Chapter 12

Samoan intonation and challenges for autosegmental-metrical theory

Abstract (147 words):

Previous work on Samoan intonation has mostly focused on prosodic interfaces. Here, we focus on intonational phonology and highlight aspects of Samoan intonation that present challenges for autosegmental-metrical theory. Building on Orfitelli and Yu (2009), Yu (2011), Calhoun (2015, 2017), Yu and Stabler (2017), Calhoun et al. (2019), Yu (2021), we describe common nuclear contours in Samoan and review analyses of sentence-medial edge tones. We show the value of transcending biases that have crept into AM theory in practice. Dissociating nuclear pitch accents from prominence can help us make sense of intonational phrase-final tonal events in Samoan that are not focal accents. Taking 'edge tone' only as a description of where a tone is realized—without any implication that the grammatical source of the tone is necessarily a prosodic constituent—helps us recognize that there could be two distinct sub-classes of edge tones in Samoan: morphosyntactic tones as well as prosodic boundary tones.

Biography: (to be added)

12.1 Introduction

Samoan is a Polynesian language in the Austronesian family from the the independent state of Samoa and the United States territory of American Samoa (Blust, 2013). Austronesian languages have long captured the interest of syntacticians for cross-linguistically unusual patterns such as verb-initial word order, voice systems and passivization, and case alignment (Clemens and Polinsky, 2017, Potsdam and Polinsky, To appear)—especially ergative-absolutive case alignment patterns in the Polynesian languages of the South Pacific (Chung, 1978). In comparison, intonation is one of the least studied aspects of Austronesian languages (Blust, 2013, p. 251), and, perhaps due to the emphasis on syntactic phenomena, has largely focused on the syntax-prosody interface in Polynesian languages. Aspects of prosodic phrasing have been used to argue for particular analyses of verb-initial word order and related phenomena in the Polynesian languages Niuean (Clemens, 2014, 2019), Tagalog (Sabbagh, 2014, Hsieh, 2016, Sabbagh, 2016, Richards, 2017), and Tongan (Ahn, 2016), as well as in Malagasy (Edmiston and Postdam, 2016, Edmiston and Potsdam, 2017) (a non-Polynesian language, but one that patterns syntactically with Philippine-type languages like Tagalog (Blust, 2013, p. 69)). Yet, the only basic descriptions of the intonational phonology of these languages are preliminary analyses of Tongan (Kuo and Vicenik, 2012) and Malagasy (Aziz, 2020*a*,*b*). There has been more work on the intonational phonology of Samoan, although also primarily in the context of interfaces with syntax, as well as information structure (Yu, 2011, Calhoun, 2015, 2017, Yu and Stabler, 2017, Yu, 2018, Calhoun et al., 2019, Yu, 2021). Work focusing on intonational phonology has been limited to a section in Mosel and Hovdhaugen's reference grammar (Mosel and Hovdhaugen, 1992, p. 36-43), a preliminary ToBI-style analysis (Orfitelli and Yu, 2009), and research on information-seeking interrogatives (Howard, 2018).

The goals of this chapter are to consolidate and build on this previous work to provide a current phonological analysis of Samoan intonation and to highlight aspects of Samoan intonation that present challenges for autosegmental-metrical theory. Following a brief background section (§12.1.1) on word-level stress and syntax, §12.2 introduces common nuclear contours (also called 'nuclear configurations'), i.e., the tonal sequence of an intonational phrase-final (IntP[1]-final) pitch accent followed by a boundary tone, see e.g., Pierrehumbert (1980), Pierrehumbert and Hirschberg (1990), Hayes and Lahiri (1991), Gussenhoven (2004), Jun (2014), Frota and Prieto (2015*b*), Grice (2022). Among these are a falling contour that appears in interrogatives and sometimes in continuations, for which we propose a different analysis from Howard (2018). We also show that this contour can appear in narratives with a distribution characteristic of "uptalk" (Warren, 2016). After the discussion of nuclear contours, §12.3 turns to the main source of contention in the analysis

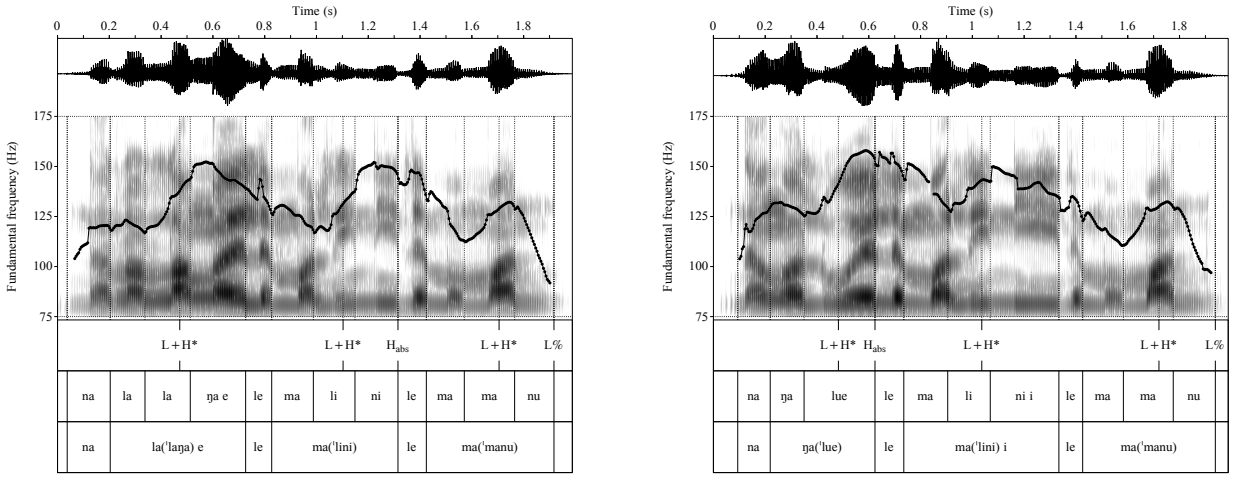
[1] We abbreviate 'Intonational Phrase' as IntP rather than IP to avoid confusion with the syntactic constituent Inflectional Phrase.

of Samoan intonation: sentence-medial edge tones. Different analyses have attributed the source of these tones to morphosyntactic spellout (Yu and Stabler, 2017, Yu, 2021), syntax-prosody constituent mapping (Yu, 2011, Calhoun, 2015), phonological grammar (Orfitelli and Yu, 2009), and information structure (Calhoun, 2017, Calhoun et al., 2019). We show how hypothesizing that there are multiple grammatical sources of these edge tones—namely, morphosyntactic spellout as well as tones inserted at prosodic phrase edges—rather than a single source, provides a way to resolve apparently conflicting data about the distribution of the edge tones.

The term 'edge tone' is often taken to be interchangeable with the term '(prosodic) boundary tone' and associated with a "major prosodic boundary", i.e., above the level of a prosodic word—a phonological phrase boundary or intonational phrase boundary (Ladd, 2008, p. 44, 47, 100). Here, we follow the use of the term 'edge tone' introduced in Yu (2021, p. 293, §6.3.2) (see §12.3 and §12.4.2 for further discussion), as a more general definition than Ladd (1995, p. 118)'s definition of 'edge tone'. Ladd (1995, p. 118) defines 'edge tone' as "a cover term for what are variously called boundary tones, phrase tones, phrase accents, non-prominence-lending pitch movements, and so on"; see also Ladd (1996, p. 285-286). The generalization we make follows the insight of Bruce (1977), as described in Pierrehumbert (2000, p. 19): "Bruce's treatment of Swedish accent and intonation... distinguish[es] the morphological source of a tone from the phonological spot where it shows up on the surface. By the morphological source, I mean the domain in which a choice of tone is contrastive." We use 'edge tone' simply to describe "the phonological spot" where a tone "shows up on the surface", without any specification of the grammatical source of the tone. The tone does not necessarily need to be inserted in the phonological grammar as a prosodic boundary tone, and could, for example, be inserted as a reflex of morphosyntactic spellout and/or pragmatic structure, or even as a lexical tone. The assumption that edge tones are necessarily prosodic boundary tones are one of the biases of Autosegmental-Metrical (AM) Theory in practice that Samoan intonation calls into question. In §12.4, we discuss how Samoan intonation raises issues for this bias of AM theory in practice, as well as the bias to relate pitch accents to word-level stress and prominence in ways that overgeneralize from West Germanic languages.

12.1.1 Background

Samoan has default VSO word order and marks ergative case on the subject of a verb-initial transitive sentence with the preposition /e/, as exemplified in the transitive sentence in (1a). 'Absolutive' case on the direct object of a transitive sentence and the subject of an intransitive sentence has been said to be unmarked (Chung 1978: p. 54–56; Ochs 1982: p. 649; Collins 2014: p. 94) and where an absolutive case marker would be expected to surface is indicated with ∅ as a place holder. The intransitive sentence (1b) with absolutive case on the subject also illustrates the prepositional ele-



(a) Transitive declarative, (1a) (b) Intransitive declarative, (1b)

Figure 12.1: F0 contours for transitive clause (1a) and intransitive clause (1b) from la_m01. As in all individual F0 contours shown in this paper, the dashed lines overlaying the F0 contour mark syllable divisions given in the Samoan word-level tier of the textual transcriptions at the bottom of the figures and the overlaid wide-band spectogram has a frequency range from 0 to 5000 Hz.

ment /i/ as a marker of oblique case on the optional *i le mamanu* 'on the design'. F0 contours for (1) are given in Figure 12.1.

(1) Basic verb-initial transitive and intransitive clauses[2]

a. Transitive clause

na la(ˈlaŋa) e le ma(ˈlini) ∅ le ma(ˈmanu).
PAST weave ERG DET marine ABS DET design
'The marine wove the design.'

b. Intransitive clause

na ŋa(ˈlue) ∅ le ma(ˈlini) i le ma(ˈmanu).
PAST work ABS DET marine OBL DET design
'The marine worked (on the design).'

The parentheses in the Samoan examples indicate footing, ˈ indicates primary stress, and ˌ indicates secondary stress. Phonotactically licit syllable shapes in Samoan are: monomoraic

[2]The following abbreviations are used in examples in this paper: ABS absolutive; Adv adverb; A agent; CONJ conjunction; COORD coordination; DET determiner; DIR directional particle; DU dual; DISJ disjunction; ERG ergative; GEN genitive; H_{abs} absolutive high edge tone; H_{coord} coordination high edge tone; H_{front} fronted expression high edge tone; H-/% high prosodic boundary tone, L-/% low prosodic boundary tone ; NEG negation; O object; OBL oblique; PART particle; PERF perfective; PRES present; RED reduplicant; S subject; SG singular; TOP topic marker; V verb.

/(C)V/, and bimoraic /(C)Vː/ and /(C)VV/. The basic stress assignment pattern observed in native monomorphemes can be described as primary stress on the penultimate mora and analyzed with inviolable constraints enforcing trochaic, bimoraic feet within a prosodic word, e.g., (ˈmanu) 'animal', (ˈlaː) 'sun', with the the head foot aligned to the right edge of the prosodic word, e.g., ma(ˈmanu) 'design', (ˌsaː)(ˈmoa) 'Sāmoa' (Zuraw et al., 2014),[3] see also Churchward (1951, p. 9, 24) and Mosel and Hovdhaugen (1992, p. 28). Exceptions come from stress on the antepenultimate mora in diphthongization of /ai, au, ei, ou/ VV sequences and additionally /aea/, /aoa/ VVV sequences, e.g., (ˈmai)le, *ma(ˈile) 'dog' but ma(ˈela), *(ˈmae)la 'hollow', (ˈmae)a, *ma(ˈea) 'get'; see Zuraw et al. (2014, §8) for further details and analysis.

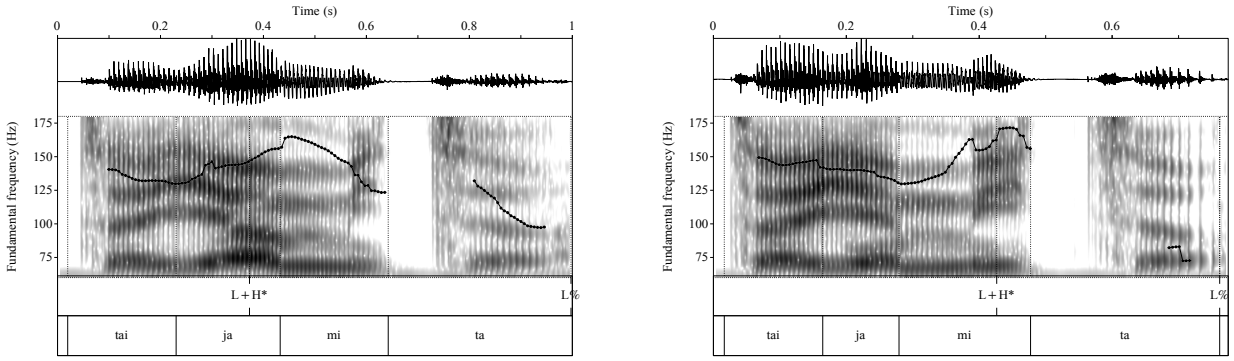
Evidence that pitch accent tones are associated to stressed moras come from comparisons like Figures 3a, 3b in Orfitelli and Yu (2009), which show that le(ˈlei) 'to be good' has a rising F0 movement in the final syllable, while (ˈlele) 'to fly' has one in the penultimate syllable; see also Figure 1 in Zuraw et al. (2014) for another comparison of F0 contours between primary stress in the final vs. the penultimate syllable. We call this tone a L+H* pitch accent and explain why in §12.2.1. Additional evidence for a stress-driven pitch accent with stress position shifted further leftward than the penultimate syllable comes from loanword adaptation. Zuraw et al. (2014, §4.4) found alternation in stress patterns of monomorphemic loanwords with five moras conditioned on the presence of epenthetic vowels. Namely, stress is avoided on epenthetic vowels (bolded in the following examples). Thus, (ˌkom**i**)pi(ˈuta), *ko(ˌm**i**pi)(ˈuta) 'computer' patterns with (ˌtemo)k**a**(ˈlasi), *te(ˌmok**a**)(ˈlasi) 'democracy' in receiving initial secondary stress, but p**u**(ˌɪini)(ˈsese), *(ˌp**u**ɪi)ni(ˈsese) 'princess' does not and receives peninitial secondary stress. There can also be some variability within/between speakers for these stress patterns. Figures 2 and 3 in Zuraw et al. (2014, §4.4) exemplify that rising F0 movements track with secondary stress position in words with initial vs. peninitial stress. Another example of variation in stress patterns within a speaker for a single loanword comes from the pronunciations Speaker la_m01 (see Appendix for speaker information) offered for *taiamita* 'diameter' (2), with primary stress on the antepenultimate mora (faithful to the English source) or on the penultimate mora, see also Orfitelli and Yu (2009, p. 2-3). The speaker's pronunciations of *taiamita* in isolation are shown in Figure 12.2. A rising F0 contour clearly occurs over the stressed mora, whether penultimate or antepenultimate.

(2) Variable antepenultimate or penultimate primary stress for *taiamita* 'diameter'

a. taiˈamita (antepenultimate)

b. (ˌtaia)(ˈmita) (penultimate)

[3] Samoan orthography marks long vowels with a macron; Samoa/Samoan is properly written as Sāmoa/Sāmoan; glottal stops are marked with an apostrophe, and /ŋ/ is written as a 'g'.



(a) Antepenult primary stress (2a)

(b) Penult primary stress (2b)

Figure 12.2: F0 contours for utterances of loanword *taiamita* 'diameter' (2) in isolation, with variable primary stress on the antepenultimate or penultimate mora, from la_m01.

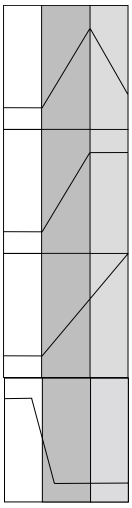
The rising pitch accent followed by a low intonational phrase boundary tone in Figure 12.2 is an example of a common nuclear contour in declaratives. This and other common nuclear contours are described in the next section, §12.2.

12.2 Nuclear contours

There are at least four common nuclear contours in Samoan, schematized and summarized in Table 12.1: a rising pitch accent followed by a fall typically seen in declaratives (§12.2.1), a rising pitch accent followed by a high plateau typically seen in enumeration (§12.2.2), a rising pitch accent followed by a continued rise typically seen in biased echo questions and in expressing disbelief (§12.2.3), and an "early" falling pitch accent followed by a low plateau typically seen in unbiased, information-seeking interrogatives (§12.2.4). The schematics assume penultimate stress; the dark grey middle section represents the accented TBU and the final light grey section represents the IntP-final TBU, and we assume that the TBU is the mora.

Following the International Prosodic Alphabet (IPrA) category labels proposed in Hualde and Prieto (2016, Figure 6) and also tonal categories and diacritics proposed in Jun and Fletcher (2014, (7)), the nuclear contours could be described as: (i) L+H* L%, (ii) L+H* H%, (iii) L+H* ˆH%, and (iii) H+>L* L%. The 'ˆ' diacritic in 'ˆH' for echo questions indicates a much higher F0 level, and the early timing '>' diacritic in 'H+>L*' for unbiased interrogatives indicates that that the fall is not over the accented TBU (like for H+L* in Hualde and Prieto (2016, Figure 6), which is the label used in Calhoun (2015), Howard (2018) et seq.), but rather, begins over the TBU preceding the accented TBU. The category label choices here are descriptive and emphasize relevant phonetic detail rather than committing to one particular phonological analysis at this point. For example,

Table 12.1: Schematic nuclear tonal contours



L+H* L%	Declaratives[4]
L+H* H%	Enumeration, continuation
L+H* ˆH%	Echo questions, disbelief
H+>L* L%	Information-seeking polar/wh-interrogatives, continuation[5]

the three-way contrast with rising contours could be analyzed without proposing phonologically contrastive high tones (H, ˆH), by proposing the absence of a final boundary tone as the phonological analysis for a final fall (see Grabe (1998, §2.4) and references therein and also brief historical overview in Ladd (2022, §6.3)) and an upstep rule for tones following L+H*, i.e., (i) L+H* L% $\rightarrow$ L+H* $\varnothing$, (ii) L+H* H% $\rightarrow$ L+H* L%, (iii) L+H* ˆH% $\rightarrow$ L+H* H% and (iii) H+>L* L% $\rightarrow$ H+>L* $\varnothing$. In addition, if there are no falling pitch accents other than the H+>L*, then the detail of alignment indicated by '>' could be introduced in phonetic implementation without introducing alignment as a source of phonological contrast.

12.2.1 Rising pitch accent followed by fall: L+H* L%

Before turning to the nuclear L+H*, we need to say a brief word about the prenuclear realization of L+H* accents. In prenuclear position, the L+H* is realized as an F0 rise that starts at the beginning of a stressed (C)V syllable and peaks at its offset or slightly after, as can be seen in Figures 12.1 and also mentioned elsewhere, e.g., in Orfitelli and Yu (2009), Calhoun (2015), Yu (2021). The phonetic alignment of the 'L' is to the beginning of the (C)V syllable, e.g., to the onset [l] of [li] in [ma(ˈlini)], [m] of [ma] in [ma(ˈmanu)], and [l] of [lu] in [ŋa(ˈlue)] in Figures 12.1 and [l] of [le] in [me(ˈleni)] in Figure 12.9, the glide [j] of [ja] in the surface realization of [tai(ˈami)ta] in Figure 12.2a, or to the nucleus, if there is no onset. The high peak of a sentence-medial pitch accent not followed by an H edge tone is typically aligned at the right edge of a stressed (C)V syllable,

[4]The fall into the post-tonic mora is schematized to end higher than the f0 contour starting point to symbolize the availability of truncation of the fall, see §12.2.1.

