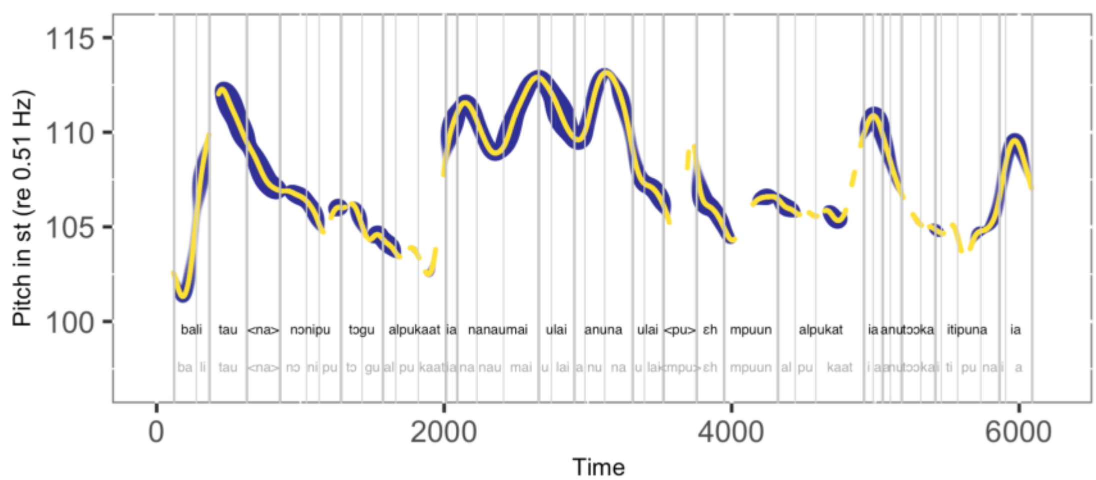

Figure 3.79: Periogram with pitch track (in st) for example (56), a CIU consisting of six embedded IUs, speaker SELP

(56) *bali | tau <na> nɔnipu tɔgu alpukaat ia | nanaumai | ulai anuna | ulai <p> ɛ puun alpukaat ia | anu tɔɔka itipuna*

bali tau nɔN-tipu tɔgu alpukaat ia
so person AV.RLS-pick possession avocado PRX

nɔ-nau=mɔ=ai uli=ai anu=na uli=ai ɛ puun
AV.RLS-go.down=CPL=VEN from=VEN FILL=3S.GEN from=VEN INTJ tree

alpukaat ia anu tɔɔka ni-tipu-0=na
avocado PRX REL finished RLS-pick-UV=3S.GEN

'so, the person picking, the owner of the avocados, goes down from the avocado tree that he was just picking.' (pearstory_36_SELP.047) ▹

In §3.2.5, I discussed the case of Heads in a THL, which are immediately followed by additional syntactic material. In this case, the Head of the THL construction constitutes an IU in CIU-initial position. Adverbial and relative clauses that fulfill this role can occasionally be quite long, consisting of five or more words. Two examples are given below in (57)–(58). Example (57) and its visualization in Figure 3.80 is a CIU with an extensive initial adverbial clause, phrased in two IUs. The adverb *laalau* 'currently' is parsed in a single IU with the remainder of the adverbial clause parsed in a single long IU, containing 6 words / 13 syllables.

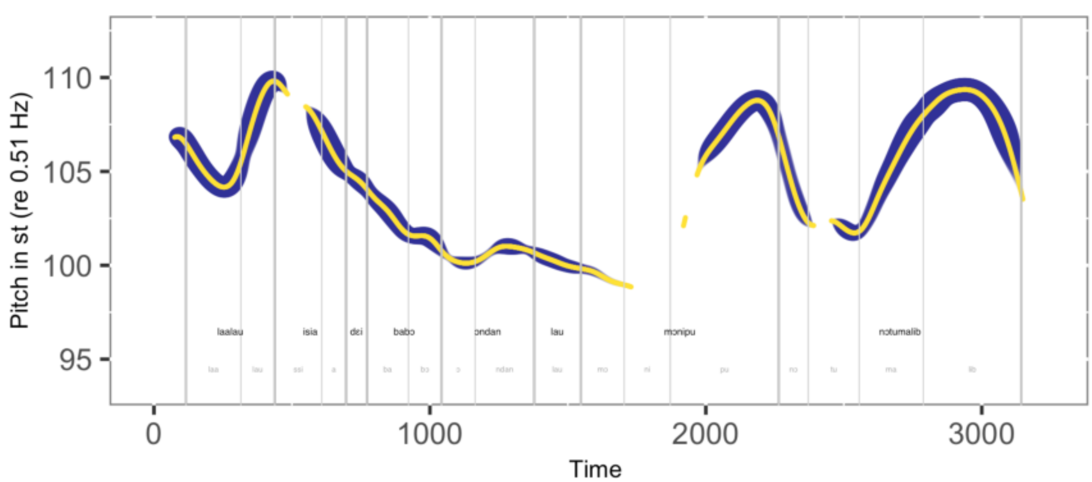

Figure 3.80: Periogram with pitch track (in st) for example (57), a CIU consisting of three embedded IUs, speaker SELP

(57) *lalau | isia dɛi babo ɔndan lau mɔnipu | nɔtumalib*

RDP-lau isia dɛi babo ondan lau mɔ-tipu
RDP-while 3S LOC above ladder while AV.RLS-pick

nɔ-t<um>alib
AV.RLS-<AUTO.MOT>pass.by
‘While he was on the ladder picking (pears),(he) passed by’ ▷

Example (58) and its realization in Figure 3.81 is an instance of a CIU with two initial IUs containing a complex adverbial clause.

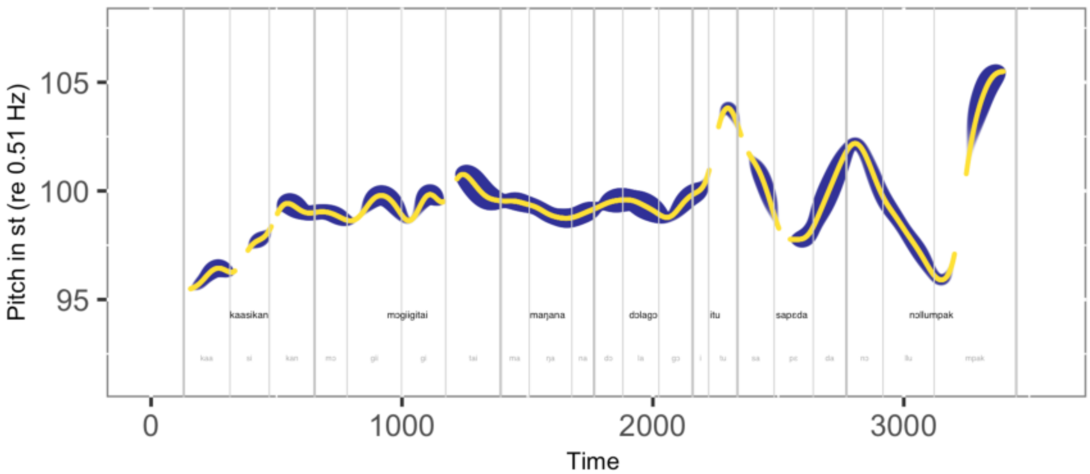

Figure 3.81: Periogram with pitch track (in st) for example (58), a CIU consisting of three embedded IUs, speaker SP

(58) *kaasikan | mɔgiigitai maŋana dɔlagɔ itu | sapɛda | nɔllumpak*

kɛasikan mɔg-RDP-ita-i maŋana dɔlagɔ itu sapɛda
excitement AV.NRLS-RDP-watch-APPL child girl DIST bicycle

nɔ-RDP-lumpak
ST-RDP-hit.against

'because of his excitement in looking at the girl, his bicycle crashed (against the stone)' (pearstory_11_SP.025) ▹

The examples demonstrate regularities in chunking. However, more insightful are instances that appear to contradict these regularities.

In many CIUs, the first word forms its own IU, which is also noted by Himmelmann (2018: 361). In fact, 31% of all CIUs in the corpus containing more than one word have their first word phrased as a separate embedded IU. This is often because words in the initial position of a CIU are connectives or one-word noun phrases that are regularly phrased as a single IU. See examples (50), (55), (56).

As explained above, adverbial clauses, relative clauses, and prepositional phrases are typically not further chunked, regardless of their length, as demonstrated by the long embedded IUs in examples (57) and (58). However, there are instances where the initial elements constitute a separate IU. Examples include the initial relative particle *anu* of a relative clause, the initial preposition of a prepositional phrase, and the initial conjunction of an adverbial clause. Consider (57) above, with its periogram in Figure 3.80. The initial word *laalau* 'while' of the adverbial clause is phrased as its own IU. The following IU contains a pronoun as well as a prepositional phrase and a predicate. These are not further chunked, as is the case for adverbial clauses (see also example (55)).

In many adverbial clauses, such as examples (57) and (58), the first element is phrased separately. When an initial element of a relative clause, adverbial clause, or prepositional phrase is phrased separately, it often co-occurs with a hesitation pause. Thus, boundary placement appears to serve as a planning device. Two examples are provided below.

The CIU in example (59) contains a verb and a lengthy prepositional phrase. The initial verb is phrased as a separate IU. However, in the following prepositional phrase, the initial preposition *dɛi* is phrased separately from the rest of the CIU and is followed by a short CIU-internal pause. The periograms with pitch tracks (in st) are shown in Figure 3.82.

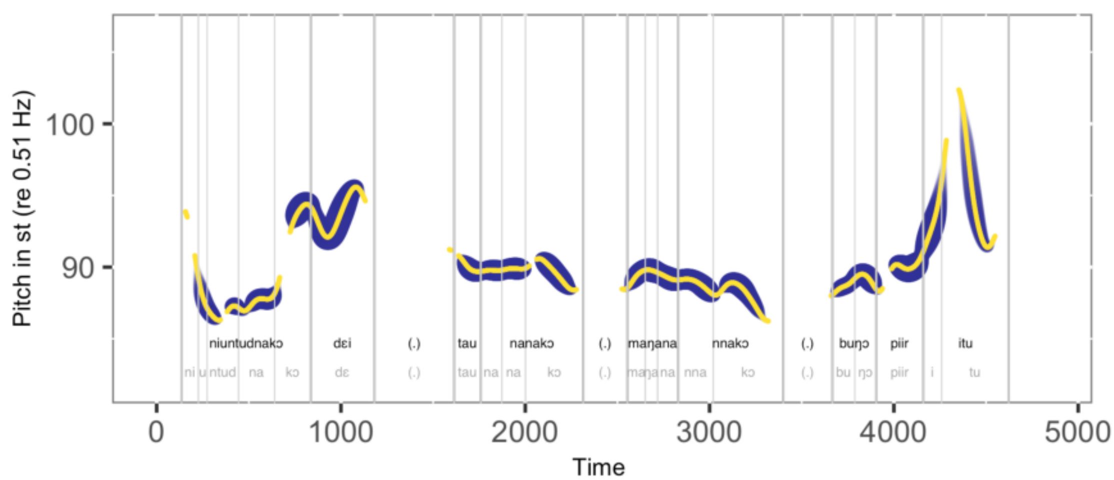

Figure 3.82: Periogram with pitch track (in st) for example (59), a CIU consisting of three embedded IUs, speaker FAH

(59) *niuntudnakɔ | dɛi | tau nanakɔ maŋana nnakɔ buŋɔ piir itu*
ni-untud-0=na=kɔ dɛi tau nɔN-takɔ maŋana nɔN-takɔ
RLS-bring-UV=3S.GEN=AND LOC person AV.RLS-steal child AV.RLS-steal
buŋɔ piir itu
fruit pear DIST
'he brought (it) to the person who stole, the child who stole the pears'

(pearstory_11_SP.025) ▷

Another example is provided in (60), and its visualization is shown in Figure 3.83. In this example, the initial preposition *untuk* 'for' is also phrased separately, followed by a CIU-internal hesitation pause.