[5]The beginning of the low plateau is schematized to begin slightly after the onset of the stressed mora to symbolize that it typically starts in the syllable nucleus rather than the onset, see §12.2.4.

e.g., in [ma(ˈlini)] in Figure 12.1b, or slightly after. Peak delay or a short high plateau into the beginning of the posttonic syllable can be observed, e.g., in la(laŋa) in Figure 12.1a and ma(lini) in Figure 12.1b. Further work is needed to check details of phonetic alignment within the syllable with respect to the onset and the rime.

The phonetic alignment of the L and H F0 inflection points just described is generally preserved when pitch accents are close together because of adjacent stresses and/or at faster speech rates. Figure 12.3 shows F0 contours for neighboring pitch accents on the subject and the object in the sentence template in (3) uttered at the speaker's normal speech rate and a faster one (1.2 times faster). There are no H edge tones between the pitch accents, which are 1 mora apart.

(3) Realization of prenuclear pitch accent under different speech rates:

na ma(ˈno)ŋi le manu / la: i le maile / liona i le afiafi / taeao
PAST smell DET bird / sun OBL DET dog / lion OBL DET morning / afternoon

'The bird/sun smelled to the dog/lion in the morning/afternoon.'

The subjects range over [ma(ˈlini)] 'marine', [(ˈmanu)] 'animal', [(ˈmoa)] 'chicken', [(ˈlaː)] 'sun', and [(ˈia)] '3.SG', and the objects shown are [(ˈmea)] 'thing' in the two left panels and [(ˌmanu)] in the two right panels. Words with VV/Vː sequences were segmented as a single unit since it was not possible to segment between the vowels in a reproducible manner. Their F0 contours are thus plotted over only 10 time slices. Regardless of syllable shape of the subject or object, or if the speech rate is faster or slower, the low F0 level on the object is reached at the onset of the stressed mora, and the high F0 peak is at the offset of the stressed mora in [(ˈmanu)] or after.

There are three distinctive properties of the rising pitch accent when it is in the nuclear context relative to in the prenuclear context: (i) early peak alignment, (ii) potential truncation of the following fall, and (iii) potential high scaling of the peak that breaks downtrend. All we mean by 'truncation' here is that the fall is not realized in its entirety, all the way down. The peak delay sometimes seen in the prenuclear context is typically absent in the nuclear context: the alignment of the peak is strictly at the right edge of the stressed mora, or even slightly before, e.g., like in the sentence-final words of Figure 12.1. Earlier alignment of peaks in nuclear relative to prenuclear contexts is well-attested cross-linguistically (Ladd, 2008, p. 142) and has been attributed to tonal crowding from the following low boundary tone (Hualde, 2002, p. 107).

According to Jun and Fletcher (2014, p. 516-17), a '+' indicates a boundary between two tone bearing units, and "if the tune–text alignment is neither consistent nor distinctive, a bitonal LH symbol without a plus sign (e.g. LH*) should be chosen to represent a rising tone" rather than choosing between 'L+H*' and 'L*+H', see also Kuo and Vicenik (2012, p. 67) on Tongan. This was the reasoning behind the adoption of 'LH*' label for the rising pitch accent in Samoan in Orfitelli and Yu (2009), Yu and Stabler (2017), Yu (2021), since the L is typically aligned to the

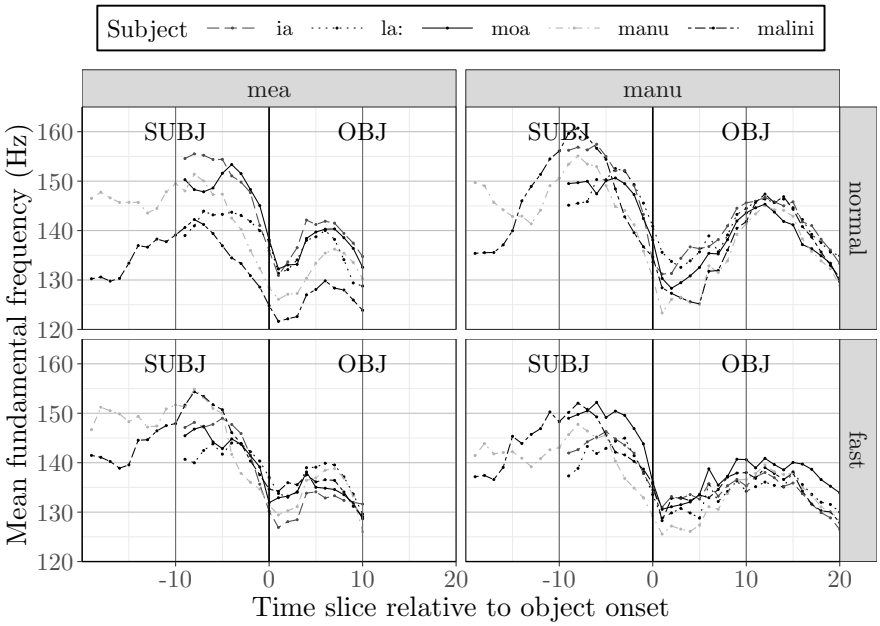


Figure 12.3: Individual F0 contours from la_m01 over varied subjects and objects in utterances of the sentence template (3) from the speech rate data set presented in Yu and Stabler (2017, (6)). These are exemplars of pitch accents one mora apart, with no intervening edge tones. The subjects range over [ma(ˈlini)] ‘marine’, [(ˈmanu)] ‘animal’, [(ˈmoa)] ‘chicken’, [(ˈlaː)] ‘sun’, and [(ˈia)] ‘3.SG’, and the objects shown are [(ˈmea)] ‘thing’ in the two left panels and [(ˌmanu)] in the two right panels.

beginning of the same TBU as the H. However, we use the 'L+H*' label here in an effort for standardization since: (i) the Samoan rising pitch accent has the same phonetic alignment as the IPrA 'L+H*' category (Hualde and Prieto, 2016, Figure 6), (ii) the focus in choosing between category labels for bitonal pitch accents seems to be on phonetic alignment of tonal targets rather than phonological association (Hualde and Prieto, 2016), and (iii) the broad category of 'L+H*' pitch accents has been refined into different phonetic alignments and phonological association not by differential diacritics in the pitch accent category label, but in differential specifications of primary and secondary phonological associations to TBUs and prosodic constituents (Prieto et al., 2005). In general, as long as the phonetic alignment and phonological association is specified in detail, the choice of category label becomes less fraught.

The earlier pitch accent peak alignment in the nuclear position might be a strategy for allowing time for the final low boundary tone to surface. But in fact, sometimes the fall to the final low boundary tone is not clearly visible in the F0 trace and sometimes it is additionally not clearly perceptible. This phenomena was typically observed in elicited declaratives in Calhoun (2015, §4.1) for IntP-final long vowels, see, e.g., [analeilaː] 'earlier' in Figures 2-4 in Calhoun (2015), repeated here as Figures 12.4a, b together with Calhoun's original intonational transcriptions. Due to an understanding of the nuclear (pitch) accent as typically marking the focus and the appearance of the final rise, nevertheless, on time expressions like [analeilaː] 'earlier' that were already given in the context, Calhoun (2015, p. 216, 219) proposes that the rise on the sentence-final word is due not to an L+H* pitch accent, but an L+H- phrase accent. Moreover, it is implied that the absence of a final fall is due to the lack of association of an underlying L% tone to a mora: "the H target of the phrase accent is associated with the end of the stressed mora . . . and there is no following mora on which to realise the fall, leaving an apparent final rise" (Calhoun, 2015, p. 216). A similar process, analyzed as L% truncation/deletion, occurs in Malagasy when stress is on the IntP-final syllable, see Aziz (2020*a*, §4.2) and Aziz (2020*b*, p. 17, 19).

We differ from Calhoun (2015) in maintaining that the IntP-final rise is due to an L+H* and not a separate kind of tonal event, i.e., it is not an L+H- phrase accent, see further discussion in §12.4.1. Here, we also show that a reduced IntP-final fall can surface not only with IntP-final long vowels (Figures 12.4, 12.14) but also with IntP-final diphthongizing VV sequences and devoiced/deleted vowels. Figure 12.5a shows an utterance of (5a) where (very late and steep) falls do occur on IntP-final /au/ and /ai/, while Figure 12.5b shows an utterance of the second part of (5b), where no fall is visible on IntP-final /ai/. The wide F0 range of the imperative in (5a) is due to the particular imagined context of utterance. This context was where the speaker is at the park with their niece, and the niece is running further and further away; the speaker is worried about heavy traffic on the street and tells her to come back to where they are.

The reduction/absence of a final F0 fall in the surface realization in Samoan can also happen

when the final vowel and/or syllable, is devoiced or deleted, e.g., in the deletion of the final vowel or syllable of [malaŋa] in Figures 12.23a, b and the deletion of the final vowel of [mamanu] in Figure 12.6, an utterance of (6). (The context for (6) is a polarity focus context. The speaker saw a picture of a woman working on a design and heard *Va'ai i le ata lea. 'O se mamanu lā e galue ai Malaga?* 'Look at this picture. Is one of those designs being worked on by Malaga?')

Due to the variability in the extent of reduction of the final F0 fall observed in the examples in this section and clear presence of a falling trajectory even if the fall is incomplete, we can rule out deletion of the L% as an analysis (Grabe, 1998, §3.1). However, elaborating on the suggestion by Calhoun (2015, p. 216), one way to analyze the distinction between a fully realized IntP-final fall and a reduced one is via a difference in association, like the "strong" vs. "weak" allophone analysis proposed by Pierrehumbert and Beckman (1988) for Japanese accentual phrase tones: the fall is reduced if the L% is associated only to the IntP-node but not a TBU.

(4) Utterance-final truncation in examples from Calhoun (2015)

a. AVO (Agent Verb Object) Calhoun (2015, (8c), Fig. 4)

ʔo si(ˈone) na to(ˈso-a) le (ˈmaea) (ˌana)(ˌlei)(ˈlaː)
TOP Sione PAST pull-ERG DET rope earlier
'It was Sione who pulled the rope earlier'

b. VAO (Verb Agent Object) Calhoun (2015, (8a), Fig. 2)

saː (ˈtoso) e si(ˈone) le (ˈmaea) (ˌana)(ˌlei)(ˈlaː)
PAST pull ERG Sione DET rope earlier
'Sione pulled the rope earlier.'

(5) Utterance-final truncation and diphthongizing VV sequences

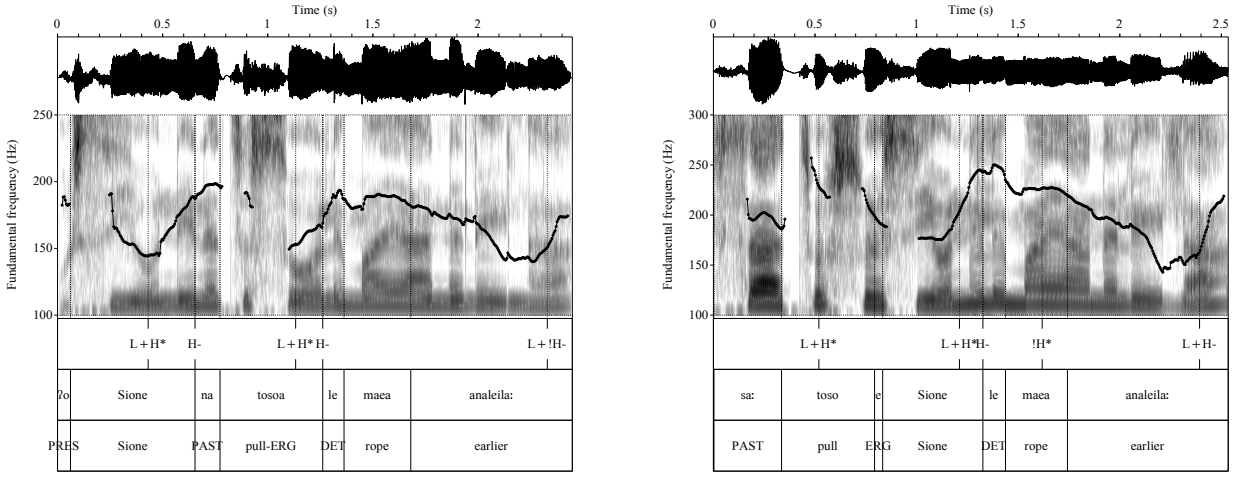
a. IntP-final /au/, /ai/

(ˈsuŋa), (ˈsau) (ˈtoe) (ˈfoʔi) (ˈmai)!
girl come again return DIR
'Hey, come back here!'

b. IntP-final /ai/

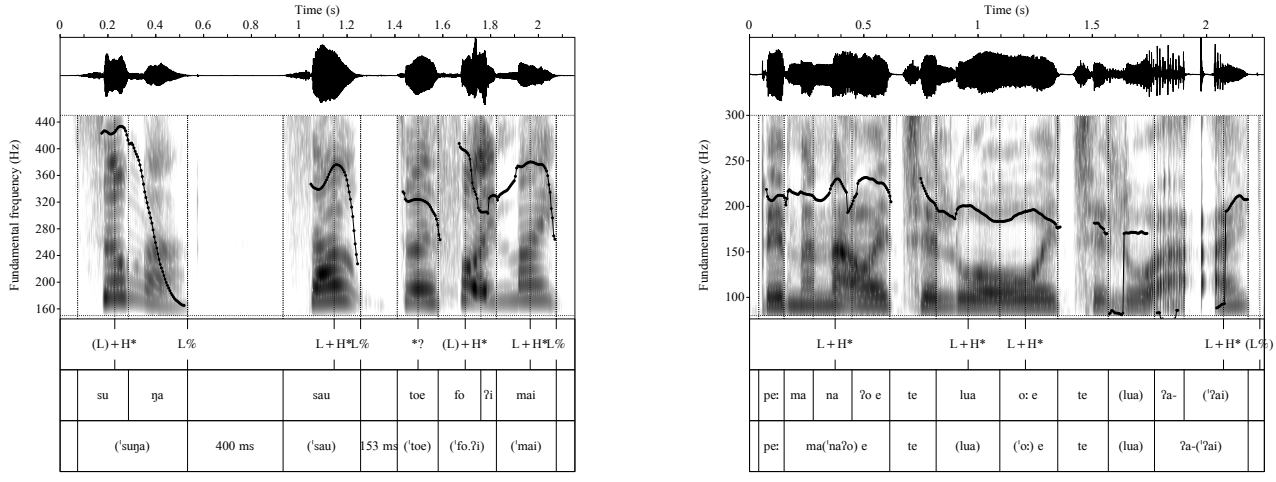
ʔua e fe(ˈsi.li) i lau u(ˈoː) peː ma(ˈnaʔo) ʔe te (ˈlua) oː
PERF 2.SG ask OBL 2.SG.POSS friend Q want 2.SG PRES 2.DU go.PL
ʔe te (ˈlua) ʔa-(ˈʔai)
2.SG PRES 2.DU RED-eat
'You ask a friend if they want to go get a feed with you.'

(6) Reduction of IntP-final fall in presence of IntP-final vowel deletion



(a) Agent Verb Object, Calhoun (2015, Fig. 4) (b) Verb Agent Object, Calhoun (2015, Fig. 2)

Figure 12.4: Examples (4a) and (4b) and accompanying transcriptions redrawn from Calhoun (2015). These exemplify truncation of the declarative-final fall in long vowels, as well as Calhoun (2015)'s analysis that the accent immediately following a sentence-medial H edge tone is suppressed (further discussed in §12.3.2).



(a) Lack of truncation of fall for IntP-final /au, ai/ (5a) (b) Truncation of fall for IntP-final /ai/ (5b)

Figure 12.5: F0 contours showing variable realization of IntP-final fall over final diphthongizing VV sequences, of (5a)from au_s02 in (a) and of (5b) from au_s11 for (b).

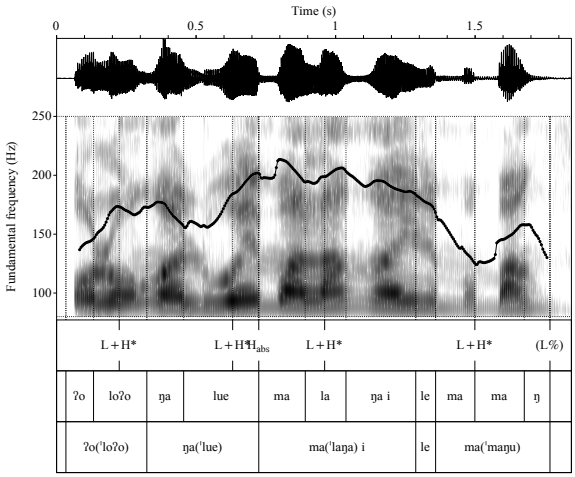


Figure 12.6: F0 contour of (6) showing utterance-final truncation of L% in the presence of final vowel deletion, from au_s04.

ʔoloʔo ŋalue Malaŋa i le mamanu
(ʔi(ˈoe) PRES work Malaga OBL DET design

‘(Yes.) Malaga is working on the design.’

Turning to the last distinctive property of the nuclear rising pitch accent—potential high scaling of the peak that breaks downtrend—an example of this phenomenon can be observed most dramatically on the final word /pea/ ‘pear’ in Figure 12.18 (see also Figure 3 in Calhoun (2015)), but also to a lesser extent in Figure 12.5b, from older speakers and au_s02 and au_s11. The downtrend-breaking phenomenon has been alluded to in early accounts of Samoan intonation. Churchward (1951, p. 9, 26) describes the last accent as “extra heavy” (and also mentions what seems like final deletion or devoicing): “The accent is extra heavy on the last word in a sentence, especially in public speeches. So heavy is this accent that the unaccented short vowel after it can scarcely be heard, and the word seems to end in a consonant.” While it’s not clear what Churchward meant by “extra heavy”, Mosel and Hovdhaugen (1992, p. 41, (2.33)) specifies that “the only obligatorily accented phrase is the last one of the sentence where the penultimate syllable is also characterized by an obligatory tone raising” and shows a final F0 peak breaking downtrend after a series of preceding F0 peaks in a schematic of an F0 contour for a declarative. The downtrend-breaking characteristic of the IntP-final F0 peak might be something that is disappearing in younger speakers, see Calhoun (2015, p. 220), and as also observed in younger speakers in the examples in this paper. Calhoun (2015) also showed that the high scaling and presence of the IntP-final pitch accent is not sensitive to focal conditions and transcribed 63/84 (75%) of elicited declaratives ending in temporal adverbials like (ˌana)(ˌlei)(ˈlaː) ‘earlier’ in (4) as having non-downstepped, i.e., downtrend-breaking

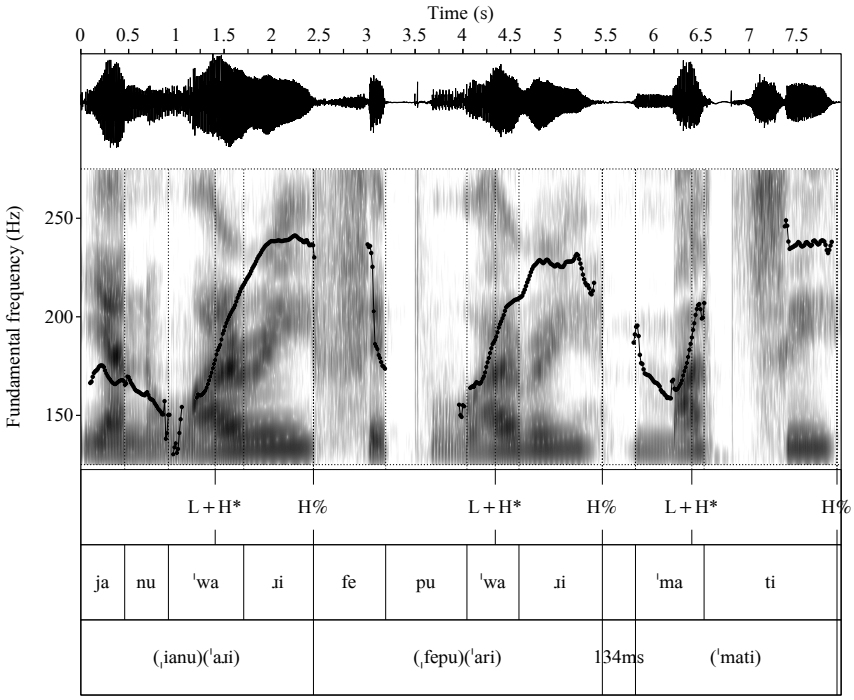
final rising pitch accents (Calhoun, 2015, (9)). Yet, all questions used as contexts for elicitation already included the temporal adverbial, so it was always given. The insensitivity of the nuclear (i.e., IntP-final) accent to focus conditions is raised as a puzzle by Calhoun (2015) et seq. Under the assumptions that the nuclear accent is expected to align with focus and that a pitch accent is necessarily prominence-lending, Calhoun (2015, p. 219) and Calhoun (2017, p. 20) propose that the nuclear L+H* accents are in fact "non-prominence lending phrase accents". In §12.4.1, we discuss how these assumptions tying (nuclear) pitch accents to focus/prominence have come into AM theory in practice, but are not in fact foundational to AM theory, see also Gussenhoven (2022).