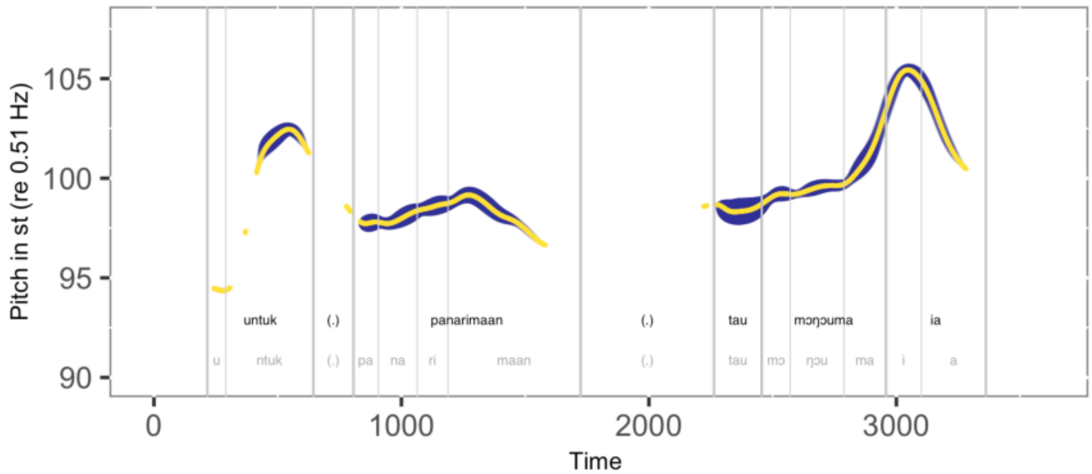

Figure 3.83: Periogram with pitch track (in st) for example (60), a CIU consisting of two embedded IUs, speaker SP

(60) *untuk | panarimaan tau mɔngɔuma ia*

untuk	pɔN-tarima-an	tau	mɔ-ngɔ-uma	ia
for	NMLZ-accept-NMLZ	person	ST-COLL-arrived	PRX

'for the reception of the visitors' (pearstory_11_SP.025) ▷

Instances of an embedded IU containing elements of two separate clauses are very rare. One such instance is presented in example (61), which is visualized in Figure 3.83. This example comprises an adverbial clause and two coordinated main clauses and involves an embedded IU containing grammatical units of both main clauses. The CIU starts with an IU containing the adverbial clause *daan nadabu* 'after he fell'. The following verb *nɔlimulas* 'scatters' is uttered as its own embedded IU. The subsequent argument *alupkaat* 'the avocado' is not parsed into its own IU, but the pitch drops continuously despite the clause boundary. An IU-final boundary tone is placed on the adverb *gaakɛ* 'also'. As a result, this IU contains the argument NP of the first clause and the verb of the second clause. The argument NP of the second clause *buludna* 'his shinbone' forms the final IU of the CIU.

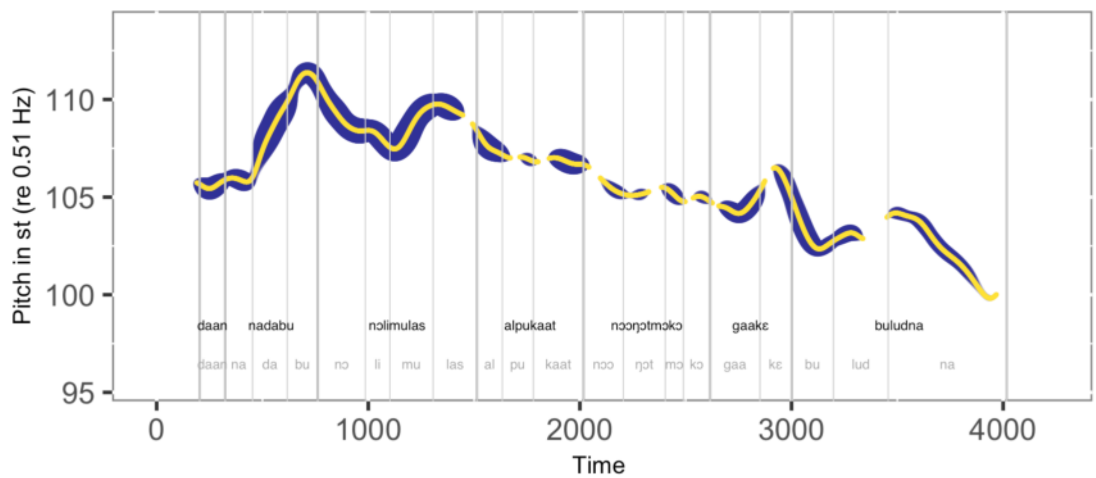

Figure 3.84: Periogram with pitch track (in st) for example (61), a CIU consisting of four embedded IUs, speaker SELP

(61) *daan nadabu | nɔlimulas | alpukaat nɔɔngɔtmɔkɔ gaakɛ | buludna*

daan nɔ-dabu nɔ-l<um>ɛlas alpukaat
later ST.RLS-fall AV.RLS-<AUTO.MOT>scatter avocado

nɔ-ɔngɔt=mɔ=kɔ gaakɛ bulud=na
ST.RLS-hurt=CPL=AND too shin=3S.GEN

'after (he) fell, the avocados scatter and his shin hurts too'

(pearstory_36_SELP.025) ▷

In this section, I have illustrated some key aspects regarding the chunking of CIUs into IUs and their regularities. The syntactic content of IUs typically constitutes a complete grammatical unit, but there are rare instances where two independent grammatical units occur within the same IU. Additionally, certain units such as adverbial clauses, prepositional phrases, and relative clauses may have their initial element phrased as a separate IU. The realization of example (61) provides an interesting case in point. In most cases, speakers consistently mark the right-edge boundary of a grammatical unit with a prosodic boundary, but boundary placement is optional. In this example, the speaker did not place a boundary at the end of the NP *alpukaat* 'avocado', which is the end of the first main clause. Consequently, the IU consists of two grammatical units that belong to two different clauses.

3.3.3 Comparing the syntax and prosody of embedded IUs of CIUs with singleton IUs

In this section, I compare the syntactic content of embedded IUs of CIUs with singleton IUs, investigating the differences or similarities between the two with regard to the grammatical structures they contain. The analysis of phrase-final tonal patterns in §3.2.1 and §3.2.2 revealed that prosodic patterns at the end of embedded IUs are essentially the same as those that occur in CIU-final position, providing evidence that these prosodic units are essentially of the same type.

To further explore this assumption, I compare grammatical structures typically found in embedded IUs of complex CIUs with those found in singleton IUs. I illustrate this with two examples below.

Compare example (62) and its visualization in Figure 3.85 with example (63) and its visualization in Figure 3.86. Both contain a similar structure: the connective *bali* 'so', followed by the verb *pɔgitata* 'look for' and a subsequent headless relative clause in undergoer function. The difference is that in example (62), the connective is parsed into one CIU together with the main clause and constitutes the initial embedded IU of the CIU. In example (63), however, the connective appears in a separate IU, clearly demarcated by further boundary phenomena, such as pitch reset, and final syllable lengthening. Note, however, that tonal targets and the tonal contours of *bali* 'so' are identical in both instances.

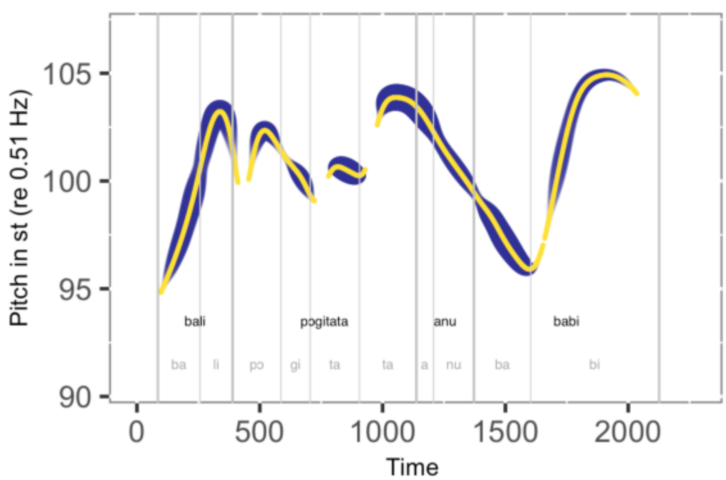

Figure 3.85: Periogram with pitch track (in st) for example (62), a CIU consisting of three embedded IUs, speaker SP

(62) *bali | pɔgitata | anu babi*
bali pɔg-ita-0=ta anu babi
so sf-look.for-uv=2s.gen rel pig
'so, look for the pig.' (spacegames_sequence4_KSR-SP.012) ▷

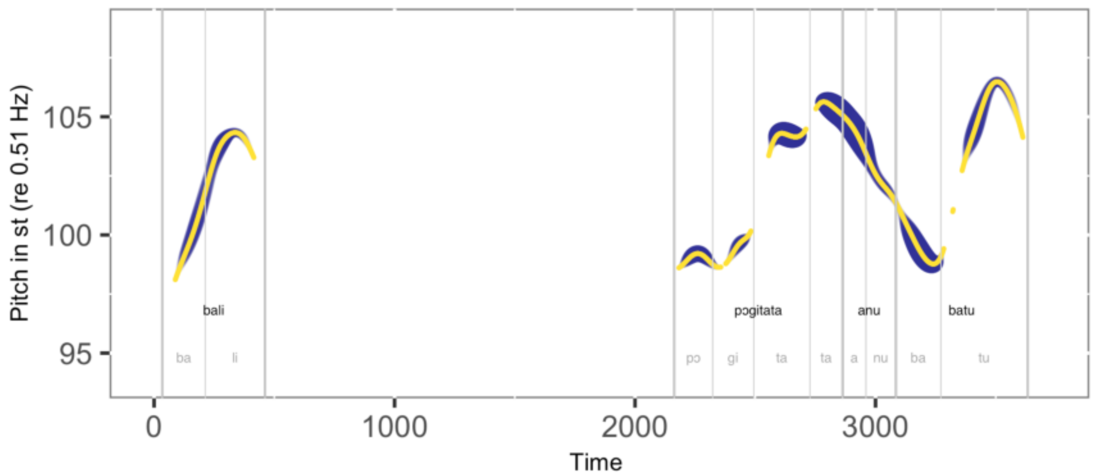

Figure 3.86: Periogram with pitch track (in st) for example (63), two CIUs, speaker SP

(63) a. *bali*
bali
so
'so'

b. *pɔgitata | anu batu*
pɔg-ita-0=ta anu batu
sf-look.for-uv=2s.gen rel stone
'look for the stone.' (spacegames_sequence4_KSR-SP.012) ▷

In both examples above, the tonal targets of the connective *bali* are the same. In example (62), pitch is interpolated between the IU-final H% boundary tone located on the last syllable of the connective *bali* 'so' and the tonal targets of the second IU *pɔgitata* 'look for'. In example (63), the connective forms its own singleton IU and pitch is reset at the beginning of the then CIU-initial word *pɔgitata* 'look for'.

A second example is given in (64) and (65). Again, in (64), the left-dislocated topic NP *ɔtɔ* 'car' is parsed into one CIU with the following clause. In (65), the same NP *ɔtɔ* ‘car’ occurs as a separate singleton IU, clearly visible by the reset in pitch and final syllable lengthening. Yet, tonal targets and the shape of the pitch contour of the NP *ɔtɔ* ‘car’ are the same in both instances.

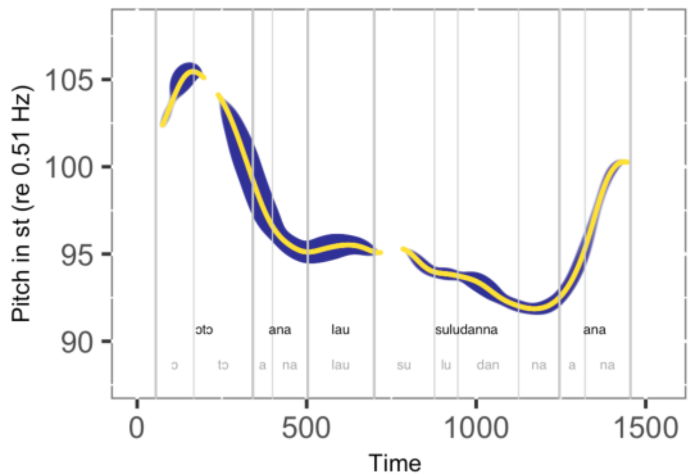

Figure 3.87: Periogram with pitch track (in st) for example (64), a CIU consisting of two embedded IUs, speaker SP

(64) *ɔtɔ | ana lau suludanna ana*
ɔtɔ ana lau sulud-an=na ana
car MED presently push-APPL=3S.GEN MED
‘it is a car that is being pushed by him’ (QUIS-focus_SP.012) ▹

(65) a. *ɔtɔ*
ɔtɔ
car
‘(it is a) car’

b. *anu laalau suludan tau mɔanɛ dɛi dulak ɔgɔbbun ana*
anu RDP-lau sulud-an tau mɔanɛ dɛi dulak ɔgɔbbun ana
REL RDP-presently push-APPL person man LOC front well MED
‘that is currently pushed by the man in front of the well’
(QUIS-focus_SP.041-42) ▹

These two example pairs illustrate that the same grammatical structures may occur as either singleton IUs or an embedded IU in a complex CIU. In the examples above, the tonal specifications remain the same.

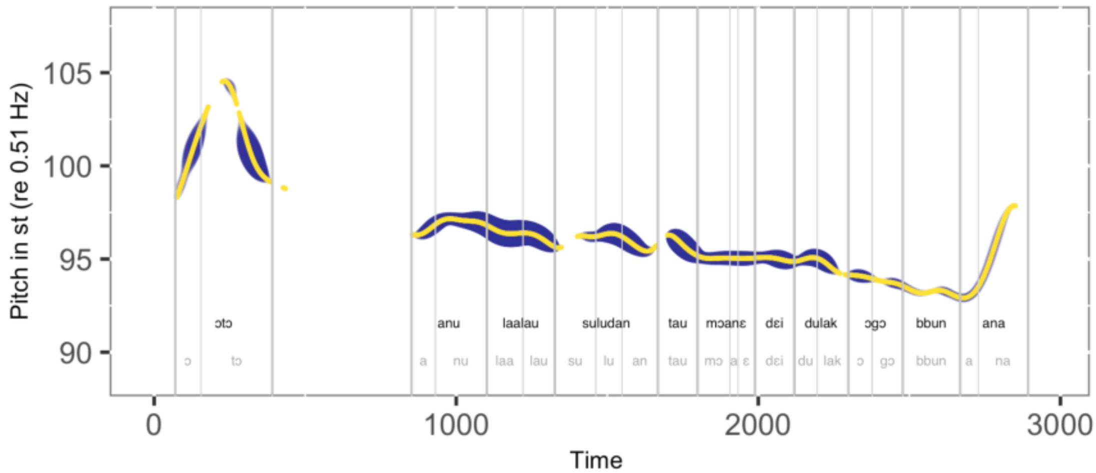

Figure 3.88: Periogram with pitch track (in st) for example (65), speaker SP

Based on the corpus, I conducted a quantitative comparison of grammatical structures found in embedded IUs of CIUs with those found in singleton IUs. Figure 3.89 illustrates the comparison. Embedded IUs of complex CIUs (exemplified here as 'y_1', 'y_2', 'y_3') are given on the right-hand side and singleton IUs which consist of a single IU (exemplified here as 'x') are displayed on the left-hand side.

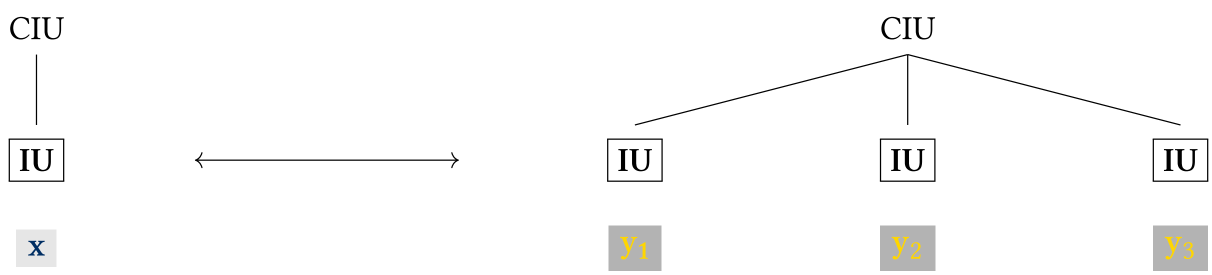

Figure 3.89: Comparison of syntactic content of embedded IUs and singleton IUs

Figure 3.90 compares the distribution of grammatical units found in singleton IUs with the distribution within embedded IUs of CIUs. The seven syntactic categories account for 82% of the 3191 embedded IUs and 78% of all 1005 unchunked IUs in the corpus.

The frequency distribution in Figure 3.90 shows that all seven major grammatical units found in embedded IUs ('y_1', 'y_2', 'y_3') also occur as unchunked singleton IUs ('x'), although with varying distribution. Specifically, 82% of all embedded IUs correspond to one of the seven categories of grammatical units,

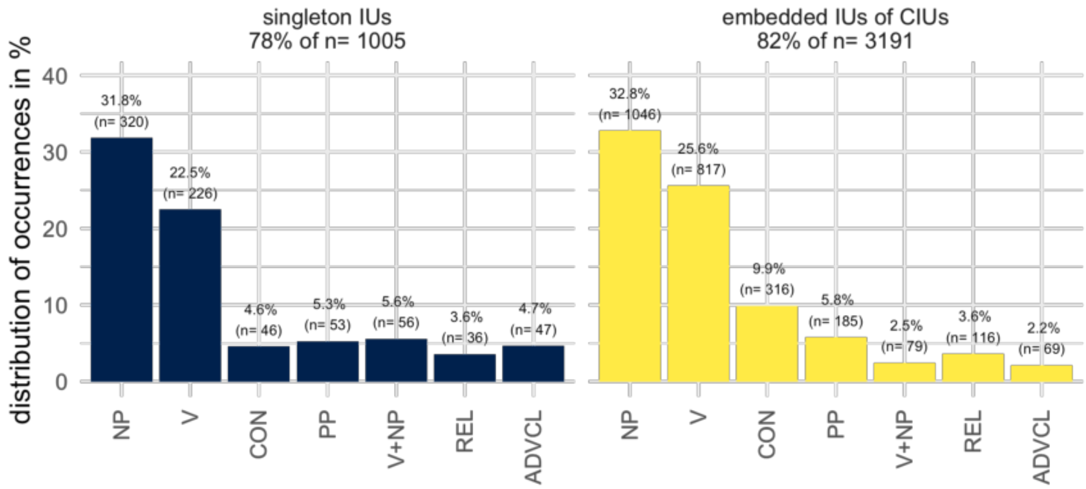

Figure 3.90: Frequency distribution of the 7 major grammatical structures of non-final IUs, within unchunked IUs: numbers are rounded to one decimal place.

and these seven grammatical units also describe 78% of all singleton IUs. The difference in distribution mainly pertains to verbs and connectives. Verbs occur less frequently as singleton IUs than as embedded IUs; 25.6% of embedded IUs are verbs, but only 22.5% of singleton IUs are verbs. Connectives, on the other hand, occur considerably less frequently as singleton IUs.

Note that the absolute numbers of IUs and embedded IUs consisting of an NP are lower here than the number of nominal IUs in Table 3.72 above. This is because the counts here are more conservative than above. Nominal IUs above include cases where the nominal element co-occurs with a connective or a relative clause, while the counts here consider only those singleton or embedded IUs that consist of a single NP only.

In summary, the results show that the distribution of grammatical units typically found in non-final IUs and singleton IUs is very similar. Elements that regularly constitute a separate IU in a complex CIU are also found as unchunked singleton IUs.

3.3.4 Discussion

In §3.3.1 above, I explored the grammatical units typically found in a CIU. I showed that 52.7% of CIUs in the corpus are clausal, 15.9% are nominal and 13.7% are interactional. Crucially, I found that proportions vary between conversational and monological data. In §3.3.2, I briefly illustrated some of the major aspects with regard to CIU-internal chunking and the major types of grammatical units found in embedded IUs of CIUs. Finally, in §3.3.3, I compared grammatical structures found in singleton IUs with those found in embedded IUs. I showed that the same grammatical units found in embedded IUs of complex CIUs also frequently occur as unchunked singleton IUs. Hence, neither the syntactic content, nor the tonal markings of singleton IUs and embedded IUs of CIUs differ. If they were of a different category, I would expect the units to differ also with regard to their syntactic content. This is clearly the case in the analysis of e.g. French and Korean intonation (Jun & Fougeron 2000, Jun 2005a) for which both the level of the IU and the lower-level Accentual Phrase are assumed, of which the latter usually only contains one word. In light of these results, I conclude that there is clear evidence that we have to assume recursive embedding of IUs into complex CIUs in order to describe the data of Totoli as proposed in the model presented in §3.2.