12.2.2 Rising pitch accent followed by high plateau: L+H* H%

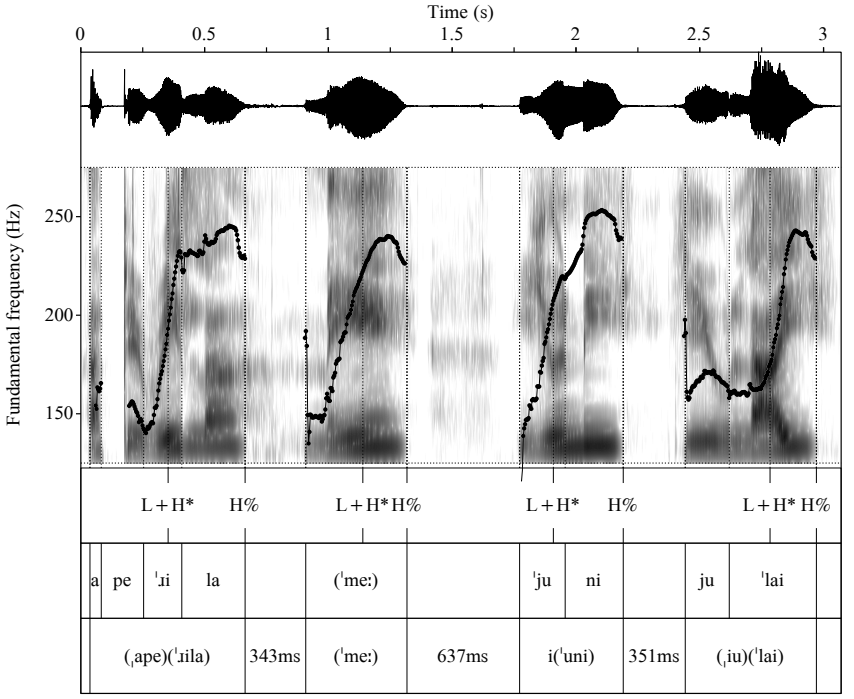
In nuclear position, a rising pitch accent over the stressed mora can also be followed by a high plateau in enumeration and as a continuation rise. Figure 12.7 displays, in two parts, the enumeration of the first seven months of the year. Examples of H% continuation rises are in Figures 12.16a and 12.17b. The enumeration of the first three months, which all end in (ˈCVCV), shows that the pitch accent rise over the penultimate stressed syllable can extend into the onset of the following syllable, but is level over the final vowel. Although there are slight falls visible in the F0 contour after the high plateau is reached when a pause immediately follows, the auditory percept is of a level high. The slight F0 falls might be a physiological consequence of a drop in sub-glottal pressure at the termination of phonation (Slifka, 2000, §6.2). For this speaker, the high F0 plateaus are around 240 Hz. The enumeration of [(ˈmeː)] 'May' and [(ˌiu)(ˈlai)] 'July' is of particular interest because the stressed syllables are bimoraic, over Vː and a diphthongizing VV sequence. The F0 rise over the final rather than the penultimate syllable of [(ˌiu)(ˈlai)] shows that the F0 rise occurs where the stressed syllable is and isn't fixed to be at the penultimate syllable. The F0 rise over the long vowel of [(ˈmeː)] is roughly over the first half of the vocalic material. The slope of the F0 rise over /ai/ in [(ˌiu)(ˈlai)] tracks with the slope of the F1 and F2 transitions and terminates in a high plateau just as the maximum distance between F1 and F2 is reached. This phonetic alignment is consistent with an analysis where the TBU is the mora, and the pitch accent is phonologically associated to the penultimate mora, while the H% boundary tone is phonologically associated or aligned to the final mora.

12.2.3 Rising pitch accent followed by superhigh: L+H* ˆH%

In addition to being followed by a high plateau, a rising pitch accent in nuclear position can also be followed by a continued rise in echo questions, e.g., in contexts of confirmation and disbelief like (7) and Figure 12.8. The intonation of echo questions was first mentioned in Orfitelli and Yu (2009, p. 8-9, Figure 8b), where it is stated that la_m01 could produce a 'confirmation contour'

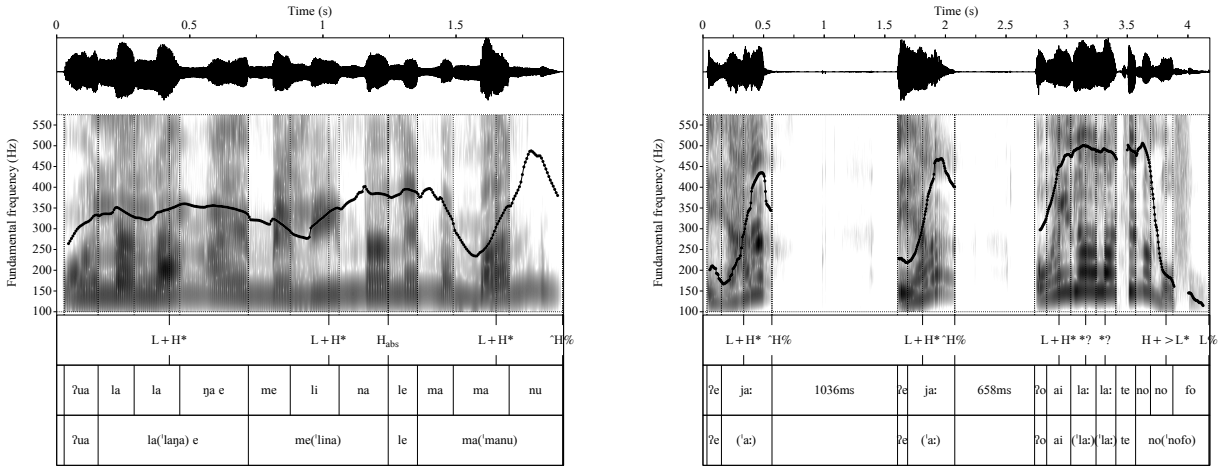


(a) January, February, March



(b) April, May, June, July

Figure 12.7: Enumeration of first seven months showing H%, broken for visibility into two sub-figures with the first three months in (a) and the next four after that in (b), from au_02, uttered out-of-the-blue.



(a) Continued rise over ma(ˈmanu) (b) Continued rise over e aː

Figure 12.8: Intonational contours of echo questions under disbelief contexts, (a) an utterance of (7a) from au_s2, and (b) an utterance of (7b) from au_s11 .

transcribed as 'H- H%', which has the same final high F0 plateau of the L+H* H% of enumeration described in this section. It was also stated that both this contour as well as the typical falling interrogative intonation could appear in the same echo question contexts.

(7) Echo questions, disbelief

a. Context: Someone tells you that your friend, Melina, is weaving a design. You can't believe it and you ask again about it.

E aːʔ (ˈsole), ʔou te (ˈleː) (ˌtali)(ˈtonu). ʔua lalaŋa e Melina ∅ le mamanu?
PRES what man 1.SG PRES NEG believe PERF weave ERG Melina ABS DET design
'What? Man, I don't believe it. Melina wove the design??'

b. Context: Ioane, a friend of yours, had told you that he would marry Melani, who is his girlfriend of many years now. A mutual friend tells you that Ioane will marry another girl named Manaia next month. This seems impossible to you because, because they've never gotten along. Very surprised, ask who he is getting married with.

E aːʔ E aːʔ ʔO ai laː laː te nonofo?
PRES what PRES what TOP who PART 3.DU PRES RED.marry
'What? What? So who two are getting married?'

In recordings with speakers from Auckland, we observed both typical falling interrogative intonation, e.g., the final question of (7b) shown in Figure 12.8b, as well as a potentially distinct

rising contour from the one observed in enumeration, e.g., in the final question of (7a) that is shown in Figure 12.8a. Like in the F0 contours for enumeration, the F0 drop seen before pauses does not create a falling percept and is likely due to physiological consequences of a cessation of phonation. Speaker au_s2 found both the typical falling interrogative intonation as well as the rising intonation produced in Figure 12.8a appropriate in the context. A typical expression in echo question contexts is *E ā?* 'What?', as in (7b) in Figure 12.8b, and this is commonly uttered with a sharp rise that continues across the long vowel. The sharp slope of rising contour at the end of Figure 12.8a also continues into the final syllable of ma(ˈmanu).

We tentatively propose that this sharp rising contour is phonologically distinct from the rising-to-plateau contour observed in enumeration because of the continued sharp rise beyond the stressed mora and because of the typical superhigh F0 range reached. In the same speaker, the rise is from 235 Hz to 484 over 200 ms in the disbelief context for ma(ˈmanu) in Figure 12.8a, a slope of 1.25 Hz/ms. In comparison, the rise is from 141 Hz to 231 over 184 ms in the enumeration of (ˌape)(ˈɹila) in Figure 12.7a, a slope of 0.49 Hz/ms. Note also that the final rise that can occur in a disbelief context is distinct from a final rise seen in truncation in declaratives because there is no deletion or devoicing of the final mora in the disbelief context and the rise continues over the final mora into a superhigh range.

12.2.4 Early falling pitch accent followed by plateau: H+>L* L%

In Samoan, as in many other languages, nuclear contours permit a wider range of pitch accents than in pre-nuclear position. Falling pitch accents are restricted to nuclear position and are followed by a low plateau. A change in illocutionary force from a declarative to a polar interrogative can be indicated solely by the use of the H+>L* L% nuclear contour rather than the L+H* L% one. The F0 contour for (8) shown in Figure 12.9 was uttered in an information-seeking context, in which the speaker has entered a store that they have never been in before and politely asks if they have melons. The H+>L* L% falls on the utterance-final politeness marker $((_{\text{ˌ}}\text{faʔa}))_{\omega}((_{\text{ˌ}}\text{mole})(\text{ˈmole}))_{\omega}$ 'please'.

(8) Polar interrogative

e (ˈmau)a ni tou me(ˈleni) $((_{\text{ˌ}}\text{faʔa}))_{\omega}((_{\text{ˌ}}\text{mole})(\text{ˈmole}))_{\omega}$[6]
PRES get DET 2PL.POSS melon please

'Have you all got some melons, please?'

[6]Whether there is primary or secondary stress on faʔa prefix is unclear, see Zuraw et al. (2014, p. 285, fn. 4), but there is evidence for treating the prefix as a separate prosodic word (Zuraw et al., 2014, p. 311).

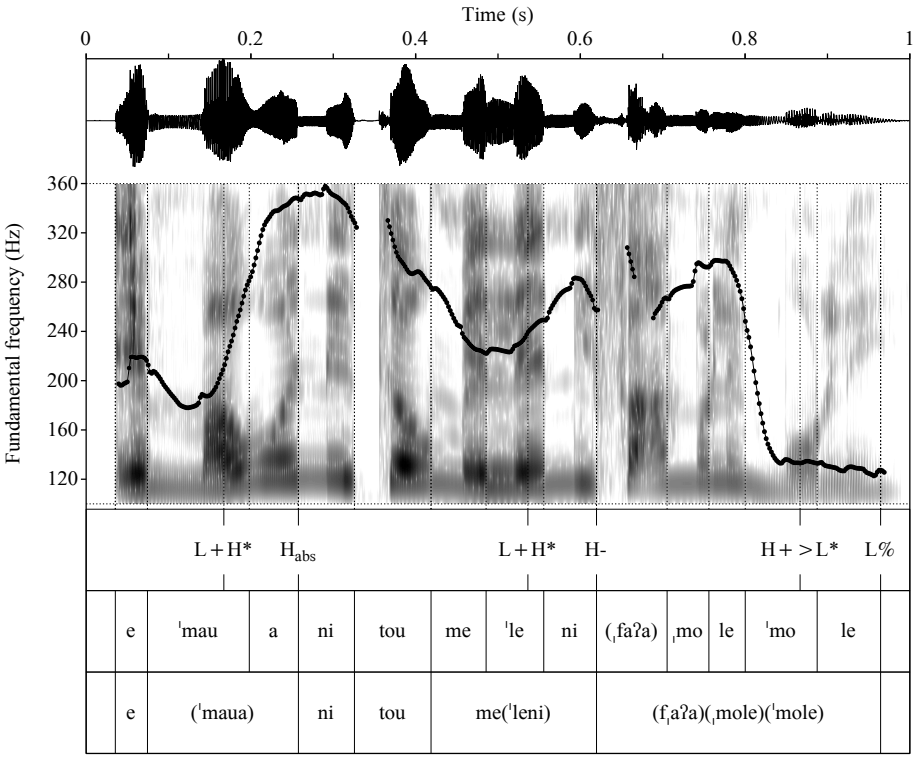
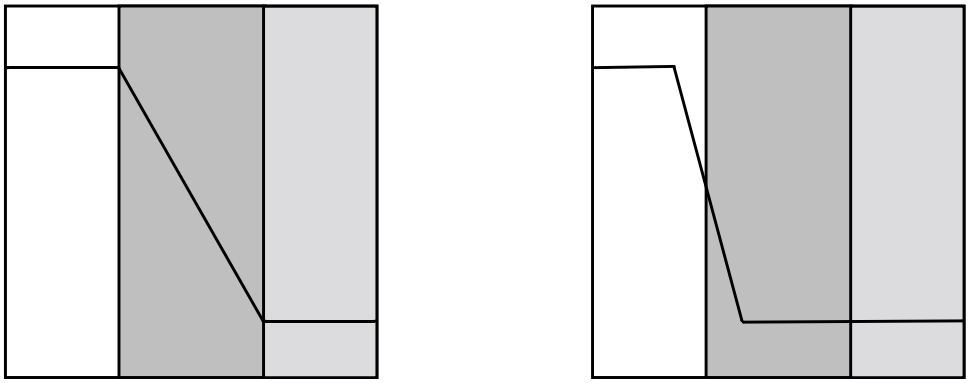


Figure 12.9: F0 contour for an information-seeking polar interrogative (8) from au_s11 exemplifying H+>L* L% nuclear contour

The nuclear accent in interrogatives has had a number of different analyses. Mosel and Hovdhaugen (1992, p. 42) states that "usually the tone is pretty low on the two to three last syllables" and therefore that there is no obligatory final accent. Orfitelli and Yu (2009, p. 5, Figure 6) proposes instead that there is in fact a nuclear pitch accent in interrogatives, but of a different type than the rising pitch accent found in declaratives, a !H*. In contrast, Calhoun (2015, p. 219) proposes that it is not a pitch accent but a final H+L- phrase accent, following the same reasoning as the proposal of a final L+H- phrase accent as opposed to an L+H* nuclear accent mentioned in §12.2.1, i.e., under the assumption that "a nucleus should always be informationally salient" (Howard, 2018), see also §12.4.1. Howard (2018) builds on this proposal and examines the production of: (i) wh- and polar-questions recorded for eliciting responses in Calhoun (2015) like (4) under broad focus, agent wh-focus, object wh-focus, contrastive agent focus, and contrastive object focus, and (ii) polar questions elicited under subject and object focus, about half of which ended in locative or temporal adverbials. Howard (2018) finds that regardless of the focus conditions, whether or not the sentence ended with an adverbial or not, and whether or not the interrogative was a wh- or polar interrogative, the final word was invariably "characterized by a high pitch plateau which begins at the H- before the final ip and extends to the H+L-" and that the "the alignment of the H+L- is consistent: the H aligns with the antepenultimate mora and the L with the penultimate mora". Howard (2018)'s schematic of a H+L- (labeled as H+L*) shows a high on the mora preceding the stressed mora and a fall over the stressed mora to a low on the posttonic mora, the same interpretation proposed in Hualde and Prieto (2016, Figure 6) for H+L*. A re-drawing of the schematic for H+L-/H+L*, with a repetition of our schematic for H+>L*, for comparison, is shown in Figure 12.10.

Our current proposal that the nuclear tune is an early falling pitch accent followed by a low plateau (H+>L* L%) differs from Howard (2018)'s proposal due to the observations that: (i) the location of the fall and following L is sensitive to the position of primary stress in the interrogative-final word rather than being positionally defined relative to the right edge of the IntP, and (ii) the fall is not typically over the penultimate mora, as suggested by a H+L-/H+L* categorization, but rather over the mora preceding the stressed mora and possibly into the onset but not the nucleus of the stressed mora (or only in the first part of the nucleus, if no onset is present). For example, in Figure 12.9, the sharp fall over the final word [((ˌfaʔa))$_\omega$((ˌmole)(ˈmole))$_\omega$] 'please' starts in the vowel preceding the primary stressed penultimate mora [mo] and then ends at the offset of the onset [m], reaching a low plateau over the nucleus [o] through the end of the utterance. The resulting percept is not of a sharp fall over the stressed mora, but rather of a low plateau—consistent with the differentiation between a level and contour tone depending on the timing of F0 movement relative to syllable structure studied in House (1996, 1990). If F0 movement is weighted differently in perception between the onset and the rime, then distinguishing between the rime and the entire

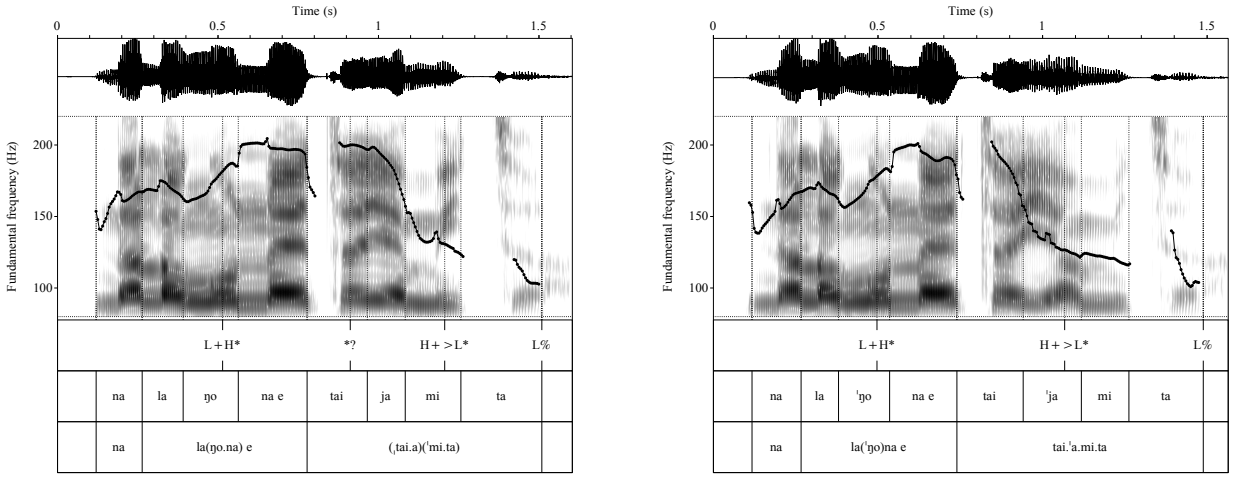


(a) H+L*/H+L- (Hualde and Prieto, 2016, Howard, 2018) (b) H+>L*

Figure 12.10: Comparison of H+L- alignment in Howard (2018) to H+>L* alignment proposed here. The H+L- schematic looks like the Hualde and Prieto (2016, Figure 6) schematic for H+L*. The fall in H+>L* alignment begins in the pretonic mora but also extends slightly into the stressed mora since the fall can sometimes extend into the syllable onset, see the main text for details.

mora or syllable as a TBU, e.g., as in Gussenhoven (2018, p. 401) and Gussenhoven (2022, §1.2), becomes important in adjudicating hypotheses about tonal alignment and association. Howard (2018) presents two sample figures of intonational contours, one with the final word [si(ˈamu)] 'jam', and one with the final word [(ˌana)(ˈpoː)] 'last night', that also display this timing. On [si(ˈamu)], a sharp fall from a high occurs on the [i] and a low plateau (with a slight fall) is reached and sustained in the stressed [a] through the end of the utterance. There is no onset in the stressed mora, but the F0 "elbow" is early on in the vowel. In [(ˌana)(ˈpoː)], there is still high F0 on the [a] preceding stressed [poː], but a low plateau over the primary stressed [oː]; the sharp fall occurs over the onset [p] of the stressed mora (although it is not clearly visible since the p is a plosive).