To conclude this discussion, I will briefly address the distribution of IU-final boundary-tone complexes and review some of the factors that may influence the choice of either complex. Consider Figure 3.91, which provides examples of adverbial clauses, noun phrases, and verbs, and illustrates their occurrence with one of the boundary-tone complexes. These three constituents are significant syntactic components, and I will use them to compare the factors affecting the choice of a boundary tone. The right-hand columns in each pair indicate the distribution of complex CIUs in embedded IUs, while the left-hand columns indicate their distribution in simple, singleton IUs.

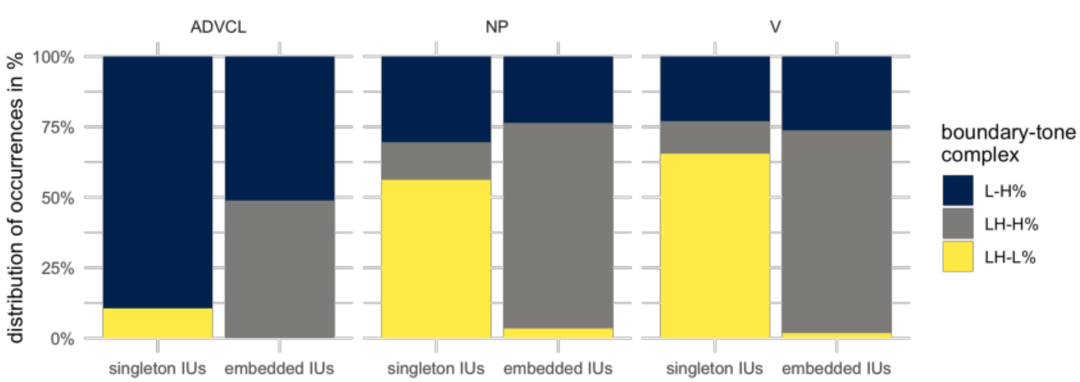

Figure 3.91: Distribution of the three boundary-tone complexes in singleton IUs and embedded IUs of CIUs, exemplified with the three grammatical categories AdvCl, NP, and VP

The figure illustrates that when adverbial clauses occur as singleton IUs, they usually take the L-H% boundary-tone complex. However, when embedded in a compound IU, many adverbial clauses take the LH-H% pattern. This preference also applies to adverbial clauses that are embedded in CIUs, although many also occur with the LH-H% pattern in this position. Noun phrases and verbs occurring as singleton IUs show a strong preference for the LH-L% pattern, although other boundary-tone complexes are possible. When occurring as embedded IUs of a complex CIU, they tend towards the LH-H% pattern. The tendency for a rise pattern in embedded IUs of complex CIUs is not surprising, as these IUs are part of a larger unit, and the final rise pattern indicates non-finality within the Compound IU.

The question then arises as to why there are two different rise patterns, i.e. L-H% and LH-H%. One possible explanation is that the difference between the three patterns LH-L%, L-H%, and LH-H% correlates with degrees of integration. The LH-H% pattern represents "full integration" and is regularly used with verbs and noun phrases in non-final positions of complex IUs. The L-H% pattern and the LH-L% pattern would then be reserved for clause-external constituents. Adverbial clauses may occur with the LH-H% full-integration pattern; however, about half of them occur with the L-H% pattern.

Above, I showed an example of a left-dislocated topic, which involved the LH-L% pattern (cf. example (64)). The dislocated element usually involves the LH-L% pattern. Other instances involve right-dislocation, afterthoughts or appositive NPs. Example (66) shows such an instance. The NP *siritana ia* 'this story' is taken up again after the verb *maaliŋ* 'to get lost', first by the pronoun *ia* 'PRX' and then by the appositive NP *baraŋ ia* 'this thing'. The verb bears the LH-L% boundary-tone complex and is immediately followed by the pronoun *ia* which takes up the NP *siritana*.

In the preceding section, I presented an example of a focused constituent, which involved the LH-L% pattern (cf. example (64)). The focused element typically involves the LH-L% pattern, while other instances involve right-dislocation, afterthoughts, or appositive NPs. Example (66) illustrates such an instance. The NP *siritana ia* 'this story' is taken up again after the verb *maaliŋ* 'to get lost', first by the pronoun *ia* PRX, and then by the appositive NP *baraŋ ia* 'this thing'. The verb bears the LH-L% boundary-tone complex and is immediately followed by the pronoun *ia*, which takes up the NP *siritana* 'this story'.

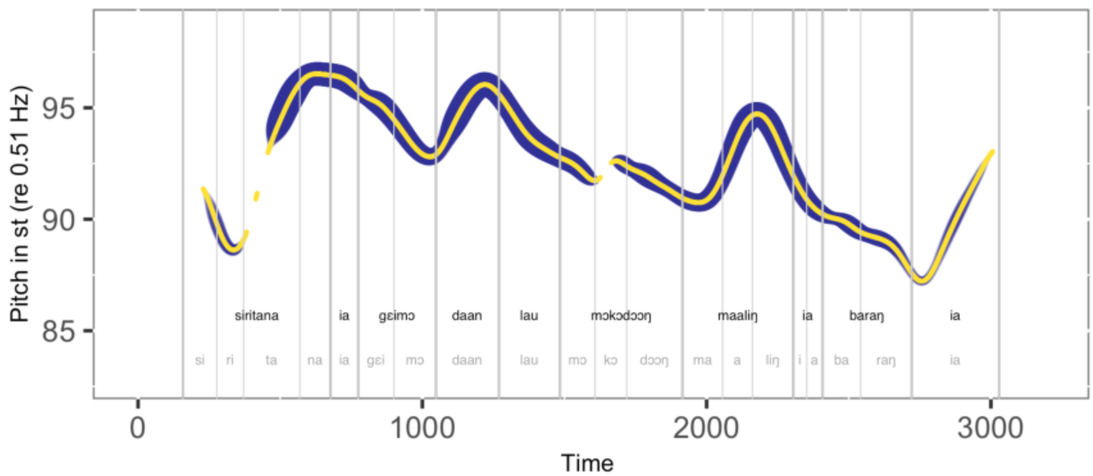

Figure 3.92: Periogram with pitch track (in st) for example (66), speaker SYNO

(66) *siritana | ia gɛimɔ daan | lau mɔkɔdɔɔng maaling | ia baraŋ ia*

sirita=na	ia	gɛimɔ	daan	lau	mɔkɔ-dɔɔŋ	mɔ-aliŋ	ia
story=3S.GEN	PRX	not	EXIST	presently	ST.AV-want	ST-disappear	PRX

baraŋ	ia
goods	PRX

‘This story will never again get lost; this thing.’

(explanation-lelegesan_SYNO.007) ▷

The question of what conditions the choice of either boundary-tone complex will remain a topic for further research.

4 Summary and Discussion

The present study aimed to achieve two primary goals: Firstly, to provide a detailed discussion of the prosody and intonation of Totoli, adopting a comprehensive approach that combined experimental evidence with data obtained from an impressionistic analysis of a wide-ranging corpus of (semi-)spontaneous speech. Secondly, to explore the grammatical structures that are typically found in (Compound) Intonation Units, which include singleton IUs, embedded IUs of CIUs, and complex CIUs as a whole.

Regarding the first objective, I have demonstrated that prominence is not a relevant concept in the prosody of Totoli, and that focus is not signaled by any prosodic cues. This was supported by evidence obtained from two experiments. The first experiment, which involved an RPT study, showed that naive native speakers generally could not agree on the location of prominences, suggesting that prominence may not be a significant category in Totoli's prosody. The results of the second experiment further supported this hypothesis, as no evidence was found that prosodic means marked focus as an information-structural category.

As Himmelmann & Kaufman (2020: 376) have noted, narrow focus on a subconstituent of a clause or noun phrase in languages is typically not signaled by intonation alone but rather by syntactic means. However, narrow focus on a constituent such as the subject NP or the object NP has only been investigated to a limited extent in Austronesian languages (cf. for example Nagaya & Hwang 2018, Kaufman 2005, Kaland et al. 2023). In this study, I have shown that in Totoli, syntactically equal SVO(O) constructions are not prosodically marked for focus when used as answers to questions that trigger focus on different constituents. This feature may be present in many other Austronesian languages, but additional data from a variety of Austronesian languages are required to determine whether it is a common feature or limited to a specific subgroup of Austronesian languages.

In §3.2, I examined the tonal patterns of the entire corpus of (semi-)spontaneous speech, consisting of 2861 Intonation Units. Based on the analysis of the tonal specifications, I presented a model of the Compound Intonation Unit in Totoli. This model assumes either singleton IUs or complex Compound IUs (CIUs). In the former, the only pitch event is the IU-final boundary complex, which occurs on the final two syllables. In the latter, the CIU comprises a series of two or more IUs, each of which bears one of the three IU-final boundary-tone complexes.

My analysis showed that the Totoli prosody is better described by assuming recursive embedding of IUs into CIUs rather than parsing of IUs into prosodic units at a level below the IU, as Himmelmann's (2018: 348) model of the IU in Austronesian languages of Indonesia and East Timor suggests. Himmelmann's model suggests that Intonation Units are further parsed into smaller prosodic units, such as intermediate phrases, and tonal patterns delimiting them consistently occur at the boundaries of major syntactic units. However, I found that the tonal patterns in my data are essentially the same, although with an inverted distribution. I have demonstrated that an embedded IU of a CIU differs substantially from a prosodic word or what is labeled as Accentual Phrase in Korean or French.

The absence of word prosody and the assignment of tone complexes to boundaries of prosodic domains fit the characteristics of what Féry (2016: 270) labels Phrase Languages:

> Phrase languages resemble intonation languages in that their tonal specifications are mostly assigned at the level of ϕ-phrases and ι-phrases. But contrary to intonation languages, specifications at the level of the word are sparse, absent or only weakly implemented. Phrase languages do not automatically associate pitch accents with stressed syllables, most tones are nonlexical (or 'postlexical'). (Féry 2016: 270)

In fact, many Austronesian languages may fall under the category Phrase Languages, following Himmelmann's (2018: 347) assertion:

> […] it seems likely that prosodic prominence does not have a major role to play in marking information-structural categories. If at all, prosodic phrasing may be of relevance in this regard inasmuch as it is not determined by syntactic or processing constraints.

Further evidence for recursive embedding of IUs into CIUs comes from an analysis of the grammatical structures that IUs typically contain. I found that a small set of categories suffices to describe the majority of their content. I compared the grammatical units typically occurring in embedded IUs with those that occur in singleton IUs which are not further segmented and I found that they are essentially similar, although again with varying proportions.

In sum, tonal patterns at the edges of singleton IUs and final IUs of CIUs are similar to those occurring at the right edge of non-final embedded IUs of CIUs.

The syntactic structures they contain also occur as simple, singleton IUs which are not further chunked. In light of these results, I concluded that singleton IUs and embedded IUs of CIUs are essentially of the same nature with the major difference being the presence or absence—i.e. the strength—of further typical boundary phenomena such as pitch reset, final lengthening, pauses and glottalization. A systematic analysis of boundary strength remains an object for future studies. Furthermore, although the tonal events at the edges of IUs are the same, it might be the case that they vary with regard to tonal scaling. That is, tonal events at the right-edge boundary of a CIU may be essentially the same as at the right-edge boundary of embedded IUs within CIUs but may vary in their tonal scaling (Riad 2018). This is an aspect which I have not systematically investigated here, and which presents a promising avenue for further research.

This research also opens many other questions. First, what does determine the choice of the final boundary tone of those IUs which are part of a CIU? Speakers are consistent in their choice of an IU-final boundary tone, and the grammatical unit contained in an IU appears to trigger the choice of boundary-tone complex. I suggested that the different patterns might be explicable by different degrees of integration, though the explanation for these different patterns requires further research. Second, what does trigger the realization of two grammatical units as either two separate IUs or a single complex one? Verbs followed by an NP often occur as a single IU. This is also observed with verbs followed by a PP, yet the tendency appears to be less strong. The analysis of the intonation of Totoli in Chapter 3 focused on the tonal patterns exclusively. Further investigations are needed in order to correlate the different boundary-tone complexes with other acoustic phenomena such as, for example, duration, intensity and possibly spectral tilt, voice quality. A particularly fruitful approach may be the description of the intonation of Totoli with the attractor-based model that encompasses categorical and also continuous components, but also accounts for the variation of their frequency (Roessig et al. 2019, Roessig 2021).

Little is known about the prosody and intonation of Austronesian languages. The study presented here pertains to one language in the region, and many of the results may well apply to other languages in the area. This study represents one of the most comprehensive investigations into the prosody and intonation of any Austronesian language to date. Further research on other languages is necessary to relate the reported results and insights to other languages in the region, and to determine which features are specific to Totoli and which are common to the region or the language family as a whole.

Appendix A: Appendix

A.1 RPT Experiment

A.1.1 Recorded speaker information

Table A.1 shows the speaker information of the RPT experiment discussed in §2.1.

Table A.1: Information of recorded speakers for the stimulus of the RPT experiment discussed in §2.1

ID	ORIGIN	GENDER	YEAR OF BIRTH	N SPEECH SAMPLES
FAH	Nalu	m	1964	14
IRN	Pinjan	m	1967	4
RSTM	Dadakitan	m	1966	17
SNG	Nalu	m	1940	15
SP	Nalu	m	1958	21

A.1.2 Participant information

Table A.2 shows the speaker information of the RPT experiment discussed in §2.1. The third column in the table refers to the participants' place of residence at the time of data collection, which almost always corresponds to the location where they grew up.

A.1.3 Instructions of boundary marking task

A.1.3.1 Indonesian original

Ketika seseorang berbicara, dia akan membagi ucapan mereka menjadi potongan-potongan. Potongan-potongan tersebut membentuk kelompok kata-kata yang memudahkan pendengar untuk memahami ucapan pembicara. Potongan-potongan

Table A.2: Speaker information of participants of RPT experiment discussed in §2.1

SPEAKER	YEAR OF BIRTH	PLACE OF LIVING	GENDER
AKR	1990	Nalu	m
BSTN	1976	Nalu	m
DHL	1988	Nalu	f
DT	1989	Nalu	m
DDN	1988	Nalu	m
FSL	1994	Binontoan	m
FTR1	1994	Binontoan	f
FTR2	1991	Binontoan	f
IFS	1986	Binontoan	m
IM	1972	Kalangkangan	f
MG	2000	Pinjan	m
NRBT	1983	Nalu	f
OCH	1994	Binontoan	f
RMDN	1994	Nalu	m
RID	1998	Binontoan	f
RST	1983	Nalu	m
RDT	1981	Nalu	m
SRN	1985	Nalu	f
STDI	1988	Kalangkangan	m
WN	1979	Kalangkangan	m

tersebut penting terutama saat pembicara memproduksi ucapan yang panjang.

Contoh potongan yang mungkin Anda ketahui adalah potongan nomor ketika Anda memberi tahu nomor telepon Anda kepada orang lain. Biasanya, Anda tidak setiap kali memberi satu nomor (0, 8, 1, 3 …), tetapi Anda akan memotong nomor hp tersebut menjadi kelompok-kelompok yang terdiri atas dua, tiga, atau empat angka (081, 358, 772 …).

Untuk rekaman yang akan Anda dengar, Anda diminta untuk menandai bagian yang terdengar sebagai satu potongan. Dengan mengklik kata terakhir dari ucapan tersebut, Anda dapat menetapkan batas, yang kemudian muncul di belakang kata yang diklik. Batas antara dua potongan tidak harus sama dengan lokasi tempat Anda akan menulis tanda koma, titik, atau tanda baca lainnya. Jadi, Anda harus benar-benar hati-hati mendengar ujaran dan tandai batas yang Anda dengar sebagai akhir sebuah potongan.

Jawaban yang Anda berikan tidak ada yang salah atau benar karena semuanya bergantung pada rasa bahasa.

Jika Anda mau memperbaiki pilihan Anda, Anda dapat mengklik kata tersebut untuk kedua kalinya, dan btas yang menjadi pilihan awal Anda akan lenyap.

Sebuah potongan mungkin saja berupa satu kata, atau mungkin terdiri atas beberapa kata, dan ukuran (jumlah kata) dalam setiap potongan dari para pembicara bisa saja berbeda-beda dalam satu ujaran. Beberapa ujaran mungkin Anda dengar konsisten, yaitu terdiri atas satu potongan saja. Jika demikian, Anda tidak perlu menandai batas potongan.

Anda dapat memutar setiap rekaman kalimat sebanyak dua kali. Akan tetapi, tidak memungkinkan untuk menghentikan rekaman pada saat contoh kalimat sedang diputar.
Contoh:

081|358|772...
0813|5877|2...
Bapak saya | sudah datang
Bapak | saya sudah datang

Selamat mengikuti eksperimen ini!
Silahkan menandai bagian yang Anda dengar sebagai satu potongan. Dengan mengklik kata terakhir salah satu potongan, batas akan muncul di belakang kata yang diklik.

A.1.3.2 English translation

When someone speaks, they divide their speech into segments. These segments form groups of words that make it easier for the listener to understand the speaker's message. These segments are especially important when the speaker produces long utterances.

An example of segments that you may be familiar with is the number segments when you give your phone number to someone else. Usually, you do not give the number one digit at a time (0, 8, 1, 3...), but you break up the phone number into groups consisting of two, three, or four digits (081, 358, 772...).

For the recordings that you will hear, you are asked to mark the sections that sound like one segment. By clicking on the last word of the speech segment, you can set a boundary, which will then appear behind the clicked word. The boundary between two segments does not necessarily have to be in the same location as where you would write a comma, period, or other punctuation marks. Therefore, you must listen carefully to the speech and mark the boundary that you hear as the end of a segment.

The answer you provided is neither right nor wrong, as it all depends on one's sense of language.

If you wish to revise your selection, you can click on the word again, and the boundary that was your initial choice will disappear.

A segment may consist of a single word or may be made up of several words, and the size (number of words) of each segment from the speakers may vary within one utterance. Some utterances may sound consistent, consisting of only one segment. If so, you do not need to mark any segment boundaries.

You can play each recorded sentence twice. However, it is not possible to stop the recording while the sample sentence is playing.
Example:

081|358|772...
0813|5877|2...
Bapak saya | sudah datang (approx. "My father already came home.")
Bapak | saya sudah datang (approx. "Father, I already came home.")

Enjoy the experiment!
Please mark the chunks you hear as one unit. By clicking on the last word of one of the chunk, a boundary will appear behind the clicked word.

A.1.4 Instructions of prominence marking task

A.1.4.1 Indonesian original

Dalam berbicara seseorang akan mengucapkan beberapa atau banyak kata dalam sebuah kalimat dengan nada yang lebih menonjol dibandingkan dengan kata-kata lain yang terdapat dalam kalimat tersebut. Kata-kata dengan nada yang menonjol ini biasanya dapat dirasakan oleh pendengarnya. Tugas Anda adalah menandai (mewarnai) kata-kata yang nadanya Anda dengar lebih menonjol dibandingkan dengan kata-kata lain dalam rekaman kalimat yang akan Anda putar.

Berikut ini Anda akan diputarkan 71 kalimat. Setiap kalimat juga akan disajikan dalam bentuk tertulis.

Tugas Anda adalah mewarnai semua kata yang nadanya Anda anggap lebih menonjol (mis. lebih tinggi) dibandingkan dengan kata-kata lain pada setiap rekaman kalimat yang Anda dengarkan. Untuk mewarnai kata, silakan menklik kata tersebut dan warnanya akan berubah menjadi merah:

Dia melihat **sapi**

Dalam hal ini, Anda dimungkinkan untuk memilih lebih dari satu kata pada setiap rekaman kalimat!

Dia melihat **sapi** dan kuda **makan** rumput

Jika Anda mau memperbaiki pilihan Anda, Anda dapat mengklik kata tersebut untuk kedua kalinya, dan kata yang menjadi pilihan awal anda akan kembali berubah warna menjadi hitam.

Anda dapat memutar setiap rekaman kalimat sebanyak dua kali. Akan tetapi, tidak memungkinkan untuk menghentikan rekaman pada saat contoh kalimat sedang diputar.

Selamat mengikuti eksperimen ini!

Tandai kata-kata yang terdengar **lebih menonjol** untuk Anda.

A.1.4.2 English translation

When speaking, individuals often emphasize certain words in a sentence through variations in tone. These prominent words are typically noticeable to the listener. Your objective is to identify and highlight (using color) the words in a recorded sentence where the speaker's tone stands out in comparison to the other words.

You will hear 71 sentences. You will also be provided with each sentence as a written transcript.

Your task is to color all the words that you deem to to stand out more (e.g. higher tone) compared to the other words in each recorded sentence that you listen to. You will also be provided with a written transcript of each sentence. To color a word, please click on the word and it will turn red:

S/he sees **a cow**

In this case, it is possible for you to choose more than one word in each recorded sentence!

S/he sees **a cow** and a horse **eating** grass

If you need to revise your selection, click on the word again, and it will revert back to its original color black.

You can play each recording twice. It will not be possible to stop the recording while it is playing.

Enjoy the experiment!

Mark the words that sound **more prominent** to you.

A.2 Focus marking

A.2.1 Recorded speaker information

Table A.3 shows the speaker information of the RPT experiment discussed in §2.2.

Table A.3: Information of recorded speakers for the stimulus of the Focus Marking experiment discussed in §2.2

SPEAKER	ORIGIN	GENDER	YEAR OF BIRTH
AKR	Nalu	m	1990
DT	Nalu	m	1989
FTR	Binontoan	f	1994
IFS	Binontoan	m	1986
SP	Nalu	m	1958
ZHRM	Tambun	m	1965

A.2.2 Participant information

Table A.4 shows the participant information of the focus marking experiment discussed in §2.2.3. The third column in the table refers to the participants' place of residence at the time of data collection, which almost always corresponds to the location where they grew up.

A.2.3 Stimuli

Examples (1)–(9) are the QA pairs that were used in §2.2.