Evidence that the location of H and L tones in the interrogative-final fall is not fixed positionally to the antepenultimate and penultimate moras, respectively, but rather, is determined by stress position is shown in Figures 12.11, 12.12, and 12.13. These display comparisons of F0 contours of the sentences in (9a) and show that the fall is on the preceding mora to the stressed mora and then a low plateau is reached in the stressed mora. The use of loanwords and code-switching allows primary stress position to be pushed back to the antepenultimate syllable. Figure 12.11 shows the polar interrogative (9b) with penultimate stress in Figure 12.11a and antepenultimate stress in Figure 12.11b, see also discussion of(2). With the penultimate stress pattern [(ˌtaia)(ˈmita)], a sharp fall occurs over the [a] preceding stressed mora [mi] and then a low plateau occurs over stressed mora and the following mora, with a slight fall in the final mora. With the antepenultimate stress pattern, [taiˈamita], the sharp fall occurs over [tai] immediately preceding stressed mora [a], with a low plateau over the stressed mora and following moras, with a slight fall in the final mora. The



(a) Antepenultimate primary stress (b) Primary stress

Figure 12.11: F0 contours for polar interrogative with penultimate and antepenultimate primary stress on 'diameter' (9b) from la_m01.

generalization shown is that the sharp fall occurs in the syllable preceding the stressed mora, and then a low plateau (possibly with a slight decline) occurs over the stressed mora and afterwards.

(9) Polar interrogative examples

a. Context: there was a loud noise.

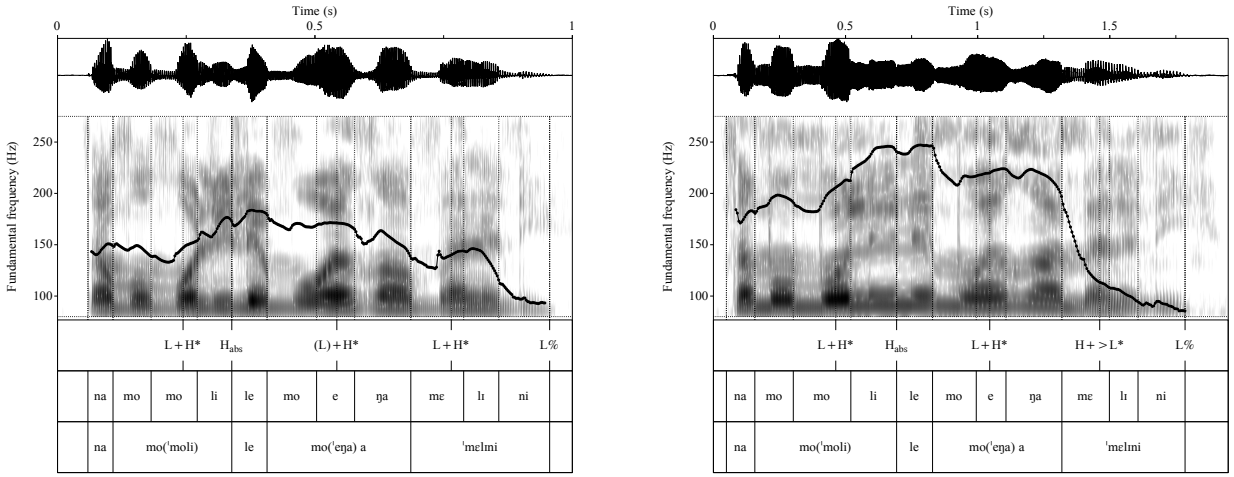
b. na laŋona e Taiamita?
PAST hear ERG Diameter
'Did Diameter hear it?'

c. Contexts: *'O ā ou mea na fai*? 'What did you do?' (declarative), *'Ua uma le galuega!* 'The work is finished.' (interrogative)

na mo(ˈmoli) le mo(ˈeŋa) a <Melanie>./?
PAST send DET bed GEN Melanie
'I sent Melanie's bed. / Did you send Melanie's bed?'

d. Context: An unbiased disjunctive polar interrogative, in the imagined context in which the speaker is having friends over for lunch and is asking if one of them would like taro or rice.

ʔe te ma(ˈnaʔo) i le (ˈtalo) (ˈpoʔo) le a(ˈlaisa)?
PRES 2.SG want OBL DET taro DISJ DET rice
'Do you want taro or rice?'



(a) Declarative (b) Polar interrogative

Figure 12.12: F0 contours comparing declarative and polar interrogative in (9c) with antepenultimate stress on sentence-final *Melanie*, from au_s4.

Figure 12.12 shows (9c) uttered as a declarative and as a polar interrogative.[7] The higher F0 range in the interrogative relative to the declarative is typical for this speaker, as also observed in Orfitelli and Yu (2009) for la_m01, but Howard (2018) reports no such range difference in their data. The nuclear rising pitch accent in the declarative over Melanie is uncharacteristically late, with the rise starting at the end of the onset of the stressed initial syllable and the peak in the posttonic syllable. The lateness might be due to tonal crowding between the initial stress and preceding stress on [mo('eŋa)]. In the polar interrogative, the sharp fall occurs over the onset [m] of the initially stressed syllable and falls to a sustained low plateau (with slight fall) afterwards. Much like in the declarative, the sharp fall is later than expected and starts in the onset of the stressed syllable rather than in the preceding syllable. This may again be due to tonal crowding.

Finally, Figure 12.13, of (9d), shows that the sharp fall is over the prenantepenultimate mora (not the antepenultimate mora) and following onset and that the low plateau starts on the stressed antepenultimate vowel rather than the penultimate mora, as proposed in Howard (2018), when the interrogative-final word, i.e., a('laisa) 'rice', has a diphthongizing VV stressed syllable with stress on the antepenultimate mora.

Since the H+>L* pitch accent already has a low target, interrogatives ending in long vowels don't show the same pattern of truncation seen in declaratives, in which a final fall may not be observable on the surface. For example, Figure 12.14 compares a declarative vs. a polar interrogative rendition of (9c) with the final word [(ˌene)('leː)] 'Henry' in place of *Melanie* and similar durations

[7] Unfortunately the speaker had a stuffy nose during the elicitation session of (9c) and related sentences, resulting in obstructed airflow during nasals and concomitant F0 perturbations.

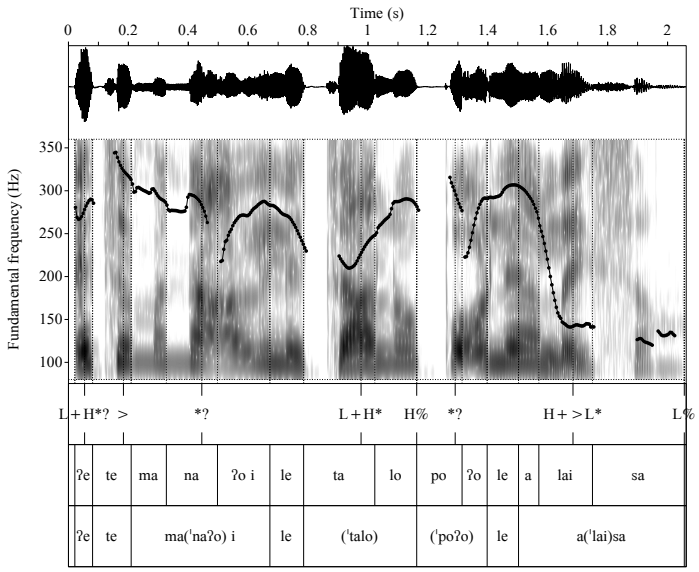
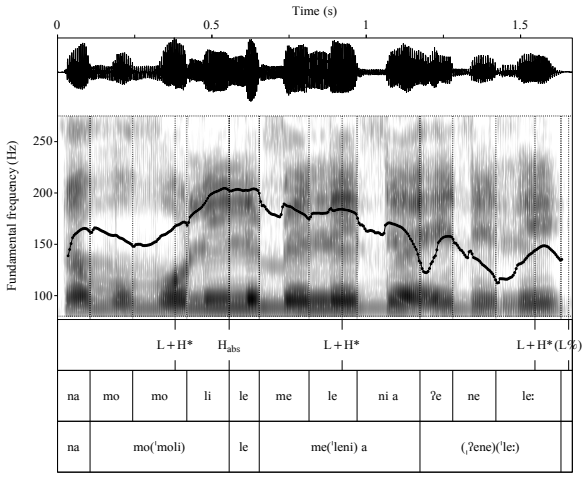


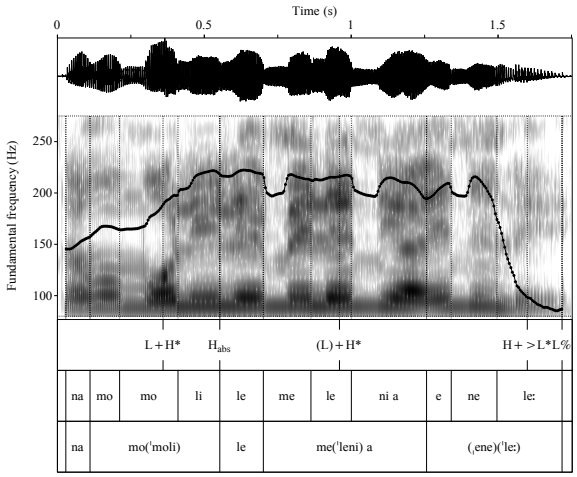
Figure 12.13: F0 contour for unbiased disjunctive polar interrogative in (9d) from au_11, showing sharp fall beginning in preantepenultimate syllable and ending in a low plateau in the rime of the stressed antepenultimate mora.

for the final syllable (22ms in the declarative and 21ms in the interrogative). While the declarative ends with the perception of a rise on (ˌene)(ˈleː), the interrogative ends with the perception of a low plateau. Similarly, Figure 12.15 shows that failure of the final vowel in [pe(ˌɹesi)(ˈkeŋe)] (a rendition of (9c) with the final word 'presidents' in place of *Melanie* and the final mora of [(ˌene)(ˈleː) ma(ˈmaŋu)] (a rendition of (9c) with the name *Enele Mamanu* in place of *Melanie*) to surface at the end of the interrogative does not prevent the L target of the H+>L* from being reached, although no low plateau follows. Figure 12.15a also shows an example of devoicing IntP-medially, of [i] immediately preceding the primary stressed mora in pe(ˌɹesi)(ˈkeŋe). Finally, Figure 12.15b shows that the overall high F0 range plateau in interrogatives that can appear to stretch over the entire utterance after the initial pitch accent and/or H edge tone can show a downtrend over longer utterances. More work is needed to investigate the domain of the high F0 range plateau.

In addition to appearing in interrogatives, the nuclear early falling accent tonal contour can appear sentence-medially, typically followed by a silent juncture, as mentioned in Yu and Stabler (2017, p. 26-27, Figure 9a) and Calhoun et al. (2019, p. 9-12, Figures 1-33). Calhoun et al. (2019) found it under three information structural conditions: on the object in VOS broad focus, on the verb in subject-clefted SVO sentences, and on the object in object-clefted OVS sentences. Figures 12.16 and 12.17 show examples in which speakers varied between a sentence-medial early falling accent contour and a high continuation rise in adjacent repetitions of reading a short paragraph or a

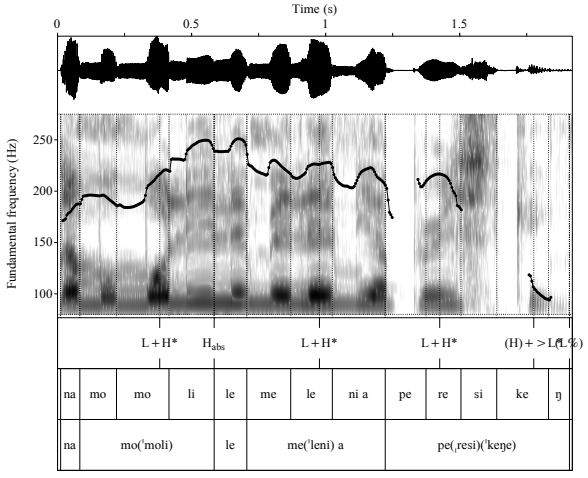


(a) Declarative

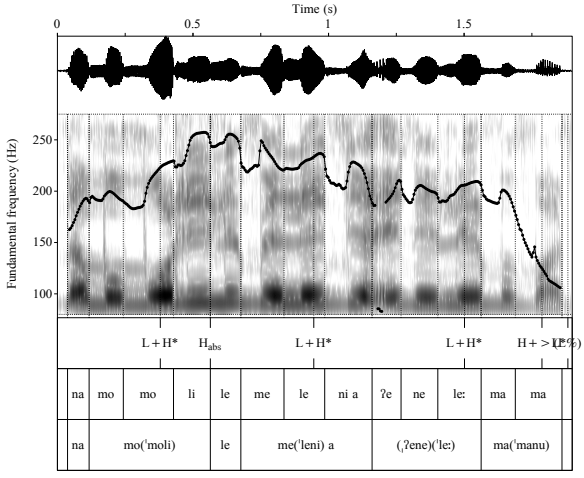


(b) Polar interrogative

Figure 12.14: F0 contours for declarative and polar interrogative renditions of utterances of (9c) from au_02 with sentence-final name (ˌene)('leː) 'Henry'

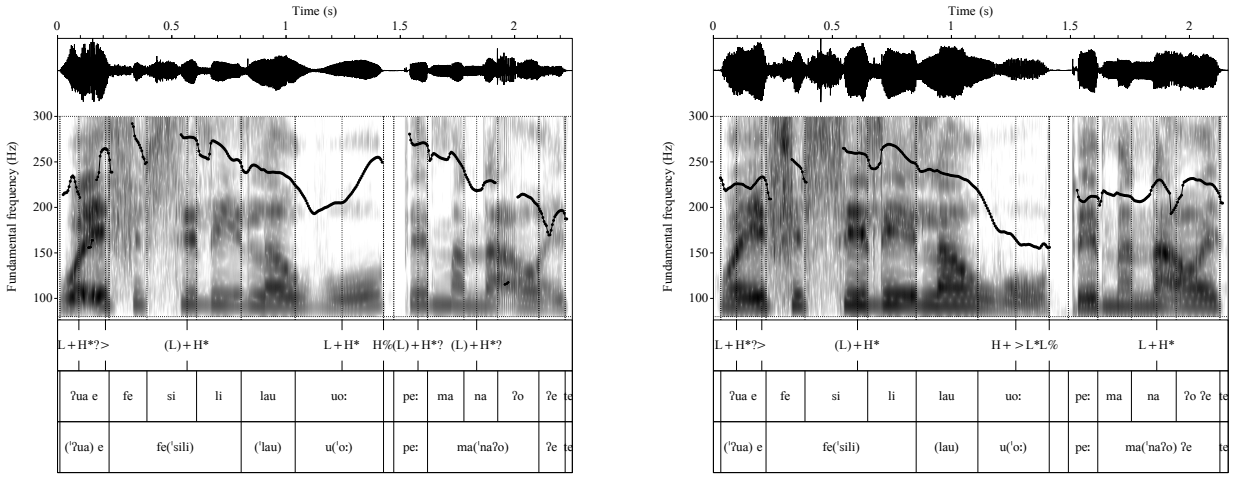


(a) pe(ˌɾesi)('keŋe)



(b) (ˌene)('leː) ma('maŋu)

Figure 12.15: F0 contours for polar interrogative renditions of utterances of (9c) from au_02 with sentence-final pe(ˌɾesi)('keŋe) 'presidents' and (ˌene)('leː) ma('maŋu) 'Enele Mamanu', showing low F0 reached despite segmental deletion/devoicing.



(a) With final H+L* L%

(b) With final L+H* H%

Figure 12.16: F0 contours for the first part of (5b), from two adjacent repetitions of the context read by au_s2, showing variability in choice of sentence-medial nuclear contours.

story. See (14) in §12.3.2 for Calhoun (2017)'s proposal about information structural-conditioning of whether a continuation is executed with an L or H tone. Figure 12.16 shows the first part of (5b). Figure 12.17 shows two renditions of the first several words of the North Wind and the Sun fable in (10), in which there is (almost) no silence after the continuation rise/fall.[8]

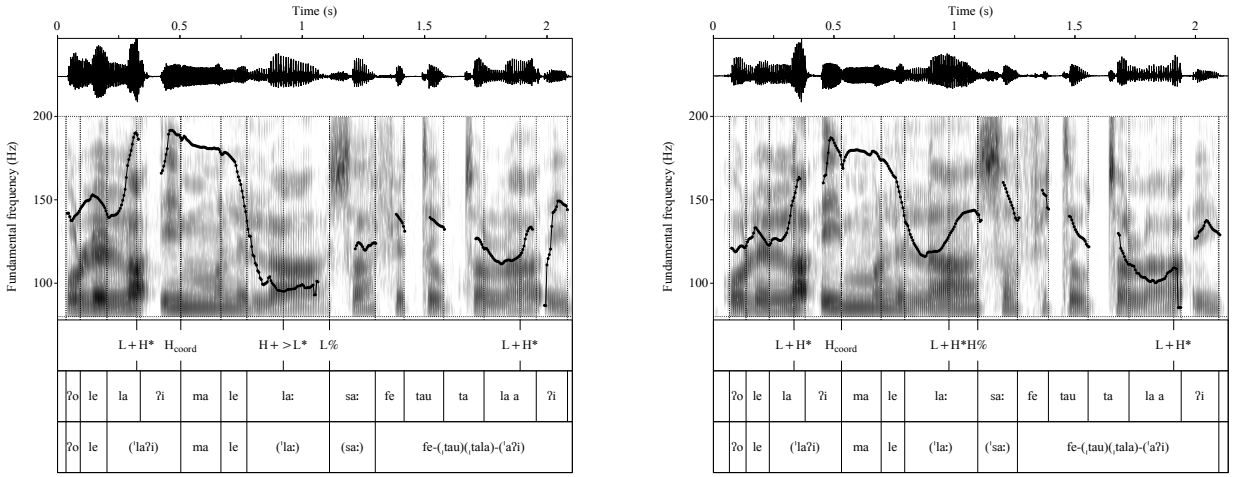
(10) Beginning of North Wind narrative

ʔo le (ˈlaʔi) ma le (laː) saː fe-(ˌtau)(ˌtala)-(ˈaʔi)...
TOPIC DET.SPEC wind CONJ DET.SPEC sun PAST disputing

'The wind and the sun were disputing...'

We have also observed repeating stretches of intonational phrases ending in early falling tonal contours in retellings of the Pear Story (Chafe, 1980). Speakers watched a silent film and then told the story as they recalled it. Figure 12.18 shows an example of a series of three early falling tonal contours before ending with a typical declarative rising pitch accent fall contour. In this context, we hypothesize that the early falling tonal contours could be instances of "uptalk" (Warren, 2016), in which, very roughly speaking, declaratives are produced with an intonation that can make them

[8]The sentence-medial edge tone in Figure 12.17b might be good candidates for an !H edge tone or an H- as opposed to an H% tone due to the relatively low F0 of the edge tone and lack of a following pause, but the jump in F0 creates a strong juncture percept; the labeled L% edge tone in Figure 12.17b similarly could be a candidate for a L-, although again, the F0 jump imparts a strong juncture percept. For both figures, we are not entirely clear yet what tonal events might be conditioning F0 movements at the beginning of the second phrase (starting with peː in Figure 12.16 and saː in Figure 12.17), including potential pitch reset and/or pitch accents on bimoraic peː and saː.



(a) With final L+H* H% (b) With final H+L* L%

Figure 12.17: F0 contours for the first several words in North Wind and Sun given in (10), from two adjacent repetitions of the story by smo_s4.

sound like interrogatives. However, since interrogatives in Samoan are characterized by a final fall rather than a final rise, "uptalk" takes the form of falling declaratives rather than rising ones. Falling "uptalk" has also been observed in Majorcan Catalan, in which polar interrogatives have a falling contour (Armstrong and del Mar Vanrell, 2016).

(11) Recalling of pear story

a. Up to 2.5s

ʔi(ˈoe) ʔo le (ˈata) (ˈsa:) (ˈʔou) (ˌmata)(ˈmata) (ˈai)
yes TOP DET.SPEC movie PAST 1.SG watch LOC
'OK, the movie that I was watching.'

b. Up to 5s

o se (ˈata) ʔo se ta(ˌma:)(ˈloa)
TOP DET.NSPEC movie GEN DET.NSPEC man
'It's a movie about a man.'

c. To end

(ˈsa:) (ˈʔou) va(ˈʔai) ⟨disfl⟩ (ˈua) (ˈʔaʔe) i (ˈluŋa) o le (ˈpea)
PAST 1.SG see PERF climb OBL up GEN DET.SPEC pear
'I saw. . .he'd been climbing up the pear tree.'

The particular utterance in Figure 12.18 is also of interest because it includes examples of the early falling contour over VV sequences. The tonal contour over the first word ʔi(ˈoe) 'yes' shows

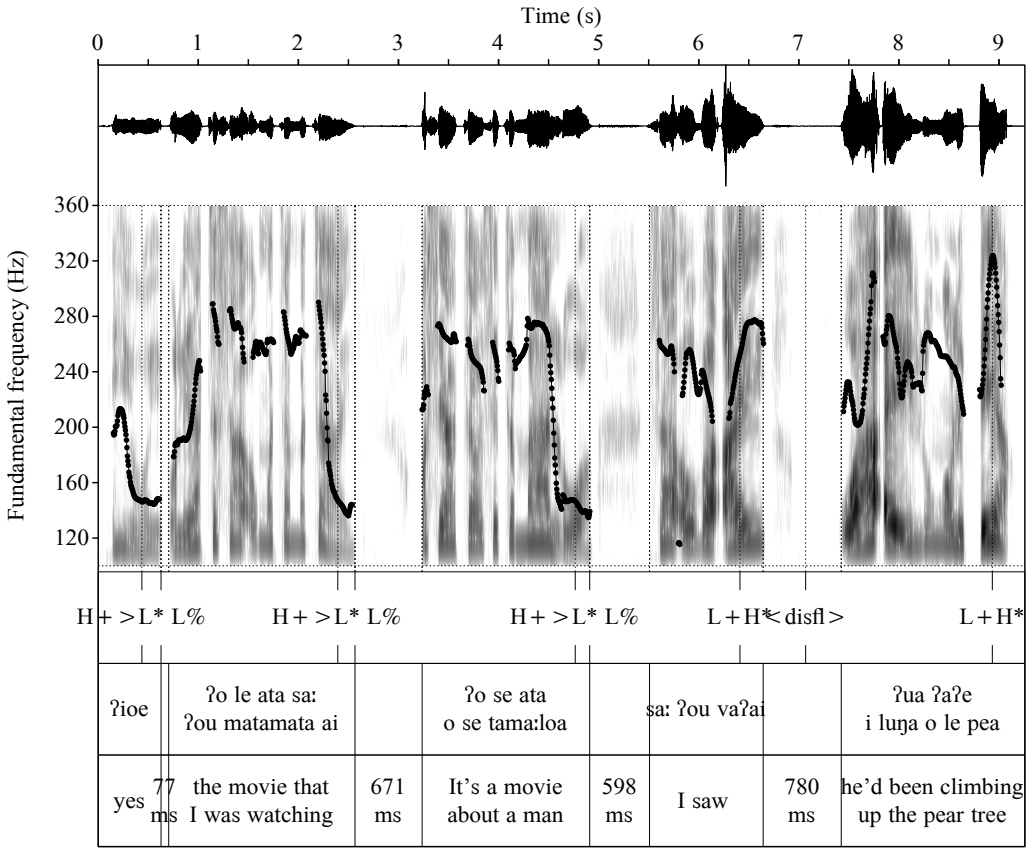
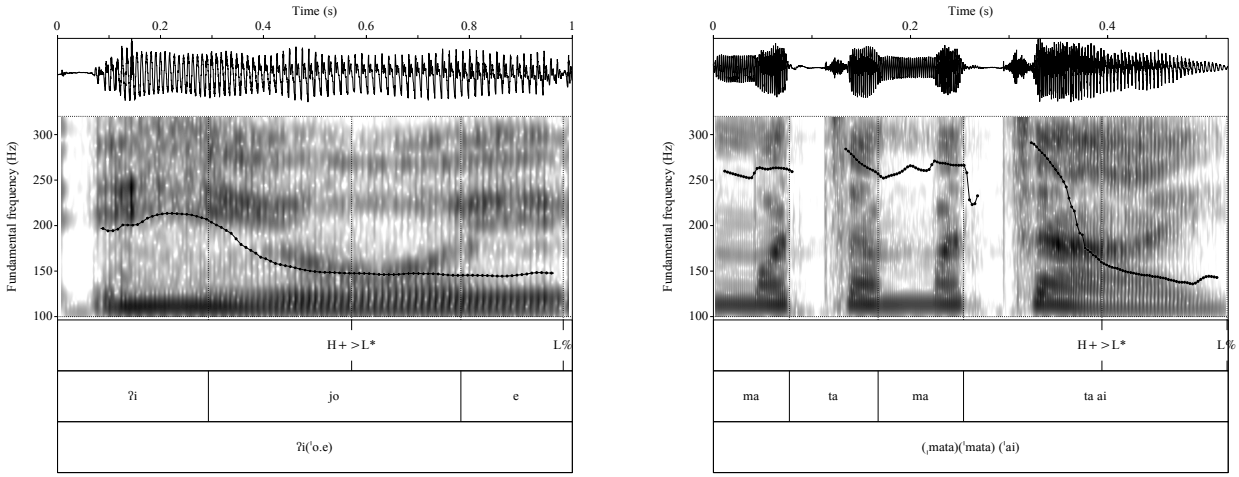


Figure 12.18: F0 contours showing a series of early falling accent tonal contours near the beginning of au_11 recalling and retelling the Pear Story.



(a) Non-diphthongizing /oe/ (b) Diphthongizing /ai/

Figure 12.19: Close-up F0 contours of IntP-final words with early falling accent over VV sequences from Figure 12.18 showing F0 alignment.

that the low plateau can be over two syllables, that is, the non-dipthongizing [oe]. (A plateau over two syllables also occurs over non-dipthongizing [oa] of ta(ˌmaː)(ˈloa) at the end of the third IntP). It also shows that word-initial [i] in ʔi(ˈoe) is not uttered as a glide (and it often is), but as a full vowel. This provides a mora preceding the stressed mora to which the H of the H+>L* can be realized. The phrase immediately following ʔi(ˈoe) shows the early falling contour over diphthongizing (ˈai). The F0 elbow occurs well before F2 diverges with the transition into [i], showing that the low F0 plateau starts in [a]. These phonetic data are consistent with an analysis where the TBU is the mora. Additional phonetic alignment data would be helpful for developing an analysis of the phonological association and alignment of the bitonal pitch accent.

12.3 Sentence-medial edge tones

The nuclear contours discussed in §12.2 already introduce two pitch accents (L+H*, H+>L*) and three boundary tones (L%, H%, ˆH%) and we saw some examples of the nuclear contours appearing sentence-medially as well as sentence-finally. But the distribution of sentence-medial edge tones in Samoan—the topic of this section—indicates that we miss generalizations if we don't consider other factors that may condition the appearance of tones, in addition to word-level stress and prosodic constituent edges. Recall from §12.1 that we use the term 'edge tone' as described in Yu (2021, p. 293,§6.3.2)—not as a term interchangeable with (prosodic) boundary tone, but as a purely descriptive term to characterize tones whose phonetic realization is conditioned

by morphosyntactic/prosodic word edges. This leaves open where the tone is introduced in the grammar, e.g., it could be inserted in building a higher-level prosodic constituent like a phonological phrase, or it could be inserted as reflex of morphosyntactic spell-out, or a lexical tone, etc. We review the empirical data on the distribution of sentence-medial high and low edge tones in §12.3.1. Then, in §12.3.2, we show how hypothesizing that there are multiple grammatical sources of these edge tones—namely, morphosyntactic spellout as well as tones inserted at prosodic phrase edges—rather than a single source, provides a way to resolve apparently conflicting data about the distribution of the edge tones. Finally, §12.3.3 contributes some new empirical data bearing on our understanding of these edge tones.

12.3.1 The distribution of sentence-medial edge tones

It is not disputed that high sentence-medial edge tones reliably occur in certain syntactic configurations in Samoan. By "reliably", we mean that it is clear that the appearance of high sentence-medial edge tones is correlated to a high degree with (but not necessarily caused by) these configurations. Orfitelli and Yu (2009), Yu (2011), Yu and Stabler (2017) showed that H edge tones reliably occur at the right edge of the first conjunct in coordinated structures, and Calhoun (2015), Yu and Stabler (2017) showed that they reliably occur at the right edge of fronted arguments in SVO and OVS word orders, e.g., in (12). The F0 contour for (12) is shown in Figure 12.20. In addition to an H appearing in the position consistent with where an absolutive case marker would be expected, at the right edge of the verb [(la(ˈlaŋa))$_\omega$-((ˈina))$_\omega$] (parsing into two prosodic words (ω) based on Zuraw et al. (2014)), an H appears at at the right edge of the fronted nominal. It also appears reliably at the right edge of a first conjunct (Orfitelli and Yu, 2009, Yu, 2011, Calhoun, 2017).[9]

(12) Non verb-initial SVO order with fronted subject

ʔo *e le ma(ˈlini) na (la(ˈlaŋa))$_\omega$-((ˈina))$_\omega$ ∅ le ma(ˈmanu).
TOPIC ERG DET marine PAST weave-INA ABS DET design

'It was the marine that wove the design.'

Yu (2011), Yu and Stabler (2017), Yu (2021) provided evidence that a high sentence-medial edge tone reliably appears at the right edge of whatever immediately precedes the absolutive argument. In addition, Yu and Stabler (2017) provided evidence that the appearance of the high edge tone before absolutive arguments is insensitive to the prosodic factors of changes in speech rate and

[9]Since disjunctions are bimoraic and initially stressed, it is difficult to confirm that Hs appear at the right edge of fit disjuncts as well.

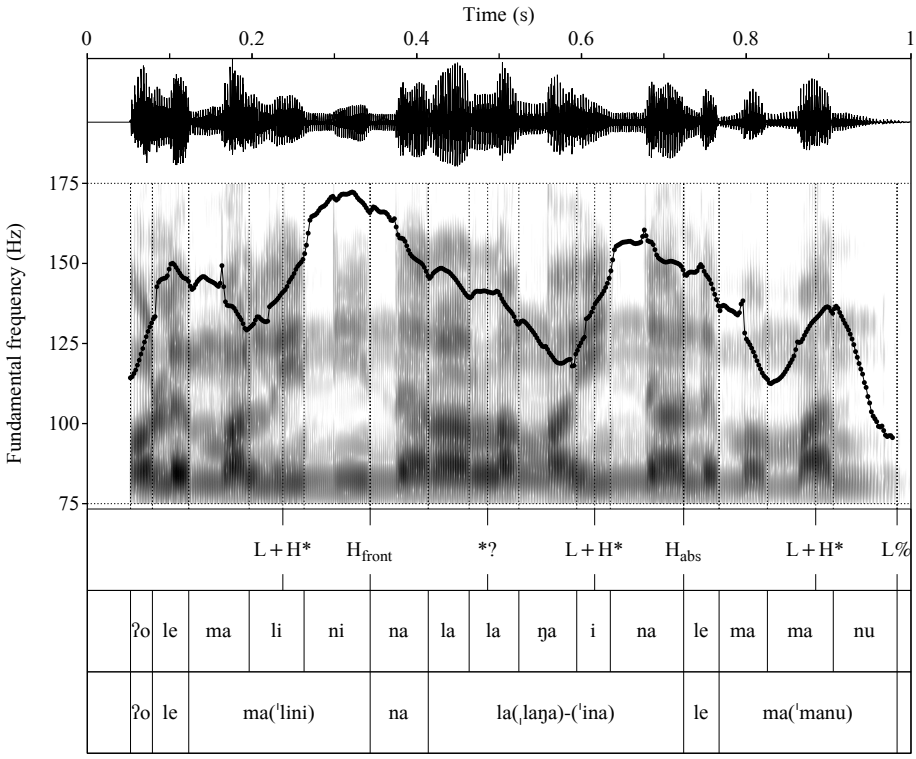


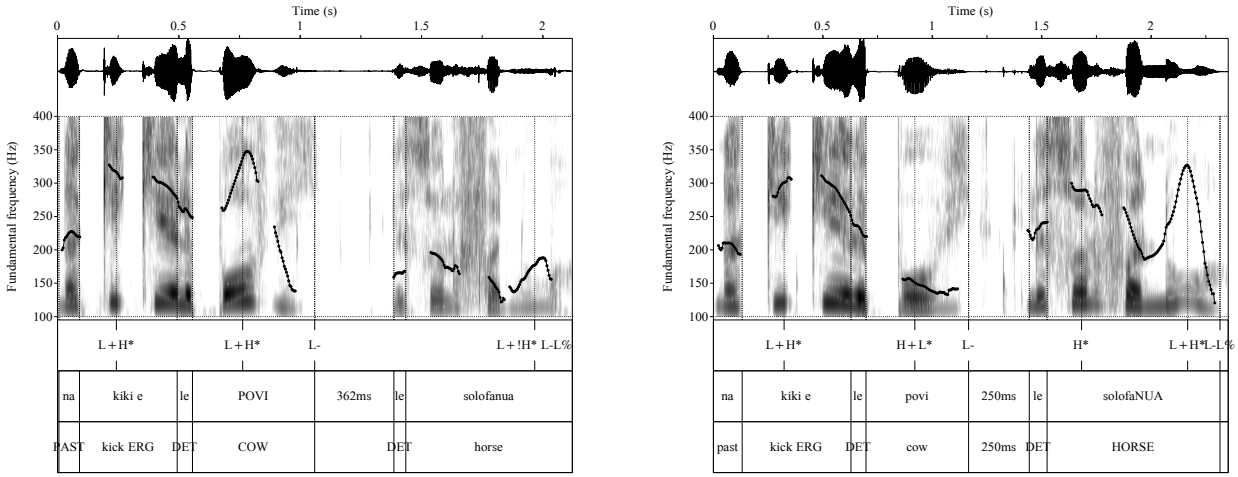
Figure 12.20: F0 contours for a fronted declarative (12) from la_m01

the length of arguments, as well as initial evidence that it appears in the *tautala leaga* register.[10] Yu (2021) found evidence that the appearance of the H edge tone before absolutives is insensitive to the syntactic nature (subject of intransitive, object of transitive predicates, proper names, pronouns, and arguments internal to nominalizations), certain semantic properties (specificity and number), and certain aspects of pragmatic context (word order, informational/contrastive focus) of the marked nominal. Moreover, the H edge tone did not appear where bare NPs are independently expected not to be case marked (pseudo noun incorporation (Massam, 2001)) or before *'o*-marked fronted nominals). Yu and Özyıldız (2016) also showed that an optional segmental absolutive particle ia_{abs} is licit in syntactic configurations where the absolutive H appears, and that ia_{abs} is illicit where absolutive H does not surface, as well as where the H in coordination and fronted arguments appears, and speculated that the absolutive H may bear a diachronic relation to this seemingly moribund particle.

If syntactic structure were the only conditioning factor regulating the distribution of sentence-medial H edge tones, then we might expect H edge tones to invariably appear in the syntactic configurations mentioned above, and not to appear otherwise. This is however, not what is observed. Calhoun (2017), Yu and Stabler (2017) found that an H edge tone could also sometimes appear before oblique and ergative arguments, where it would not be expected. Calhoun (2017), Yu and Stabler (2017) also found that a L edge tone could sometimes appear in the syntactic configurations where an H edge tone is expected, see also, for example, the absence of an H immediately preceding absolutive argument [le solofanua] 'the horse' in sample utterances of (13) from Calhoun et al. (2019, Figure 1), shown in Figure 12.21. The utterances in Figure 12.21 were produced by a university teacher of Samoan in New Zealand, asked to read (13) with "as clear and consistent an accenting difference between them as possible while still being natural for the language" (Calhoun et al., 2019, p. 9): with nuclear accent (focus) on the subject (Figure 12.21a) and on the object (Figure 12.21b). (The use of the term 'nuclear accent' in Calhoun et al. (2019) refers to a sense coming from information structural prominence, rather than IntP-final position, see §12.4 for further discussion).

(13) VSO sentence in Calhoun et al. (2019, (13), Figure 1), shown in Figure 12.21

[10] Samoan is well-known to have socially-conditioned speech registers with systematic lexical and phonological properties. One distinction between registers is between *tautaula lelei* 'good language'—which has been described as being used in literary contexts and and Westernized institutional contexts like in church and school, as well as with foreigners, and *tautaula leaga* 'bad language'—described as being used in traditional ceremonies and meetings, as well as between family members and between friends (Shore 1977, 1980; Duranti 1981, p. 165–168; Ochs 1988, p. 196; Duranti 1990, p. 4–5; Mosel and Hovdhaugen 1992, p. 7–11) (although the difference in practice is not so clear-cut (Mayer, 2001)). The register distinction is marked by differences in segmental phonology: /t/ and /k/ → /k/ and /n/ and /ŋ/ → /ŋ/ from *tautaula lelei* to *tautala leaga*, as well as dropping of the ergative case marker (Mosel and Hovdhaugen, 1992, p. 9).