(1) a. *inaŋ nɔŋinum sɔpa?*
inaŋ nɔN-inum sɔpa
mother AV.RLS-drink what
'What does the mother drink?'

b. *inaŋ nɔŋinum* ***ɔgɔ!***
inaŋ nɔN-inum ɔgɔ
mother AV.RLS-drink water
'The mother drinks water.'

(2) a. *isɛi nɔŋinum ɔgɔ?*
isɛi nɔN-inum ɔgɔ
who AV.RLS-drink water
'Who drinks the water?'

b. ***inaŋ*** *nɔŋinum ɔgɔ!*
inaŋ nɔN-inum ɔgɔ
mother AV.RLS-drink water
'The mother drinks water.'

Table A.4: Participant information of focus marking perception experiment discussed in §2.2.3

SPEAKER	YEAR OF BIRTH	PLACE OF LIVING	GENDER
AM	1978	Binontoan	m
ANDR	1997	Dapalak	f
AAL	1988	Nalu	m
BLW	1994	Binontoan	m
DAT	1989	Nalu	m
DWS	1994	Binontoan	f
EKW	1989	Binontoan	f
HLM	1986	Binontoan	f
IFS	1986	Binontoan	m
IRM	1972	Laulalang	f
IWRM	1978	Binontoan	m
ISRW	1999	Gio	f
JMTR	1993	Binontoan	f
SRMN	1986	Nalu	f
NSK	1980	Binontoan	f
NRM	1981	Binontoan	f
MRB	1994	Nalu	m
SRM	1985	Nalu	f
WN	1979	Kalangkangan	m
YK	1985	Binontoan	m

(3) a. *sɔpa niinum inaŋ?*
sɔpa ni-inum-0 inaŋ
what RLS-drink-UV mother
'What does the mother drink?'

b. ***ɔgɔ** niinum inaŋ!*
ɔgɔ ni-inum-0 inaŋ
water RLS-drink-UV mother
'The mother drinks water.'

(4) a. *ɔgɔ niinum isɛi?*
ɔgɔ ni-inum-0 isɛi
water RLS-drink-UV who
'Who drinks the water?'

b. *isɛi niinum **inaŋ**!*
ɔgɔ ni-inum-0 inaŋ
water RLS-drink-UV mother
'The mother drinks water.'

(5) a. *sɔpa nɔlugud sɛsɛŋ?*
sɔpa nɔN-lugud sɛsɛŋ
what AV.RLS-chase cat
'Who/What chases the cat?'

b. ***dɛuk** nɔlugud sɛsɛŋ!*
dɛuk nɔN-lugud sɛsɛŋ
dog AV.RLS-chase cat
'The dog chases the cat.'

(6) a. *dɛuk nɔlugud sɔpa?*
dɛuk nɔN-lugud sɔpa
dog AV.RLS-chase what
'What does the dog chase?'

b. *dɛuk nɔlugud **sɛsɛŋ**!*
dɛuk nɔN-lugud sɛsɛŋ
dog AV.RLS-chase cat
'The dog chases the cat.'

(7) a. *maŋana dɔlagɔ nɛmɛɛnan buuk dɛi isɛi?*
maŋana dɔlagɔ nɔN-bɛɛn-an buuk dɛi isɛi
child girl AV.RLS-give-APPL book LOC who
'Who does the girl give the book to?'

b. *maŋana dɔlagɔ nɛmɛɛnan buuk **dɛi inaŋna!***
maŋana dɔlagɔ nɔN-bɛɛn-an buuk dɛi inaŋ=na
child girl AV.RLS-give-APPL book LOC mother=3.SG
'The girl gives the book to her mother.'

(8) a. *maŋana dɔlagɔ nɛmɛɛnan sɔpa dɛi inaŋna?*
maŋana dɔlagɔ nɔN-bɛɛn-an sɔpa dɛi inaŋ=na
child girl AV.RLS-give-APPL what LOC mother=3.SG
'What does the girl give to her mother?'

b. *maŋana dɔlagɔ nɛmɛɛnan* ***buuk*** *dɛi inaŋna!*
maŋana dɔlagɔ nɔN-bɛɛn-an buuk dɛi inaŋ=na
child girl AV.RLS-give-APPL book LOC mother=3.SG
'The girl gives the book to her mother'

(9) a. *isɛi nɛmɛɛnan buuk dɛi inaŋna?*
isɛi nɔN-bɛɛn-an buuk dɛi inaŋ=na
who AV.RLS-give-APPL what LOC mother=3.SG
'Who gives the book to the mother?'

b. ***maŋana dɔlagɔ*** *nɛmɛɛnan buuk dɛi inaŋna!*
maŋana dɔlagɔ nɔN-bɛɛn-an buuk dɛi inaŋ=na
child girl AV.RLS-give-APPL book LOC mother=3.SG
'The girl gives the book to her mother'

A.2.4 Instructions

A.2.4.1 Indonesian original

Please listen carefully.
You will hear two question-answer pairs.
Only one of them is correct!
Your task is to choose a compatible pair.
You will hear each question-answer pair twice.
After the second time, you should choose the one that sounds more compatible.
Which pair is more compatible?

A.2.4.2 English translation

Tolong mendengar dengan seksama.
Anda akan mendengar dua pasangan pertanyaan-jawaban.
Hanya salah satunya adalah yang benar!
Tugas Anda adalah memilih pasangan yang cocok.
Anda akan mendengar setiap pasangan pertanyaan-jawaban dua kali.
Setelah kedua kalinya, Anda harus pilih salah satu yang kedengarannya lebih cocok.
Pasangan yang mana lebih cocok?

A.3 The corpus

Table A.5: Overview of conversational recordings

FILENAME	SPEAKER	GENDER	N IUs	DURATION
QUIS-animalgame	RSM, AKR	m, m	112	00:03:08
spacegames	KSR, SP	m, m	1215	00:37:00
			1327	00:40:08

Table A.6: Overview of monological recordings

FILENAME	SPEAKER	GENDER	N IUs	DURATION
explanation_lelegesan	SYNO	m	164	00:06:51
explanation_wedding-tradition	ZBR	m	321	00:16:13
pearstory	SP	m	46	00:02:03
pearstory	RSTM	m	192	00:06:23
pearstory	RD	f	72	00:03:11
pearstory	SP	m	51	00:02:33
pearstory	IRN	m	31	00:01:27
pearstory	MLI	f	44	00:04:22
pearstory	SNG	m	131	00:05:57
pearstory	SELP	f	70	00:02:38
pearstory	FAH	m	74	00:03:09
QUIS-animalgame	SP	m	89	00:11:09
QUIS-focus	SP	m	41	00:04:37
lifestory	RDA	m	198	00:07:35
story-monkey-butterfly	RSM	m	69	00:02:25
story-monkey-crocodile	RSM	m	47	00:01:35
story-monkey-python	RSM	m	27	00:02:33
story-monkey-turtle	RSM	m	72	00:02:14
story-session	MMN	f	88	00:03:13
			1899	01:30:08

References

Albert, Aviad, Francesco Cangemi, T. Mark Ellison & Martine Grice. 2022. *ProPer: Prosodic analysis with periodic energy*. DOI: 10.17605/OSF.IO/28EA5. osf.io/28ea5.

Albert, Aviad, Francesco Cangemi & Martine Grice. 2018. Using periodic energy to enrich acoustic representations of pitch in speech: A demonstration. In *Proceedings of Speech Prosody*, 804–808. Poznań, Poland. DOI: 10.21437/SpeechProsody.2018-162.

Allen, Shanley E. M. & Catherine Dench. 2015. Calculating mean length of utterance for eastern Canadian Inuktitut. *First Language* 35(4-5). 377–406. DOI: 10.1177/0142723715596648.

Altenberg, Bengt. 1987. *Prosodic patterns in spoken English: Studies in the correlation between prosody and grammar for text-to-speech conversion.* Lund, Sweden: Lund University Press.

Anikin, Andrey. 2019. Soundgen: An open-source tool for synthesizing nonverbal vocalizations. *Behavior Research Methods* 51(2). 778–792. DOI: 10.3758/s13428-018-1095-7.

Arnhold, Anja & Aki-Juhani Kyröläinen. 2017. Modelling the interplay of multiple cues in prosodic focus marking. *Laboratory Phonology* 8(1). 4. DOI: 10.5334/labphon.78.

Arvaniti, Amalia & Janet Fletcher. 2020. The autosegmental-metrical theory of intonational phonology. In Carlos Gussenhoven & Aoju Chen (eds.), *The Oxford handbook of language prosody*, 78–95. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198832232.013.4.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.

Baumann, Stefan, Martine Grice & Susanne Steindamm. 2006. Prosodic marking of focus domains - Categorical or gradient. In *Proceedings of Speech Prosody*, 301–304. Dresden, Germany. https://www.isca-speech.org/archive/speechprosody_2006/baumann06_speechprosody.html.

Baumann, Stefan & Bodo Winter. 2018. What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *Journal of Phonetics* 70. 20–38. DOI: 10.1016/j.wocn.2018.05.004.

Berkovits, Rochele. 1993. Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics* 21(4). 479–489. DOI: 10.1016/S0095-4470(19)30231-1.

Bickel, Balthasar & Fernando Zúñiga. 2017. The 'word' in polysynthetic languages: Phonological and syntactic challenges. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford handbook of polysynthesis*, 158–185. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.52.

Blust, Robert. 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics* 43(1). 73–148. DOI: 10.1353/ol.2004.0004.

Boersma, Paul & David Weenink. 2023. *Praat: Doing phonetics by computer*. http://www.praat.org/.

Bolinger, Dwight Le Merton. 1986. *Intonation and its parts: Melody in spoken English*. Stanford: Stanford University Press.

BPS Tolitoli. 2019. *Kecamatan Tolitoli Utara dalam angka 2019*. Indonesian. https://tolitolikab.bps.go.id/publication/2019/09/26/6d4cf9debe2e88ee8a38a48e/kecamatan-tolitoli-utara-dalam-angka-2019.html (3 January, 2023).

Bracks, Christoph. 2023. *Compound intonation units in Totoli: Postlexical prosody and the prosody-syntax interface*. OSF. DOI: 10.17605/OSF.IO/8CMQ4.

Bracks, Christoph. Submitted. Illustrations of the IPA: Totoli. *Journal of the International Phonetic Association*.

Bracks, Christoph, Maria Bardají i Farré, Datra Hassan, Sumitro Pogi & Nikolaus P. Himmelmann. 2023. *Totoli Documentation Corpus 2*. Data Center for the Humanities. DOI: 10.18716/dch/a.00000014.

Bracks, Christoph & Fabian C. Moss. 2022. Totoli's art of Lelegesan: Analyzing sociocultural context and musical content. In *Proceedings of the 10th International Workshop on Folk Music Analysis*. DOI: 10.31219/osf.io/5tsxa.

Brown, Gillian, Karen Currie & Joanne Kenworthy. 2015. *Questions of intonation* (Routledge Library Editions: The English Language). London: Routledge. DOI: 10.4324/9781315688664.

Byrd, Dani, Jelena Krivokapić & Sungbok Lee. 2006. How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America* 120(3). 1589–1599. DOI: 10.1121/1.2217135.