(a) VSO, nuclear accent (focus) on subject (b) VSO, nuclear accent (focus) on object

Figure 12.21: F0 contours for nuclear accent (focus) placement by speaker on subject *povi* 'cow' versus object *solofanua* 'horse' in VSO word order redrawn from Calhoun et al. (2019, Figure 1), including Calhoun et al. (2019)'s original intonational transcriptions. An L edge tone rather than an H edge tone occurs before the absolutive argument [le solofanua] 'the horse' in both renditions.

na (ˈki)ki e le povi le solofanua
PAST kick ERG DET cow DET horse

'The cow kicked the horse'

Additionally, Calhoun (2017) found that H edge tones appeared as expected before absolutive nominals, but only 25% of the time before an absolutive argument under focus-sensitive *na'o* 'only' constructions. With all of the distributional facts together at hand, it is undisputed that the sources of sentence-medial edge tones must extend beyond syntactic structure. What has been disputed about sentence-medial edge tones is: (i) if the systematic co-occurrence of long pauses (on the order of hundreds of milliseconds) together with some of these edge tones suggests that those edge tones may be distinct from edge tones that systematically do not appear with such pauses, (ii) if sentence-medial prosodic boundary tones are sometimes intonational phrase tones, or necessarily phonological phrase boundary tones, (iii) if there is a unified source or multiple sources of high and low edge tones arising from different parts of the grammar, and (iv) what this source or these sources are. The next section, §12.3.2, introduces two alternative proposals with different ways of resolving the issues under dispute. The issues under dispute are also further discussed in §12.4 in the context of how biases from typical practice in AM theoretical analyses can affect the development of analyses.

12.3.2 Current analyses of sentence-medial edge tones

The two most current analyses of sentence-medial edge tones attribute their regulation either to the unified source of information structure (Calhoun, 2017, Calhoun et al., 2019), or to two distinct sources: morphosyntactic spellout and phonological phrasing (which itself could be influenced by mapping from syntactic to prosodic domains and/or information structure) (Yu and Stabler, 2017, Yu, 2021).

Calhoun (2017, p. 36-37) proposes that phonological phrases correspond to information units and that all sentence-medial high and low edge tones in Samoan are phonological phrase boundary tones (H-, L-) that are regulated by information structure in terms of the rheme and the theme. As reviewed in Calhoun (2017, p. 7), the rheme is "the part of the utterance which updates the common ground, or is new in relation to the current question-under-discussion", while the theme "establish[es] a link between the current utterance and the common ground or what the current question-under-discussion is." Calhoun (2017) examined copular constructions (equative sentences) in rheme-theme and theme-reme orders. For example, in the context of "Tell me who's in the family", "Amani is the youngest in the family" is in rheme-theme order, while "The youngest in the family is Amani" is in theme-rheme order (Calhoun, 2017, p. 16-17). Based on patterns of productions of sentences like these with varying rheme-theme orders, Calhoun (2017) made the proposal quoted in (14).

(14) Summary of Calhoun (2017)'s proposal quoted from (Calhoun, 2017, p. 37)

a. The default ordering of information in Samoan is rheme-theme. In this order, the rheme is normally phrased separately to the theme.

b. If the theme contains a focus, it should normally precede the rheme, a focused theme following the rheme is dispreferred. In theme-rheme order, a prosodic boundary between the constituents is optional.

c. H- phrase tones mark an information unit as incomplete. Typically, this marks the end of a rheme with a following theme. However, H- tones can also mark coordinated information units.

d. L- phrase tones mark a completed information unit.

e. A weak ((!)H*) or no accent on a constituent marks it as backgrounded.

The proposal states that variability in whether a sentence-medial tone occurs or not is due to variability in whether a prosodic boundary is present or not, the likelihood of which is dependent to some extent on, but not determined by theme/rheme ordering. If a phonological phrase boundary is present, the choice of an H- vs. an L- is due to differences in whether an information unit is incomplete or complete, following Pierrehumbert and Hirschberg (1990, p. 302). Speakers may interpret

the same discourse context differently in terms of themes and rhemes, resulting in variability in the presence of a phrase tone, as well as whether the choice is H- or L-. For example, speakers could interpret the response "Melina only ate the melon earlier" in VSO order in the discourse context "Why was Melina hungry?": (i) as a single information unit and produce no boundary tone after Melina (the subject) or (ii) as two information units (two rhemes)—one offering that Melina had eaten in the past, and another offering that what she was eaten was only a melon—with a L phrase tone between the subject and object (Calhoun, 2017, p. 24). Similarly, speakers could choose to produce sentences in the same discourse context with different rheme/theme structures, resulting in variability in phrase tone choice e.g., two rhemes (separated by L-) or rheme-theme (separated by H-). The allowance for this extent of variability in speakers' choices for parceling out utterances, even under the same discourse context, into information units and themes and rhemes—as well as variability in how those choices may condition the presence and type of edge tones—makes it difficult to falsify Calhoun (2017)'s proposal (Yu, 2021, §6.4). While Calhoun (2017)'s proposal can account for the current empirical data, this difficulty of falsification makes it especially valuable to see if there are alternative proposals that may be more easily ruled out.

With the exception of two of the focus-sensitive *na'o* 'only' constructions, all syntactic constructions examined in Calhoun (2017) were non-verb-initial. Sentence-medial edge tones were observed after fronted nominals, between coordinated clauses, and before sentence-final adverbials. Before absolutive objects under *na'o*, sometimes no phrase tone was observed at all, and other times an L- or H- was observed—even though an H edge tone would be expected to invariably occur before the absolutive argument. The sentence-medial edge tones observed in Calhoun (2017) typically co-occurred with silent pauses on the order of hundreds of milliseconds, as did the sentence-medial edge tones in Calhoun et al. (2019), as shown in figures of sample F0 contours. Yu (2021, §5.3) argues that the sentence-medial edge tones in Calhoun (2017)—which typically co-occur with long silent pauses—are inserted by phonological grammar as IntP boundary tones, while the H edge tones observed reliably before absolutive arguments, in coordination, and after fronted arguments—which appear without lengthening and/or pauses—are inserted by morphosyntax: "Hs inserted in spellout *reliably* appear in a small, restricted set of distinct syntactic configurations (absolutive case, coordination, fronting; more may be found with additional fieldwork). Moreover, the distribution of H_{abs} patterns with the distribution of segmental case markers rather than with the distribution of H_{coord} and H_{front}. Hs (and Ls) inserted by the phonological grammar, denoted as H% and L%, don't reliably appear in particular syntactic configurations. Instead, they are triggered by prosodic domains, which are sensitive to both syntactic structure and prosodic factors. Thus these Hs *variably and sporadically* appear in variable syntactic configurations. These syntactic configurations do include those that trigger the spellout of an H, but also others, e.g., ergative case and oblique modifiers and objects" (Yu, 2021, p. 336). The plausibility of having

tones mark specific syntactic constructions, including marking case, is discussed at length in Yu (2021), but will not be discussed here as the focus of this paper is the intonational phonology.

In Yu (2021)'s proposal, what governs the variation of the edge tones inserted by prosodic grammar (H%, L%) could include information structure, in the ways proposed by Calhoun (2017) in (14), although mediated via syntax (Yu, 2021, §6.4). It is relatively easy to falsify the strong hypothesis that there are distinct classes of edge tones: (i) H edge tones that are invariably inserted in certain syntactic configurations in spellout, and (ii) L% and H% tones inserted in the phonological grammar at the edge of a prosodic constituent. If one of the syntactic configurations expected to have an H tone reflex is present, yet no edge tone at all appears, then the only recourse is to provide credible evidence for an alternative syntactic analysis. This is the strategy taken in explaining productions in which no H_{abs}appears before absolutive arguments under *na'o* Yu (2021, §5.1.2). Although Mosel and Hovdhaugen (1992, p. 526) refers to "the absolutive noun phrase of *na'o*", three speakers across a wide range of ages found it acceptable to have ergative and oblique nominals under *na'o* as well, but only if the corresponding segmental case markers were absent—suggesting that case-marking, whether segmental or tonal, is illicit under na'o for those speakers. If an H edge tone appears in a position where no morphosyntactic H is expected, then it must be attributed to an H% and should exhibit variability in its presence, and an L edge tone (L%) can also be expected to variably occur in the same environment. Accordingly, Yu (2021, §5.1.2) analyzed the H edge tone co-occurring with a pause/lengthening that appeared only 25% of the time before an absolutive argument under focus-sensitive *na'o* 'only' constructions (where case-marking is illicit) as an H%. Similarly, L edge tones co-occurring with pauses that sometimes appear where a morphosyntactically-conditioned H is expected—like before the absolutive argument in Figure 12.21 or at the right edge of the fronted argument in Figure 12.17a—are analyzed as L% tones. The hypothesis is then that IntP tones like an L% can "override" H edge tones inserted in spellout.

Yu (2021, p. 336)'s proposal is stated as strongly as possible to make it easily falsifiable. Thus H% and L%s are hypothesized to reliably co-occur with a pause, distinct from Hs inserted in spellout, which are hypothesized to not co-occur with a pause. As expected, the proposal may need to be weakened to account for IntP tones (and possibly prosodic boundary tones at lower levels in the prosodic hierarchy) which do not co-occur with pauses, but potentially other (potentially weaker) juncture cues. For example, we the edge tones at the right edge of the fronted phrases in Figure 12.17a,b as IntP tones, but they do not co-occur with pauses. To our ears, there is still a strong percept of a juncture where the IntP tones are marked due to a jump in F0 (fn. 8), but of course, systematic perceptual work would need to be done with native speakers. Moreover, Yu (2021, §5.3)'s analysis reflects a general bias in intonational work to rely on a pause as necessary and sufficient evidence for a juncture level coinciding with an Intonational Phrase.[11] Other factors

[11]As a rule of thumb, pauses have long been used to diagnose Intonational Phrase boundaries, see e.g., Selkirk

contributing to a percept of juncture may be as or even more important. For example, one reason that Calhoun et al. (2019, p. 13) analyzes sentence-medial tones like those in Figure 12.21 as L-, H- tones rather than IntP tones is because they observe "coherent intonation contours" across the sentence, in spite of the pause. Calhoun (2015, p. 220) also notes two consistent properties of the scaling of pitch accents immediately following H edge tones: (i) the dip of the local F0 minimum expected from the L target can be reduced, and (ii) the H of the pitch accent is lower than the F0 reached by the high edge tone, e.g., see the downstep marked in transcriptions on the tonal event immediately following the sentence-medial high tone in Figures 12.4. Relatedly, a third property is that the low F0 level at the beginning of F0 rise of the pitch accent can sometimes be reached "late", after the beginning of the stressed syllable, as shown in Figure 12.22, to be discussed in the next section, §12.3.3.

12.3.3 On-going analyses of sentence-medial edge tones

Clearly more production and perception work on F0 scaling across prosodic domains delineated by edge tones with and without pauses is needed to provide evidence bearing on claims of prosodic constituents below the IntP in the prosodic hierarchy as well as claims of a class of phonologically-conditioned edge tones at prosodic junctures distinct from Hs inserted in spellout. There are two issues from intonational phonology we'd like to draw attention to in informing ongoing work on F0 scaling and edge tones in Samoan. First, effects of tonal crowding on phonetic implementation play an important role in conditioning the scaling of pitch accents, including pitch accents that neighbor H edge tones. Second, we should revisit questions raised by Calhoun (2015, p. 216) about the phonological analysis of the tonal events we've been calling H edge tones. Yu (2021, §4.6) showed some evidence that these tonal events are edge tones because they are still pronounced with high F0 at the right edge of words with initial stress like *Melanie* (in code-shifting). But how the H edge tone interacts with the immediately preceding pitch accent remains an open question.

Information structural factors may condition the scaling of pitch accents, as proposed in Point (e) in Calhoun (2017, p. 37)'s proposal quoted in (14). However, those factors must interact with the effects of tonal crowding. In particular, the reduction of the L target and reduced height of the H peak in the pitch accent following an H edge tone noticed in Calhoun (2015, p. 220) are gradiently scaled by how much tonal crowding there is between the high edge tone and the immediately following pitch accent. This scaling of the low and high F0 of the pitch accent and delay in reaching

(1978/1981, p. 135), Pierrehumbert (1980, p. 19), Ladd (1986, p. 315-317), Nespor and Vogel (1986, p. 188), Krivokapić (2007, p. 163), Jun and Fletcher (2014, p. 501-502). For instance: "It is *between* intonational phrases (and only between them, we would claim) that one finds potential pauses." (Selkirk, 1978/1981, p. 135) But it is possible, since much of this early work relied on auditory transcription rather than instrumental measurements, that what was identified as a pause could have been a confluence of other junctural cues, including pitch reset.

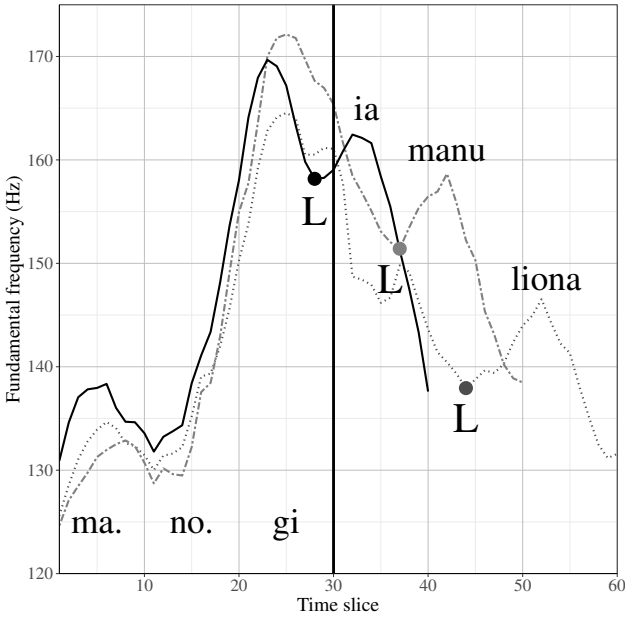


Figure 12.22: Comparison of individual F0 contours from la_m01 (same data set as in §4.2 in Yu (2021)) illustrating the reduction of L target F0 valley and increase in H target F0 maximum on the subject as tonal crowding increases with preceding H edge tone at right edge of the verb ma(ˈnoŋi).

the L target by tonal crowding with respect to the preceding H edge tone is summarized in the comparison of individual F0 contours in Figure 12.22 of the verb [ma(ˈnoŋi)] followed by a series of subjects with stress progressively closer to the stressed syllable of the verb: [li(ˈona)] ‘lion’, [(ˈmanu)] ‘animal’, [(ˈia)] ‘3.SG’. Figure 12.22 comes from la_m01, from the same intransitives data set as in §4.2 in Yu (2021), using a sentence template very similar to (3).

While Figure 12.22 draws attention to the interaction between H edge tones and pitch accents immediately following them, there are also unresolved questions about the pitch accents immediately preceding H edge tones. For example, Calhoun (2015, p. 216) points out that an alternative analysis of the tonal event sequence of an L+H* followed by an H edge tone could be a single L+H- phrase accent. Figure 12.23, shows the excerpt of The North Wind and the Sun fable in (15) read by three different speakers, with quite different choices for the placement of pitch accents (as evidenced by F0 movements). To reduce the density of tonal events in a single tier for readability, pitch accent and edge tones are split up into separate tiers in the figures. In addition to providing information about which kinds of edge tones are invariable, the renditions raise questions about the relation between H edge tones and the pitch accents that immediately precede them.

(15) Excerpt from North Wind narrative

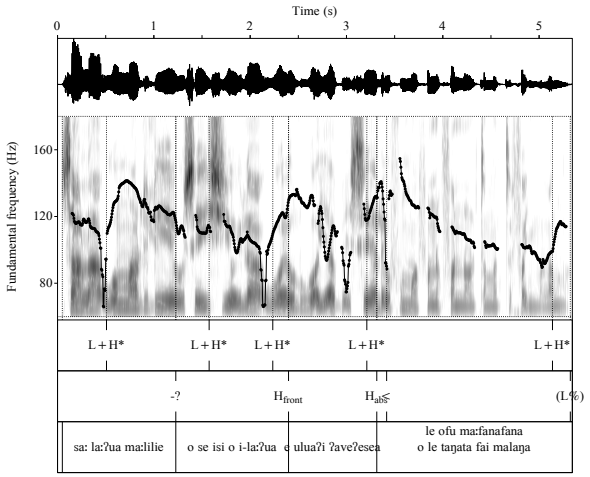
sa: laːʔua malilie ʔo se isi o i-laːʔua e ulua ʔi ʔaveʔese-a ∅ le
PAST 3.DU agree.RED TOP DET other GEN 3.DU PRES first remove-ERG ABS DET

ofu-mafanafana o le tagata fai malaga
cloak GEN DET man do journey

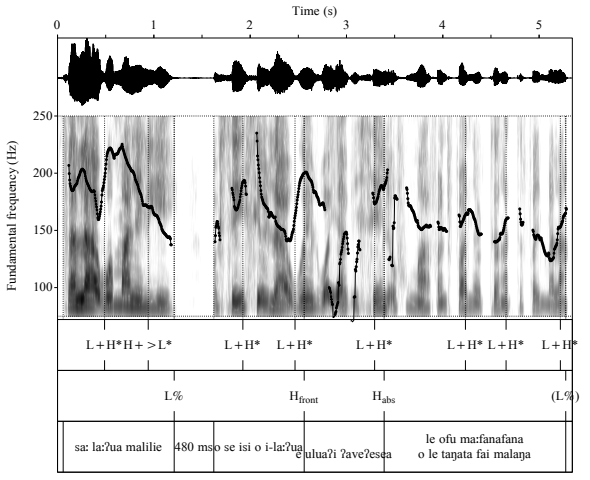
'They agreed that the one who first succeeded in making the traveler take his cloak off'

The invariability and variability in the distribution of sentence-medial edge tones among the speakers has a generalization: all speakers exhibit an H edge tone right before the absolutive argument at the right edge of [ʔaveʔese-a], marked as H_{abs}. Two of the speakers exhibit an H edge tone at the right edge of [i-laːʔua] inside the relative clause before the predicate, marked as H_{front}, and all three speakers exhibit a prosodic juncture marked by lengthening and/or a pause between the root and embedded clauses, marked as H_{RC} for la_m01, although the tonal characteristics and size of the juncture varies across speakers. The variability in the juncture at the onset of the relative clause suggests that edge tones at this site might come in the prosodic grammar, potentially arising from mapping of syntactic domains. The absence of the H_{front} inside the relative clause for la_m01 could be because the syntactic environment inside a relative clause is different from the configuration resulting in H_{front} in matrix clauses; more work is needed to investigate this and the other preliminary observations in this section.

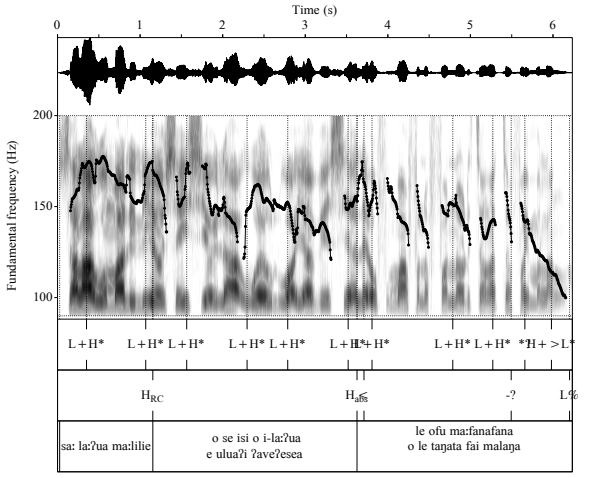
The distribution of pitch accents in the multiple renditions of the North Wind excerpts is also of interest. First, the lack of clear pitch accents on every content word suggests that pitch accents are not in fact obligatory on each content word, contra Orfitelli and Yu (2009). However, it is apparent that the speaker who provided the utterances analyzed for Orfitelli and Yu (2009), la_m01 in Figure 12.23c, is a speaker who has a high density of pitch accents relative to other speakers. Second, different choices between speakers for where pitch accents are placed suggest different choices of information structure between the speakers for the same text, and/or other factors of individual variability. Speaker smo_s4 has noticeably fewer pitch accents than the other two speakers, while speaker la_m01 has the most. A common property of the distribution of pitch accents across speakers, though, is that a sentence-medial word bearing a high edge tone at the right edge (but no pause or obvious lengthening) always has a downtrend-breaking rising pitch accent. Thus, whatever might be conditioning where high edge tones appear (see §12.3), it seems that a rising pitch accent obligatorily occurs before it: the H edge tone "protects" the pitch accent from deletion or suppression. This protection could be a consequence of the H edge tone and the immediately preceding pitch accent being a single tonal event, e.g., a L+H- phrase accent, or of the "protected" L+H* accent immediately preceding H edge tone serving a demaracative function, like in Aziz (2020*a*,*b*)'s analysis of Malagasy. Either analysis showcases one way Samoan intonation challenges AM theory in practice, which supposes a sharp division between pitch accents and edge tones and their form and function. We discuss this and other biases of AM theory in practice in the context of Samoan intonation in the next section.