Chafe, Wallace L. (ed.). 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex Publ. Co.

Chafe, Wallace L. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, (Typological Studies in Language 11), 21–51. Amsterdam: John Benjamins. DOI: 10.1075/tsl.11.03cha.
Chafe, Wallace L. 1988. Punctuation and the prosody of written language. *Written Communication* 5(4). 395–426. DOI: 10.1177/0741088388005004001.
Chafe, Wallace L. 1993. Prosodic and functional units of language. In Jane A. Edwards & Martin D. Lampter (eds.), *Talking data: Transcription and coding in discourse research*, 33–44. New York: Psychology Press.
Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4. 109–149. DOI: 10.1017/S0952675700000798.
Cho, Taehong. 2005. Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ, i/ in English. *The Journal of the Acoustical Society of America* 117(6). 3867–3878. DOI: 10.1121/1.1861893.
Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
Cicourel, Aaron V. 2007. A personal, retrospective view of ecological validity. *Text & Talk* 27(5-6). 735–752. DOI: 10.1515/TEXT.2007.033.
Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1). 37–46. DOI: 10.1177/001316446002000104.
Cole, Jennifer, Timothy Mahrt & José I. Hualde. 2014. Listening for sound, listening for meaning: Task effects on prosodic transcription. In *Proceedings of Speech Prosody*, 859–863. DOI: 10.21437/SpeechProsody.2014-160.
Cole, Jennifer, Timothy Mahrt & Joseph Roy. 2017. Crowd-sourcing prosodic annotation. *Computer Speech & Language* 45. 300–325. DOI: 10.1016/j.csl.2017.02.008.
Cole, Jennifer, Yoonsook Mo & Mark Hasegawa-Johnson. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1(2). 425–452. DOI: 10.1515/labphon.2010.022.
Cole, Jennifer & Stefanie Shattuck-Hufnagel. 2016. New methods for prosodic transcription: Capturing variability as a source of information. *Laboratory Phonology* 7(1). 1–29. DOI: 10.5334/labphon.29.
Cramer, Duncan & Dennis Howitt. 2004. *The SAGE dictionary of statistics: A practical resource for students in the social sciences*. English. 1st edn. London: SAGE.
Croft, William. 1995. Intonation units and grammatical structure. *Linguistics* 33(5). DOI: 10.1515/ling.1995.33.5.839.

Croft, William. 2007. Intonation Units and grammatical structure in Wardaman and in cross-linguistic perspective. *Australian Journal of Linguistics* 27(1). 1–39. DOI: 10.1080/07268600601172934.

Cruttenden, Alan. 1997. *Intonation*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139166973.

Crystal, David. 1976. *Prosodic systems and intonation in English* (Cambridge Studies in Linguistics 1). Cambridge: Cambridge University Press.

de Vries, Lourens. 2005. Towards a typology of tail–head linkage in Papuan languages. *Studies in Language* 29(2). 363–384. DOI: 10.1075/sl.29.2.04vri.

Dombrowski, Ernst & Oliver Niebuhr. 2010. Shaping phrase-final rising intonation in German. In *Proceedings of Speech Prosody*. Chicago, USA. https://www.isca-speech.org/archive/pdfs/speechprosody_2010/dombrowski10_speechprosody.pdf.

Downing, Laura J. 2005. On the ambiguous segmental status of nasals in homorganic NC sequences. In Marc van Oostendorp & Jeroen van de Weijer (eds.), *The internal organization of phonological segments*, 183–216. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110890402.183.

Downing, Laura J. & Al Mtenje. 2017. *The phonology of Chichewa*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198724742.001.0001.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4). 805–855. DOI: 10.2307/415719.

Du Bois, John W, Susanna Cumming, Stephan Mark Schuetze-Coburn & Danae Paolino. 1992. Discourse transcription. *Santa Barbara papers in linguistics* 4.

Du Bois, John W, Stephan Mark Schuetze-Coburn, Susanna Cumming & Danae Paolino. 1993. Outline of discourse transcription. In Jane A. Edwards & Martin D. Lampter (eds.), *Talking data: Transcription and coding in discourse research*, 45–89. New York: Psychology Press.

Edwards, Jane A. & Martin D. Lampter (eds.). 1993. *Talking data: Transcription and coding in discourse research*. New York: Erlbaum. DOI: 10.4324/9781315807928.

Fenk-Oczlon, Gertraud & August Fenk. 2002. The clausal structure of linguistic and pre-linguistic behavior. In Betram F. Givón T. Malle (ed.), *The evolution of language out of pre-language* (Typological Studies in Language 53), 215–229. Amsterdam: John Benjamins. DOI: 10.1075/tsl.53.11fen.

Féry, Caroline. 2016. *Intonation and prosodic structure*. Cambridge: Cambridge University Press. DOI: 10.1017/9781139022064.

Fleiss, Joseph L. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76(5). 378–382. DOI: 10.1037/h0031619.

Ford, Cecilia E. & Sandra A. Thompson. 1996. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and grammar* (Studies in Interactional Sociolinguistics 13), 134–184. New York & Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620874.003.

Fougeron, Cécile & Patricia A Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6). 3728–3740. DOI: 10.1121/1.418332.

Goedemans, Rob & Ellen van Zanten. 2007. Stress and accent in Indonesian. *LOT Occasional series* 9. 35–62. https://dspace.library.uu.nl/bitstream/handle/1874/296762/bookpart.pdf?sequence=2&isAllowed=y.

Grice, Martine, James Sneed German & Paul Warren. 2020. Intonation systems across varieties of English. In Carlos Gussenhhoven & Aoju Chen (eds.), *The Oxford handbook of language prosody*, 285–302. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198832232.013.18.

Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2). 143–185. DOI: 10.1017/S0952675700003924.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110859263.

Gussenhoven, Carlos & Aoju Chen (eds.). 2020. *The Oxford handbook of language prosody*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198832232.001.0001.

Halliday, M. A. K. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics* 3(2). 199–244. DOI: 10.1017/S0022226700016613.

Haspelmath, Martin. 2015. *The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses*. https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf.

Herbert, Robert K. 1986. *Language universals, markedness theory, and natural phonetic processes*. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110865936.

Himmelmann, Nikolaus P. 1991. *Sourcebook on Tomini-Tolitoli languages: General information and word lists*. Canberra: Pacific Linguistics, Research School of Pacific & Asian Studies, The Australian National University. DOI: 10.15144/PL-511.

Himmelmann, Nikolaus P. 2006. The challenges of segmenting spoken language. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of* *language documentation*, 253–274. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110197730.253.

Himmelmann, Nikolaus P. 2010. Notes on Waima'a intonational structure. In Michael C. Ewing & Marian Klamer (eds.), *East Nusantara: Typological and areal analyses*, 47–69. Canberra: Pacific Linguistics, Australian National University. DOI: 10.15144/PL-618.

Himmelmann, Nikolaus P. 2018. Some preliminary observations on prosody and information structure in Austronesian languages of Indonesia and East Timor. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 347–374. Berlin: Language Science Press. DOI: 10.5281/zenodo.1402555.

Himmelmann, Nikolaus P. & Daniel Kaufman. 2020. Austronesia. In Carlos Gussenhoven & Aoju Chen (eds.), *The Oxford handbook of language prosody*, 370–383. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198832232.013.28.

Himmelmann, Nikolaus P. & D. Robert Ladd. 2008. Prosodic description: An introduction for fieldworkers. *Language Documentation & Conservation* 2(2). 244–274.

Himmelmann, Nikolaus P. & Sonja Riesberg. 2013. Symmetrical voice and applicative alternations: Evidence from Totoli. *Oceanic Linguistics* 52(2). 396–422. DOI: 10.1353/ol.2013.0021.

Himmelmann, Nikolaus P., Meytal Sandler, Jan Strunk & Volker Unterladstetter. 2018. On the universality of intonational phrases: A cross-linguistic interrater study. *Phonology* 35(2). 207–245. DOI: 10.1017/S0952675718000039.

Hyman, Larry M., Francis Katamba & Livingstone Walusimbi. 1987. Luganda and the strict layer hypothesis. *Phonology Yearbook* 4. 87–108. DOI: 10.1017/S0952675700000786.

Ipek, Canan & Sun-Ah Jun. 2013. Towards a model of intonational phonology of Turkish: Neutral intonation. *Proceedings of Meetings on Acoustics* 19(1). 060230. DOI: 10.1121/1.4799755.

Iwasaki, Shoichi. 1996. The syntactic and functional structures of intonation unit in Thai. *Pan-Asiatic Linguistics* 2. 750–762.

Iwasaki, Shoichi & Hongyin Tao. 1993. A comparative study of the structure of the intonation unit in English, Japanese, and Mandarin Chinese. In *Annual meeting of the Linguistic Society of America*.

Jun, Sun-Ah. 2005a. Korean intonational phonology and prosodic transcription. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 201–229. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.003.0008.

Jun, Sun-Ah. 2005b. Prosodic typology. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 430–458. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.003.0016.

Jun, Sun-Ah (ed.). 2005c. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.001.0001.

Jun, Sun-Ah (ed.). 2014a. *Prosodic typology II*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199567300.001.0001.

Jun, Sun-Ah. 2014b. Prosodic typology: by prominence type, word prosody, and macro-rhythm. In Sun-Ah Jun (ed.), *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199567300.003.0017.

Jun, Sun-Ah & Cécile Fougeron. 2000. A phonological model of French intonation. In Antonis Botinis (ed.), *Intonation: Analysis, modelling and technology* (Text, Speech and Language Technology (TLTB) 15), 209–242. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-011-4317-2_10.

Kaland, Constantijn, Marc Swerts & Nikolaus P. Himmelmann. 2023. Red and blue bananas: Time-series f0 analysis of contrastively focused noun phrases in Papuan Malay and Dutch. *Journal of Phonetics* 96. 1–20. DOI: 10.1016/j.wocn.2022.101200.

Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. In Malcom Ross & I Wayan Arka (eds.), *The many faces of Austronesian voice systems: some new empirical studies*, 175–196. Canberra: Pacific Linguistics, Research School of Pacific & Asian Studies, The Australian National University. DOI: 10.15144/PL-571.

Kember, Heather, Jiyoun Choi, Jenny Yu & Anne Cutler. 2021. The processing of linguistic prominence. *Language and Speech* 64(2). 413–436. DOI: 10.1177/0023830919880217.

Kreckel, Marga. 1981. Tone units as message blocks in natural discourse: Segmentation of face-to-face interaction by naive, native speakers. *Journal of Pragmatics* 5(5). 459–476. DOI: 10.1016/0378-2166(81)90029-1.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276. DOI: 10.1556/aling.55.2008.3-4.2.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511808814.

Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174. DOI: 10.2307/2529310.

Lee, Yong-cheol, Bei Wang, Sisi Chen, Martine Adda-Decker, Angélique Amelot, Satoshi Nambu & Mark Liberman. 2015. A crosslinguistic study of prosodic focus. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4754–4758. IEEE. DOI: 10.1109/icassp.2015.7178873.

Leto, Claudia, Winarno S. Alamudi, Nikolaus P. Himmelmann, Jani Kuhnt-Saptodewo, Sonja Riesberg & Hasan Basri. 2005–2010. *DoBeS Totoli documentation*. https://dobes.mpi.nl/projects/totoli/.

Levinson, Stephen C., Penelope Brown, Eve Danzinger, Lourdes de León, John B. Haviland, Eric Pederson & Gunter Senft. 1992. Man and tree & space games. In Stephen C. Levinson (ed.), *Space stimuli kit 1.2*, 7–14. Nijmegen: Max Planck Institute for Psycholinguistics. DOI: 10.17617/2.2458804.

Lieberman, Philip. 1966. *Intonation, perception, and language*. Massachusetts Institute of Technology. (Doctoral dissertation). DOI: 1721.1/12996.

Lieberman, Philip, Masayuki Sawashima, Katherine S. Harris & Thomas Gay. 1970. The articulatory implementation of the breath-group and prominence: Crico-thyroid muscular activity in intonation. *Language* 46(2). 312–327. DOI: 10.2307/412281.

Määttä, Taisto. 1993. Prosodic structure indices in a folk tale. *Phonum* 2. 107–120.

Mahrt, Tim. 2016. *LMEDS: Language markup and experimental design software*. http://prosody.beckman.illinois.edu/lmeds.html.

Maskikit-Essed, Raechel & Carlos Gussenhoven. 2016. No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology* 33(2). 353–389. DOI: 10.1017/s0952675716000154.

Mathôt, Sebastiaan, Daniel Schreij & Jan Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2). 314–324. DOI: 10.3758/s13428-011-0168-7.

Matsumoto, Kazuko. 2000. Japanese intonation units and syntactic structure. *Studies in Language* 24(3). 515–564. DOI: 10.1075/sl.24.3.03mat.

Matsumoto, Kazuko. 2003. *Intonation units in Japanese conversation: Syntactic, informational and functional structures*. Amsterdam: John Benjamins. DOI: 10.1075/slcs.65.

Miller, George A. 1994. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 101(2). 343–352. DOI: 10.1037/0033-295x.101.2.343.

Mo, Yoonsook, Jennifer Cole & Eun-Kyung Lee. 2008. Naïve listeners' prominence and boundary perception. *Proceedings of Speech Prosody*. 735–738. https://www.isca-speech.org/archive/speechprosody_2008/mo08_speechprosody.html.

Mücke, Doris & Martine Grice. 2014. The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *Journal of Phonetics* 44. 47–61. DOI: 10.1016/j.wocn.2014.02.003.

Nagaya, Naonori & Hyun Kyung Hwang. 2018. Focus and prosody in Tagalog. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 375–388. Berlin: Language Science Press. DOI: 10.5281/zenodo.1402557.

Nespor, Marina & Irene Vogel. 1983. Prosodic structure above the word. In Anne Cutler & D. Robert Ladd (eds.), *Prosody: Models and measurements*, 123–140. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-69103-4_10.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110977790.

Ots, Nele & Eva Liina Asu. 2019. The perception of prominence: A rapid prosody transcription study with Estonian listeners. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 10(1). 145–165. DOI: 10.12697/jeful.2019.10.1.08.

Park, Joseph Sung-Yul. 2002. Cognitive and interactional motivations for the intonation unit. *Studies in Language* 26(3). 637–680. DOI: 10.1075/sl.26.3.07par.

Pater, Joe. 2004. Austronesian nasal substitution and other NC effects. In John J. McCarthy (ed.), *Optimality theory in phonology: A reader*, chap. 14, 269–289. Malden, USA: John Wiley & Sons, Ltd. DOI: 10.1002/9780470756171.ch14.

Pierrehumbert, Janet Breckenridge. 1980. *The phonology and phonetics of English intonation*. Massachusetts Institute of Technology. (Doctoral dissertation). DOI: 1721.1/16065.

Pitt, Mark A., Laure Dilley, Keith Johnson, Scott Kieslin, William Raymond, Hume Elizabeth & E Fosler-Lussier. 2007. *Buckeye corpus of conversational speech*. Columbus, OH. www.buckeyecorpus.osu.edu.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language* (General Grammar Series). London: Longman.

Reid, Lawrence A. 2000. Sources of Proto-Oceanic initial prenasalization: The view from outside Oceanic. In Videa P. De Guzman & Byron Bender (eds.), *Grammatical analysis: Morphology, syntax, and semantics, studies in honor of Stanley Starosta* (Oceanic Linguistics Special Publications 29), 30–45. Honolulu: University of Hawai'i Press.

Riad, Tomas. 2018. The phonological typology of North Germanic accent. In Larry M. Hyman & Frans Plank (eds.), *Phonological typology*, 341–388. Berlin, Boston: Mouton De Gruyter. DOI: 10.1515/9783110451931-010.

Riehl, Anastasia. 2008. *The phonology and phonetics of nasal-obstruent sequences*. Cornell University. (Doctoral dissertation). DOI: 10.5281/zenodo.3976876.

Riesberg, Sonja. 2014. *Symmetrical voice and linking in Western Austronesian languages* (Pacific Linguistics 646). Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9781614518716.

Riesberg, Sonja. 2019. 'Fern leaf, so you are challenging me?'Some observations on the Lelegesan, a form of verbal combat in Totoli. *NUSA: Linguistic studies of languages in and around indonesia* 66. 83–96. DOI: 10.15026/93966.

Riesberg, Sonja, Janina Kalbertodt, Stefan Baumann & Nikolaus P. Himmelmann. 2018. On the perception of prosodic prominences and boundaries in Papuan Malay. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 389–414. Berlin: Language Science Press. DOI: 10.5281/zenodo.1402559.

Riesberg, Sonja, Janina Kalbertodt, Stefan Baumann & Nikolaus P. Himmelmann. 2020. Using rapid prosody transcription to probe little-known prosodic systmes: The case of Papuan Malay. *Labratory Phonology* 11(1). 1–35. DOI: 10.5334/labphon.192.

Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561. DOI: 10.1353/lan.2019.0055.

Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2021. The many ways of transitivization in totoli. In Silvia Luraghi & Elisa Roma (eds.), *Valency over time*, 235–264. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110755657-008.

Roessig, Simon. 2021. *Categoriality and continuity in prosodic prominence* (Studies in Laboratory Phonology 10). Berlin: Language Science Press. DOI: 10.5281/zenodo.4121875.

Roessig, Simon, Doris Mücke & Martine Grice. 2019. The dynamics of intonation: Categorical and continuous variation in an attractor-based model. *PLoS ONE* 14(5). 1–32. DOI: 10.1371/journal.pone.0216859.

Schiering, René, Balthasar Bickel & Kristine A. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709. DOI: 10.1017/S0022226710000216.

Schuetze-Coburn, Stephan Mark. 1994. *Prosody, syntax and discourse pragmatics: Assessing information flow in German conversation*. University of California, Los Angeles. (Doctoral dissertation).

Schuetze-Coburn, Stephan Mark, Marian Shapley & Elizabeth G. Weber. 1991. Units of intonation in discourse: A comparison of acoustic and auditory analyses. *Language and Speech* 34(3). 207–234. DOI: 10.1177/002383099103400301.

Schwiertz, Gabriele. 2009. *Intonation & Prosodic Structure in Beaver (Athabaskan)–Explorations on the language of the Danezaa*. Universität zu Köln. (Doctoral dissertation). http://kups.ub.uni-koeln.de/id/eprint/2935.

Selkirk, Elisabeth O. 1986. *Phonology and syntax: The relationship between sound and structure*. Cambridge, MA: MIT press.

Selkirk, Elisabeth O. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 435–485. London: John Wiley & Sons, Ltd. DOI: 10.1002/9781444343069.ch14.

Skopeteas, Stavros, Ines Fiedler, Sam Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS): Refence manual* (ISIS 4). Potsdam: Universitätsverlag Potsdam. https://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/Information_Structure_Questionnaire_Manual.pdf.

Stoel, Ruben B. 2006. The intonation of Banyumas Javanese. In *Proceedings of Speech Prosody*, 827–830. https://www.isca-speech.org/archive/speechprosody_2006/stoel06_speechprosody.html.

Stoel, Ruben B. 2005. *Focus in Manado Malay: Grammar, particles, and intonation.* Leiden: CNWS publications.

Swerts, Marc, Emiel Krahmer & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics* 30(4). 629–654. DOI: 10.1006/jpho.2002.0178.

Tallman, Adam J. R. 2020. Beyond grammatical and phonological words. *Language and Linguistics Compass* 14(2). e12364. DOI: 10.1111/lnc3.12364.

Tao, Hongyin. 1996. *Units in Mandarin conversation: Prosody, discourse, and grammar.* Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/sidag.5.

Team, R Core. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

’t Hart, Johan. 1981. Differential sensitivity to pitch distance, particularly in speech. *The Journal of the Acoustical Society of America* 69(3). 811–821. DOI: 10.1121/1.385592.

’t Hart, Johan, René Collier & Antonie Cohen. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody* (Cambridge Studies in Speech Science and Communication). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511627743.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30(2). 219–255. DOI: 10.1162/002438999554048.

Turk, Alice & Stefanie Shattuck-Hufnagel. 2020. *Speech timing: Implications for theories of phonology, phonetics, and speech motor control*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198795421.001.0001.

Vogel, Irene. 2019. Life after the strict layer hypothesis. In Hongming Zhang & Youyong Qian (eds.), *Prosodic studies: Challenges and prospects*, 9–60. London: Routledge. DOI: 10.4324/9781351212878-2.

Wagner, Michael. 2012. Focus and givenness: A unified approach. In Ivona Kučerová & Ad Neeleman (eds.), *Contrasts and positions in information structure*, 102–147. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511740084.007.

Wouk, Fay. 1998. Solidarity in Indonesian conversation: The discourse marker *kan*. *Multilingua* 17(4). 379–406. DOI: 10.1515/mult.1998.17.4.379.

Wouk, Fay. 2008. The syntax of intonation units in Sasak. *Studies in Language* 32(1). 137–162. DOI: 10.1075/sl.32.1.06wou.

You, Hie-Jung. 2012. Determining prominence and prosodic boundaries in Korean by non-expert rapid prosody transcription. In *Proceedings of Speech Prosody*, 318–321. https://www.isca-speech.org/archive/speechprosody_2012/you12_speechprosody.html.

Name index

Language index

Subject index

Compound Intonation Units in Totoli

This book is an investigation into aspects of prosody, intonation and the prosody-syntax interface in Totoli, an endangered Austronesian language. With a strongly data-driven approach, the study integrates a combination of experimental evidence from both production and perception with corpus-based evidence through descriptive and inferential statistics.

The study takes the prime structuring unit of speech – the Intonation Unit – as its principal unit of investigation. It presents a thorough description of the IU, develops an intonational model of it, and investigates the syntactic units it contains. The author argues that the data is best analysed by assuming recursive embedding of Intonation Units into Compound Intonation Units.

This research represents a significant advancement in our understanding of the nature of prosodic systems found in the languages of the region and in intonational systems in general. It is one of the few investigations into the intonation of Austronesian languages and its analytical proposals are relevant both to prosodic theory and to phonological typology.