(a) Speaker smo_s4



(b) Speaker la_s22



(c) Speaker la_m01

Figure 12.23: F0 contours for North Wind excerpt in (15) showing invariability and variability in pitch accent distribution.

12.4 Challenges for the practice of AM theory

Both properties of Samoan intonation itself, as well as progress towards understanding it, illuminate challenges for the practice of AM theory. By specifying the *practice* of AM theory, we draw a distinction between the concepts of AM theory and the various ways in which AM theory has been understood in informing intonational fieldwork and analysis in the past few decades. Ladd (2000, p. 37) states: "The central claim of the AM view is that an intonation contour is represented phonologically as a sequence of tones associated in well-defined ways with the segmental string." Samoan intonation poses no challenge for this bare statement of an autosegmental approach (Goldsmith, 1976) to intonation. But perhaps the first place that many people would look for a definition of AM theory is Ladd (1996, 2008, Ch. 2), which coined the term 'autosegmental-metrical' theory and lists "four basic tenets" of the AM approach. The first basic tenet elaborates on the central claim of sequential tone structure mentioned above from Ladd (2000, p. 37): "tonal structure consists of a string of local events associated with certain points in the segmental string. . . .In languages like English, the most important events of the tonal string are pitch accents, which are associated with prominent syllables in the segmental string, and edge tones [i.e., prosodic boundary tones], which are associated with the edges of intonational tunes at major prosodic boundaries" (Ladd, 2008, p. 44). Note that Ladd (2008, p. 48-49) defines a pitch accent here as "a local feature of a pitch contour. . .which signals that the syllable with which it is associated is prominent in the utterance". He further clarifies, regarding prominence, that "pitch accents must be associated with metrically strong syllables", i.e., he means metrical prominence, and that "in some languages like French or Indonesian pitch accents may associate to syllables which are not necessarily stressed and which may not seem 'prominent' either to native speakers or phonetically trained listeners" (Ladd, 2008, p. 62).

The problems Samoan intonation and its analysis pose for the practice of AM theory arise when we neglect the proviso "in languages like English" prefacing the remark in Ladd (2008, p. 44)'s first tenet, overgeneralize the connection between stress and pitch accent in English to other intonational systems, and consider *only* the kinds of tonal events named, i.e., pitch accents and prosodic boundary tones. First, the apparent analytic difficulties presented by IntP-final pitch accents insensitive to focus conditions raised by Calhoun (2017) and Howard (2018) is a consequence of how AM theory in practice overgeneralizes the link between pitch accents, metrical strength, and information structure in English to all languages (§12.4.1). Second, the hypothesis that some Samoan sentence-medial edge tones arise from spellout and are not prosodic boundary tones challenges the assumption—in practice—that tones in the string arise exclusively from the two named types of events—either as pitch accents or prosodic boundary tones (§12.4.2).

Default assumptions of a relation between various kinds of prominence and pitch accents and

concentrated attention to phonological sources of tones are exemplified in Himmelmann and Ladd (2008)'s first sentence in a description of a "minimal framework for investigating intonation in a new language" based on AM theory: "The most important phonological distinction to be drawn is the one between intonational features at major prominent syllables and intonational features at boundaries: in current terminology, the distinction is between 'pitch accents' and 'boundary tones"' (Himmelmann and Ladd, 2008, p. 253-254). The same biases of AM theory in practice are evident in Arvaniti and Fletcher (2020, p. 82)'s overview of AM theory: "Independently of the particular version of prosodic structure adopted in an AM analysis, it is widely agreed that tones associate with phrasal boundaries or constituent heads (informally, stresses) or both. Tones that associate with stressed syllables are called 'pitch accents' and one of their roles is prominence enhancement. . ." And in recommendations for approaching intonational fieldwork from an AM perspective, Jun and Fletcher (2014, p. 514) describe two possible classifications for tones: as ones that "can mark either the head (pitch accent) or the edge (phrasal or boundary tone) of a prosodic unit."

12.4.1 Positionally-defined (nuclear) pitch accents in AM theory

The automatic categorization of tones that are not prosodic boundary tones as pitch accents—which in turn are narrowly defined as tones necessarily associated to syllables with word-level stress—comes in only in AM theory in practice, even as early as Hayes and Lahiri (1991, p. 51)'s analysis of Kolkata Bengali intonation, which states: "Pitch accents are tones that get linked to stressed syllables". A broader definition of tones associated with accent is clearly implied in AM theory,[12] considering that seminal works for AM theory include Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988)'s analysis of Japanese, which treat pitch accent "as a lexically linked H tone" (Beckman and Pierrehumbert, 1986, p. 256), see also Pierrehumbert and Beckman (1988, p. 122). Unlike English, then, Japanese pitch accent location is assigned in the lexicon rather than conditioned on word-level stress assigned by the phonological grammar. Furthermore, Beckman (1986) explicates that Japanese accent is not a "stress accent" because its identifying acoustic features are primarily f0-based, rather than a conglomeration of not only f0 but also other features such as intensity.

What could a broader definition of pitch accent that encompasses Japanese be? Gussenhoven (2004, §3.5), Gussenhoven (2011, §3.3) and Gussenhoven (2022, §1.2) lay out one way to do it and

[12] See also Ladd (2008, p. 61) statement that "a syllable can be metrically strong or prominent without necessarily being stressed", and further that "The AM approaches to pitch accent sketched here draw no fundamental distinction between tones that are assigned in the lexicon and tones that are assigned or modified postlexically" (Ladd, 2008, p. 166).

define accent as "a place marker in the phonological structure where tones are to be inserted" or "an abstract phonological location indicator of tone", following Goldsmith (1976, 1981, 1984); a 'pitch accent' under this view is simply "the tone melody that must or can be inserted in the accented syllable" (Gussenhoven, 2022, §1.2). This purely locational definition of 'accent' and 'pitch accent' removes the puzzle of how a pitch accent can be dissociated from notions of prominence, like how a nuclear (IntP-final) accent can be invariant under different focus conditions in Samoan (§§12.2.1, 12.2.4). The location and presence of accents can be determined by the lexicon, phonology, morphology, syntax, and/or information structure, as exemplified in Gussenhoven (2022). In English, word-level stress is assigned in lexical phonology, pitch accents can only appear on stressed syllables, and one factor among additional determinants of accent presence is information structure, via pervasive post-focal de-accenting. In Samoan, word-level stress is also assigned by the phonological grammar, and accents can only appear on stressed syllables (§12.1.1). However, Samoan does not have pervasive post-focal de-accenting like English, as evidenced by the invariable IntP-final pitch accent in declaratives and interrogatives across focus conditions (§12.2).

When all words in an IntP following a focused element are de-accented—a typical situation in a language like English—the IntP-final pitch accent necessarily appears on the focused element. Due to West Germanic languages like English with this pervasive post-focal de-accenting, 'nuclear accent' has taken on an additional definition besides IntP-final in AM theory in practice: the focal accent (Grice, 2022, p. 69), and more broadly, the most salient or prominent accent, in some acoustic, perceptual, and/or information structural sense (Ladd, 2008, § 4.1.1, Ch. 7). For example, Féry (2017, p. 2, 61) states that "the primarily accented word of a sentence" bears the nuclear accent (or nuclear stress), which can be "changed if needed by context", e.g., when different elements are under contrastive focus. Pierrehumbert (1980, p. 40) describes locating pitch accents in English based on "prominence", in the sense of "the aggregate of metrical strength and emphasis, as it pertains to the control of tonal values". From this metrical perspective that integrates word-level stress and information structural salience together into a single calculation of prominence (Pierrehumbert, 1980, p. 37-38), the identification of a tonal event as a pitch accent requires not only the temporal location of a stressed TBU, but also a percept of some kind of prominence (Ladd, 2008, §7.1.3). For example, ToBI transcription guidelines for Mainstream American English (MAE-ToBI) describe pitch accents as "events or patterns that mark syllables as more salient or more prominent than neighboring syllables by reason of their intonation" (Veilleux et al., 2006, Ch. 2). While MAE-ToBI is only one instantiation of intonational transcription conventions for one variety of a language and is also distinct from Pierrehumbert (1980), Beckman and Pierrehumbert (1986)'s AM-theoretic analyses of English (Pierrehumbert, 2000, Arvaniti, 2022, Jun, 2022), MAE-ToBI transcription guidelines have nevertheless had an outsized influence on AM theory in practice (Ladd, 2022).

Calhoun (2015) et seq. follow the English-centric prominence-based notion of (nuclear) pitch accent of AM theory in practice, whereby: (i) pitch accent presence is determined by a percept of prominence, e.g., "Accents were sometimes produced without a clear initial L target (e.g. maea in Fig. 2 [adapted as Figure 12.4 in this paper]); these were labelled H* (or !H*). In these cases, more weight was placed on other cues to prominence, such as duration and amplitude, in deciding accent status" (Calhoun, 2015, p. 216), and (ii) metrical prominence is directly related to information structural salience and focus (Calhoun, 2015, p. 219), see also Calhoun (2010). Therefore, because of the invariable occurrence of the IntP-final pitch accent in declaratives and interrogatives across focus conditions and the downtrend-breaking high scaling of IntP-final pitch accents in declaratives, Calhoun (2015, p. 219) et seq. propose that the IntP-final pitch accent is in fact a non-prominence-lending phrase accent that is solely "edge-marking": L+H- in declaratives and H+L- in interrogatives.[13] But if we retreat from an English-centric definition of (nuclear) pitch accent and step back to a purely locational/positional definition of (nuclear) pitch accent in AM theory, there is no conflict between a tonal event being a pitch accent and being dissociated from contrastive/informational focus. The tendency in AM theory in practice to define pitch accents and prosodic boundary tones simultaneously by how they are located (via accents or prosodic constituent edges) and(/or) whether they are prominence-lending or not "can be a straitjacket for the analysis of new languages" (Grice, 2022, p. 65). If we eliminate the need for proposing phrase accents like L+H- and H+L- in Samoan, then there is no need to posit duplicate sets of tones in the tonal inventory of Samoan like in Calhoun (2015, p. 219), based on whether they are nuclear or pre-nuclear (in the positional sense).

By revealing Samoan's de-coupling of focal accent from IntP-final position, Calhoun (2015, 2017), Calhoun et al. (2019) show that the assumption of AM theory in practice that the focal accent is necessarily IntP-final can be misleading and confining. Attention to IP-final tones has been useful for understanding typological variation and the role of discourse context in intonation, e.g., in Romance languages (Frota et al., 2007). Partitioning accents into two classes—IntP-final and non-final—has also been useful for identifying generalizations in tonotactics, e.g., like in Samoan, as discussed in this paper, as well as many other languages, e.g., see Hayes and Lahiri (1991), Frota and Prieto (2015*a*), Post (2000), Dainora (2006). But the assumption of a special status of the IntP-final accent arising from pervasive postfocal deaccenting and associated with various kinds of prominence is called into question by Samoan, Egyptian Arabic (Hellmuth, 2009, 2006), Swedish varieties (Bruce, 1977, Bruce and Gårding, 1978), and Italian varieties (D'Imperio, 2001, Swerts et al., 2002), among others.

[13] More generally, Calhoun (2015, p. 219) et seq. , propose that L+H-, H+L- and other phrase accents occur not only at the end of intonational phrases (immediately prior to an L%), but at the end of phonological phrases, see §12.4.2 for further discussion.

Samoan and other verb-initial languages offer a rich source for building our understanding beyond a West Germanic-centric view of the relation between nuclear accent and prominence. Calhoun (2015) et seq., proposes that the default position of the nuclear accent (in the prominence sense) in canonical VSO order is not IntP-final, but IntP-initial on the verb, and Calhoun et al. (2019, p. 6) suggests that this initial position is common for verb-initial languages (Kügler and Calhoun, 2020), which have received less attention in intonational fieldwork. The rationale is that this initial position of phonetic prominence, since the first accent in an IntP has the largest range of all pitch accents in the IntP, and also because sentence-initial position is a position where new or focal information is placed, e.g., clefted SVO/OVS constructions.

Calhoun et al. (2019) also reveals a striking role of syntactic structure in modulating perceptual and information structural prominence, as well as variability systematically conditioned on language dominance of bilingual speakers. Calhoun et al. (2019) investigated: (i) perceptual prominence, by asking listeners to identify the most prominent/heavy word (*fa'amamafa*) in the the produced sentences described above, and (ii) information structural prominence, by asking listeners to judge whether the most likely question preceding the sentence heard was one consistent with wh-focus on the object or wh-focus on the subject. Listeners, and especially the most Samoan-dominant speakers (all speakers had some degree of fluency with English) chose the word with the nuclear accent (as defined in the study, see discussion of Figure 12.21 in §12.3.1) as the most prominent word from around 60 to 80% of the time on average, depending on the word order condition. The likelihood of choosing the nuclear accented word was highest when the nuclear accented word was also fronted (in SVO, OVS orders) or the subject was nuclear accented in default VSO order, and least likely when the nuclear accented word was in SVO and OVS orders, but not the fronted argument, with likelihood of choosing the nuclear accented word when the object was nuclear accented in VSO order falling in between. Thus, while perceptual prominence tracks with acoustic prominence, it is also modulated by expected informational structural prominence due to syntactic structure. These syntactic effects played an even bigger—even, the only—systematic role in listeners' assessment of information structural prominence, which were very variable: listeners chose the object as the most informationally structural prominent around 45-70% of the time on average in VSO and OVS sentences regardless of nuclear accent placement, with no significant difference in likelihood across conditions, and listeners chose the subject as most information structurally prominent about 70% of the time on average in SVO sentences, regardless of nuclear accent placement.

12.4.2 Tones that are neither pitch accents nor prosodic boundary tones

The bias to assume that pitch accents are assigned to stressed syllables, discussed in the previous section, is part of a general bias of AM theory in practice towards assuming that determinants of pitch accent placement are necessarily phonological. Gussenhoven (2011, 2022) point out that this phonological bias has come at the expense of recognizing other determinants of pitch accents such as lexical and morphosyntactic factors. In this section, we extend that point to edge tones. As we mentioned at the beginning of §12.4, AM theory in practice slips into assuming that tones in the string that are not pitch accents must be prosodic boundary tones. The development of accounts of the distribution of sentence-medial edge tones in Samoan presents an illustrative case study in the value of widening the possibilities considered beyond the AM-theoretical tendency to conflate edge tones and prosodic boundary tones. Among other things, this bias can lead to proposals of prosodic domains based only on the distribution of edge tones, without additional evidence from other correlates of prosodic junctures.

For example, Orfitelli and Yu (2009, p. 8) proposed that Samoan has an intermediate phrase solely because of the presence of sentence-medial high edge tones, e.g., in coordination, and thus analyzed the high edge tone as an H- phrase accent associated to the intermediate phrase node. Orfitelli and Yu (2009, p. 7-8) also assumed that Samoan had an L- phrase accent that was yet to be uncovered. In this first intonational analysis of Samoan (Orfitelli and Yu, 2009), we were certainly influenced by having our first exposure to intonational analysis and AM theory via MAE-ToBI, thinking—a tone that regularly appears sentence-medially that isn't a pitch accent... since it isn't sentence-final,[14] it can't be an intonational phrase tone, so it must be an intermediate phrase tone, and so, aha, Samoan must have an intermediate phrase!

Gussenhoven (2004, §7.3.4) and (Gussenhoven, 2016, p. 1) have pointed out that the proposal of an intermediate phrase level in American English itself in Beckman and Pierrehumbert (1986) could be viewed as a consequence of this same kind of reasoning. Pierrehumbert (1980)'s En-

[14]The analysis against sentence-medial edge tones being analyzed as intonational phrase tones is reflective of a bias in intonational phonology to consider sentences as the unit that maps to intonational phrases. Calhoun et al. (2019, p. 13) states that the fact that the edge tones "occur regularly in mid-sentence position" supports their analysis as boundary tones associated with phonological phrases, below the Intonational Phrase. Part of this bias may be because most ToBI training material examples, as well as elicited materials for developing an intonational phonological analysis, tend to be single sentences and single intonational phrases. In general, there is less attention in intonational phonology to syntactic structures at the level of clauses and across multiple clauses, and to phonological structures at the level of the Intonational Phrase or Utterance and across multiple Intonational Phrases. Sentences are often conflated with clauses and Intonational Phrases with Utterances. The syntax-prosody interface constraint MATCHCLAUSE (Selkirk, 2009, p. 40) states that syntactic clauses are mapped to intonational phrases, but little work examines clauses and sentences and how different analyses of clauses might matter for the syntax-prosody interface (although see, e.g., Kawahara and Shinya (2008), Selkirk (2009) for exceptions).

glish phrase accent, sandwiched between pitch accents and the IntP-final boundary tone, could be seen as inspired by Bruce (1977)'s Swedish 'sentence accent' (Pierrehumbert and Beckman, 1988, p. 246), which can also be sandwiched between a (lexical) pitch accent and a 'terminal juncture' tone at the end of the sentence. Beckman and Pierrehumbert (1986) further re-analyzed Pierrehumbert (1980)'s English phrase accent tone as a a prosodic boundary tone at the right edge of a newly proposed prosodic constituent, the intermediate phrase. But Bruce (1977, p. 20,24) in fact stated only that "sentence accent is placed on the last element of the rheme", where the rheme contains "the greatest amount of new information in the sentence", i.e., Bruce (1977) positioned the 'sentence accent' in relation to a information structural unit, not necessarily a prosodic unit. It is far from clear that the Swedish focal 'sentence accent' H tone at the right edge of a focal constituent/unit is necessarily a higher-level prosodic boundary tone like an ip tone. It has received a variety of phonological analyses (Gussenhoven, 2004, §7.3.4), e.g., as a floating tone temporally aligned (but not phonologically associated) to a stress group (foot) boundary (Bruce, 1983, 1987). Gussenhoven (2004, p. 212) analyzes it as a H tone related to a "focal constituent" (as does D'Imperio (2001, p. 349)) and states that: "by not aligning it with a prosodic constituent like ω or φ, we account for the fact that it does not behave like a typical boundary tone."

In the case of Samoan, the jump to an assumption that sentence-medial edge tones are necessarily prosodic boundary tones led to further puzzles when Yu (2011), Calhoun (2015) proposed further that all high sentence-medial edge tones (they also didn't find any low sentence-medial edge tones at the time) marked prosodic domains that arise from the mapping of syntactic constituents to prosodic constituents, e.g., syntactic XPs to phonological phrases. The assumption at the time that the high edge tones arising in coordination, absolutive case, and fronting all have a unified source via a general mapping of syntactic XPs to prosodic constituents reflects an even more general attentional bias to fixate on prosodic constituents in work on syntax-prosody interface, to the exclusion of other aspects of the interface such as morphosyntactic spellout (Yu, 2021). Later, Calhoun (2017), Yu and Stabler (2017), Yu (2021) rejected the XP-φ mapping analysis because of the asymmetry in the distribution of H edge tones in VSO and VOS structures—both VOS orders from scrambling (Calhoun, 2017, § 2.2) as well as from pseudo-noun incorporation (Yu, 2021, §6.3.3). H edge tones appear before the object (and not after the verb) in VSO structures, but after the verb (and not before the object) in VOS structures.[15]

The current proposals—Yu and Stabler (2017), Yu (2021)'s proposal that a subset of sentence-medial H edge tones are inserted in morphosyntactic spellout and Calhoun (2017), Calhoun et al. (2019)'s proposal that that all sentence-medial edge tones are regulated by information structure—

[15]Note that this distribution of H edge tones in Samoan also does not fit with an analysis whereby the predicate and subject are analyzed as being mapped to separate prosodic domains, as has been proposed in Niuean (Clemens, 2014, 2019) and Malagasy (Aziz, 2020*a*,*b*).

may be divergent, but they are unified in drawing attention to determinants of edge tones beyond prosodic constituency. While Calhoun (2017), Calhoun et al. (2019) still links information structural units directly to prosodic constituents (phonological phrases), one could adopt their proposal with an eye towards the case of the Swedish focal H and remain open about how the edge tones become temporally sequenced, e.g., one option might follow Gussenhoven (2004, p. 212) ALIGN constraint with respect to an information structural unit rather than a prosodic constituent like a φ-phrase. Regardless of how the current proposals about the regulation of sentence-medial edge tones in Samoan fare in the face of future empirical data, we hope that the history of how their analyses have developed serve as a reminder that part of Bruce (1977)'s fundamental autosegmental insight that the mapping of tones onto the F0 contour arises via the interplay of a linear sequence of tones is that those tones can come from multiple sources in the grammar (Pierrehumbert, 2000, p. 19-20). In addition to tones assigned in the lexicon (lexical accent) and those inserted in the phonological grammar (intonational phrase tones), Bruce (1977)'s analysis also included the focal H tone, whose distinct grammatical source seems to have been perhaps misleadingly lumped together with prosodic boundary tones at times in AM theory lore.

12.5 Conclusion

Previous work on intonation in Austronesian languages, including Samoan, has mostly focused on prosodic interfaces with syntax and semantics/pragmatics. Here, we have put a bigger emphasis on intonational phonology proper and also highlighted aspects of Samoan intonation that present challenges for autosegmental-metrical theory in practice. Building on previous work, we described four common nuclear contours in Samoan. We explicated a new analysis of early falling nuclear contours for information-seeking interrogatives that showcased the importance of considering subsyllabic/moraic structure in phonetic alignment of tonal events. Namely, it is important to consider the timing of the F0 fall with respect to the onset and rime. A fall over only the onset of the stressed syllable/mora is distinct from a fall that extends over the rime. A fall over only the onset into a low plateau in the rime can result in the percept of a level tone over the stressed syllable/mora (House, 1996, 1990). We also provided new empirical data showing that the falling nuclear contour typical for information-seeking interrogatives appears regularly in non-interrogative contexts, perhaps with a distribution characteristic of "uptalk".

In the discussion of the analysis of both the nuclear contours and sentence-medial edge tones of Samoan, we showed the value of transcending biases that have crept into AM theory in practice. Dissociating nuclear pitch accents, and pitch accents more generally, from various notions of prominence can help us make sense of IntP-final tonal events in Samoan. On the one hand, the placement of IntP-final tonal events appears to be insensitive to variation in discourse context.

But on the other hand, where they show up in the speech signal is clearly conditioned by stress placement. If we move away from the West Germanic-centric assumption that IntP-final accents are necessarily focal accents, then there is no puzzle about nuclear accents in Samoan. Instead, we can recognize the potential of a different positional generalization about focal accents in Samoan and perhaps other verb-initial languages: that they are IntP-initial by default rather than IntP-final (Calhoun, 2015, 2017, Calhoun et al., 2019, Kügler and Calhoun, 2020).

We also explicated the value of remembering Bruce (1977)'s fundamental insight to dissociate the grammatical source of a tone from "the phonological spot where it shows up on the surface" (Pierrehumbert, 2000, p. 19). Taking 'edge tone' only as a description of this "phonological spot"—without any implication that the grammatical source of the tone is necessarily a prosodic constituent—opens up the hypothesis space for intonational analysis. In Samoan, expanding possible grammatical sources of edge tones we consider to include morphosyntactic spellout helps us recognize that there could be sub-classes of edge tones from distinct grammatical sources, potentially with distinct signatures. In Samoan, distinct signatures might come in the distribution of edge tones with respect to syntactic structure, as proposed in Yu (2021): tones inserted in spellout invariably occur with certain syntactic structures, while tones inserted in phonological grammar do not. In constraint-based analyses of tonal association in Gussenhoven (2004), distinct signatures arise from how tones from different grammatical sources are ranked with respect to their priority to be associated to a TBU. For example, Gussenhoven (2004, §11.3)'s analysis of Stockholm Swedish ranks association to a TBU as highest priority for lexical accent tones, followed by the focal tone, and then IntP tones. In Samoan, a way we could account for the "overriding" of H_{abs} tones by L prosodic boundary tones would be to give priority of association to IntP tones over morphosyntactic H tones. Similarly, Torreira and Grice (2018) proposes an analysis of tunes in Spanish where different tones in the tune have different priority associations to TBUs rather than being classified statically as pitch accents and edge tones. Allowing flexibility in tonal association, together with opening up what grammatical sources we consider for tones, are just some of the ways that we can build on AM-theoretic foundations in approaching the intonation of understudied languages.

Appendix: speaker information

Demographic information about speakers and the list of figures that came from recordings of them are given in the tables below.

Table 2: Demographic information by speaker

Code	Location/date	Gender	Age	Samoan usage
la_m01	Los Angeles 2008-2016	M	19-22	
smo_s4	Apia Dec. 2011	M	26	
la_s22	Los Angeles Jan. 2012	F	57	
au_s02	Auckland Aug. 2019	F	51	50%
au_s04	Auckland Aug. 2019	M	18	40%
au_s11	Auckland July 2015, Aug. 2019	F	48-52	80%

Table 3: Figure list by speaker

Code	Figures
la_m01	12.1, 12.2, 12.11, 12.22, 12.23c
smo_s4	12.23a, 12.17
la_s22	12.23b
au_s02	12.5b, 12.7a, b; 12.8a, 12.16
au_s04	12.6, 12.12, 12.14, 12.15
au_s11	12.9, 12.5a; 12.8b, 12.13, 12.18, 12.19

Bibliography

Ahn, B. (2016), 'Syntax-phonology mapping and the Tongan DP', *Glossa* **1**(1), 4.1–36.

Armstrong, M. E. and del Mar Vanrell, M. (2016), Intonational polar question markers and implicature in american English and Majorcan Catalan, *in* 'Proceedings of Speech Prosody 2016', pp. 158–162.

Arvaniti, A. (2022), The Auotosegmental-Metrical model of intonational phonology, *in* J. A. Barnes and S. Shattuck-Hufnagel, eds, 'Prosodic Theory and Practice', MIT Press, Cambridge, MA, pp. 25–63.

Arvaniti, A. and Fletcher, J. (2020), The Autosegmental-Metrical theory of intonational phonology, *in* C. Gussenhoven and A. Chen, eds, 'The Oxford Handbook of Language Prosody', Oxford University Press, chapter 6, pp. 78–95.

Aziz, J. (2020*a*), Intonational phonology of Malagasy: pitch accents demarcate syntactic constituents, *in* 'Proceedings of the 10th International Conference on Speech Prosody 2020'.

Aziz, J. (2020*b*), A preliminary model of Malagasy intonation, Master's thesis, University of California Los Angeles, Los Angeles, CA.

Beckman, M. E. (1986), *Stress and non-stress accent*, Foris Publications, Dordrecht, The Netherlands.

Beckman, M. and Pierrehumbert, J. (1986), 'Intonational structure in Japanese and English', *Phonology Yearbook* **3**, 255–309.

Blust, R. (2013), *The Austronesian Languages: revised edition*, Vol. A-PL 008, Asia-Pacific Linguistics, Canberra, Australia.
URL: *http://hdl.handle.net/1885/10191*

Bruce, G. (1977), *Swedish word accents in sentence perspective*, CWK Gleerup, Lund.

Bruce, G. (1983), 'Accentuation and timing in Swedish', *Folia Linguistica* **17**, 221–238.

Bruce, G. (1987), How floating is sentence accent?, *in* K. Gregersen and H. Basbøll, eds, 'Nordic prosody IV', Odense University Press, Odense, pp. 41–49.

Bruce, G. and Gårding, E. (1978), A prosodic typology for Swedish dialects, *in* E. Gårding, G. Bruce and R. Bannert, eds, 'Nordic prosody: papers from a symposium', Vol. 13, Travaux de l'Institut de Linguistique de Lund, pp. 219–228.

Calhoun, S. (2010), 'The centrality of metrical structure in signaling information structure: A probabilistic perspective', *Language* **86**(1), 1–42.

Calhoun, S. (2015), 'The interaction of prosody and syntax in Samoan focus marking', *Lingua* **165, Part B**, 205–229.

Calhoun, S. (2017), 'Exclusives, equatives and prosodic phrases in Samoan', *Glossa* **2**(1), 11.1–43.

Calhoun, S., Wollum, E. and Va'ai, E. K. (2019), 'Prosodic prominence and focus: expectation affects interpretation in Samoan and English', *Language and Speech* .

Chafe, W. L., ed. (1980), *The pear stories: cognitive, cultural, and linguistic asspects of narrative production*, Ablex, Norwood, New Jersey.

Chung, S. (1978), *Case Marking and Grammatical Relations in Polynesian*, University of Texas Press, Austin, TX.

Churchward, S. (1951), *A Samoan Grammar*, 2nd edn, Spectator Publishing Co. Pty. Ltd., Melbourne, Australia.

Clemens, L. (2019), 'Prosodic noun incorporation: the relationship between prosody and argument structure in Niuean', *Syntax* **22**(4), 337–377.

Clemens, L. E. (2014), Prosodic noun incorporation and verb-initial syntax, PhD thesis, Harvard University, Cambridge, MA.

Clemens, L. E. and Polinsky, M. (2017), Verb-initial word orders, primarily in Austronesian and Mayan languages, *in* M. Everaert and H. C. van Riemsdijk, eds, 'The Wiley Blackwell Companion to syntax', 2nd edn, John Wiley & Sons, Inc., pp. 1–50.

Dainora, A. (2006), Modeling intonation in English: A probabilistic approach to phonological competence, *in* C. T. B. Louis Goldstein, D . H. Whalen, ed., 'Laboratory Phonology 8', Mouton de Gruyter, pp. 107–132.

D'Imperio, M. (2001), 'Focus and tonal structure in Neapolitan Italian', *Speech Communication* **33**, 339–356.

Duranti, A. (1981), *The Samoan fono: a sociolinguistic study*, number 80 *in* 'Pacific Linguistics Series B', Linguistic Circle of Canberra, Canberra, Australia.

Duranti, A. (1990), 'Code switching and conflict management in Samoan multiparty interaction', *Pacific Studies* **14**(1), 1–30.

Edmiston, D. and Postdam, E. (2016), Extraposition in Malagasy, *in* H. Hsieh, ed., 'AFLA 22: the proceedings of the 22nd meeting of the Austronesian Formal Linguistics Association', Asia-Pacific Linguistics, Canberra, Australia, pp. 121–138.

Edmiston, D. and Potsdam, E. (2017), Linearization at PF: Evidence from Malagasy extraposition, *in* A. Lamont and K. Tetzloff, eds, 'Proceedings of the 47th Annual Meeting of the North East Linguistic Society', Vol. 1, GLSA, Amherst, MA, pp. 295–308.

Féry, C. (2017), *Intonation and prosodic structure*, Cambridge University Press, Cambridge, UK.

Frota, S., D'Imperio, M., Elordieta, G., Prieto, P. and Vigaário, M. (2007), The phonetics and phonology of intonational phrasing in Romance, *in* P. Prieto, J. Mascaró and M.-J. Solé, eds, 'Segmental and prosodic issues in Romance phonology', John Benjamins Publishing Company, Amsterdam and Philadelpia, pp. 131–153.

Frota, S. and Prieto, P. (2015*a*), Intonation in Romance: systemic simliarities and differences, *in* S. Frota and P. Prieto, eds, 'Intonation in Romance', Oxford University Press, Oxford, England.

Frota, S. and Prieto, P., eds (2015*b*), *Intonation in Romance*, Oxford University Press, Oxford, England.

Goldsmith, J. (1981), English as a tone language, *in* D. L. Goyvaerts, ed., 'Phonology in the 1980's', John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 287–308.

Goldsmith, J. (1984), Tone and accent in Tonga, *in* G. N. Clements and J. Goldsmith, eds, 'Autosegmental studies in Bantu tone', Foris Publications, Dordrecht, The Netherlands, chapter 2, pp. 19–51.

Goldsmith, J. A. (1976), Autosegmental phonology, PhD thesis, Massachusetts Institute of Technology.

Grabe, E. (1998), 'Pitch accent realization in English and German', *Journal of Phonetics* **26**(2), 129–143.

Grice, M. (2022), Commentary on Chapter 1: introducing flexibility into Autosegmental-Metrical Phonology, *in* S. Shattuck-Hufnagel and J. A. Barnes, eds, 'Prosodic Theory and Practice', MIT Press, Cambridge, MA, pp. 64–75.

Gussenhoven, C. (2004), *The phonology of tone and intonation*, Cambridge University Press, Cambridge, UK.

Gussenhoven, C. (2011), Sentential prominence in English, *in* 'The Blackwell companion to phonology', Wiley-Blackwell, pp. 2778–2806.

Gussenhoven, C. (2016), 'Analysis of intonation: the case of MAE_ToBI', *Laboratory Phonology* **7**(1), 10.

Gussenhoven, C. (2018), Prosodic typology meets phonological representations, *in* L. M. Hyman and F. Plank, eds, 'Phonological typology', Walter de Gruyter GmbH, Berlin/Boston, pp. 389–418.

Gussenhoven, C. (2022), Just how metrical is the Autosegmental-Metrical model? evidence from pitch accents in Nubi, Persian, and English, *in* H. Kubozono, J. Ito and A. Mester, eds, 'Prosody and Prosodic Interfaces', Oxford University Press.

Hayes, B. and Lahiri, A. (1991), 'Bengali intonational phonology', *Natural Language & Linguistic Theory* **9**, 47–96.

Hellmuth, S. (2009), The (absence of) prosodic reflexes of given/new information status in Egyptian Arabic, *in* J. Owens and A. Elgibali, eds, 'Information structure in spoken Arabic', Routledge, Oxford, pp. 165–188.

Hellmuth, S. J. (2006), Intonational pitch accent distribution in Egyptian Arabic, PhD thesis, School of Oriental & African Studies, University of London.

Himmelmann, N. P. and Ladd, D. R. (2008), 'Prosodic description: an introduction for fieldworkers', *Language documentation and conservation* **2**(2), 244–274.

House, D. (1990), *Tonal perception in speech*, Lund University Press, Lund, Sweden.

House, D. (1996), Differential perception of tonal contours through the syllable, *in* 'Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on', Vol. 4, pp. 2048–2051.

Howard, M. (2018), The intonational phonology of Sāmoan questions. Poster presented at ProsLang – Workshop on the Processing of Prosody across Languages and Varieties.

Hsieh, H. (2016), Prosodic indicators of phrase structure in Tagalog transitive sentences, *in* H. Nomoto, T. Miyauchi and A. Shiohara, eds, 'AFLA 23: the Proceedings of the 23rd meeting of the Austronesian Formal Linguistics Association', Asia-Pacific Liinguistics, Canberra, Australia, pp. 111–122.

Hualde, J. I. (2002), Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestionis, *in* C. R. Wiltshire and J. Camps, eds, 'Romance Phonology and Variation : Selected papers from the 30th Linguistic Symposium on Romance Languages', John Benjamins Publishing Company.

Hualde, J. I. and Prieto, P. (2016), 'Towards an International Prosodic Alphabet (IPrA)', *Laboratory Phonology* **7**(1), 5.

Jun, S.-A. (2014), *Prosodic typology II: the phonology and phonetics of intonation and phrasing*, Oxford University Press, Oxford, England.

Jun, S.-A. (2022), Author response to the commentary: ToBI is not designed to be phonetically transparent, *in* J. A. Barnes and S. Shattuck-Hufnagel, eds, 'Prosodic Theory and Practice', MIT Press, Cambridge, MA, pp. 204–211.

Jun, S.-A. and Fletcher, J. (2014), Methodology of studying intonation: from data collection to data analysis, *in* S.-A. Jun, ed., 'Prosodic typology II: the phonology and phonetics of intonation and phrasing', Oxford University Press, Oxford, England, chapter 16, pp. 493–519.

Kawahara, S. and Shinya, T. (2008), 'The intonation of gapping and coordination in Japanese: evidence for Intonational Phrase and Utterance', *Phonetica* **65**, 62–105.

Krivokapić, J. (2007), 'Prosodic planning: Effects of phrasal length and complexity on pause duration', *Journal of Phonetics* **35**, 162–179.

Kügler, F. and Calhoun, S. (2020), Prosodic encoding of information structure: A typological perspective, *in* C. Gussenhoven and A. Chen, eds, 'The Oxford Handbook of Language Prosody', Oxford University Press, Oxford, UK, pp. 454–467.

Kuo, G. and Vicenik, C. (2012), 'The intonation of Tongan', *Working Papers in Phonetics, Department of Linguistics, UCLA* **111**, 63–91.

Ladd, D. R. (1986), 'Intonational phrasing: The case for recursive prosodic structure', *Phonology Yearbook* **3**, 311–340.

Ladd, D. R. (1995), "linear" and "overlay" descriptions: an autosegmental-metrical middle way, *in* K. Elenius and P. Branderud, eds, 'Proceedings ICPhS 95 Stockholm', Vol. 2, pp. 116–123.

Ladd, D. R. (1996), *Intonational phonology*, Cambridge University Press, Cambridge.

Ladd, D. R. (2000), Bruce, pierrehumbert, and the elements of intonational phonology, *in* M. Horne, ed., 'Prosody: theory and experiment', Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 37–59.

Ladd, D. R. (2008), *Intonational phonology*, 2nd edn, Cambridge University Press.

Ladd, D. R. (2022), The trouble with ToBI, *in* J. Barnes and S. Shattuck-Hufnagel, eds, 'Prosodic Theory and Practice', MIT Press, Cambridge, Massachusetts and London, England, pp. 247–257.

Massam, D. (2001), 'Pseudo noun incorporation in Niuean', *Natural Language & Linguistic Theory* **19**(1), 153–197.

Mayer, J. F. (2001), Code-switching in Samoan: t-style and k-style, PhD thesis, University of Hawaii.

Mosel, U. and Hovdhaugen, E. (1992), *Samoan reference grammar*, Scandinavian University Press, Oslo.

Nespor, M. and Vogel, I. (1986), *Prosodic phonology*, Foris Publications, Dordrecht, The Netherlands.

Orfitelli, R. and Yu, K. (2009), Intonational phonology of Samoan, *in* 'Presented at Austronesian Formal Linguistics Association XVI, University of California, Santa Cruz.'.

Pierrehumbert, J. (1980), The phonology and phonetics of English intonation, PhD thesis, Massachusetts Institute of Technology.

Pierrehumbert, J. (2000), Tonal elements and their alignment, *in* M. Horne, ed., 'Prosody: theory and experiment', Kluwer Academic Publishers, pp. 11–36.

Pierrehumbert, J. and Beckman, M. (1988), *Japanese Tone Structure*, The MIT Press.

Pierrehumbert, J. and Hirschberg, J. (1990), The meaning of intonational contours in the interpretation of discourse, *in* P. Cohen, J. Morgan and M. Pollack, eds, 'Intentions in communication', MIT Press, Cambridge, MA, pp. 271–311.

Post, B. (2000), *Tonal and phrasal structures in French intonation*, Thesus (Subsidiary of Holland Academic Graphics), The Hague.

Potsdam, E. and Polinsky, M. (To appear), Austronesian syntax, *in* B. Palmer, ed., 'Oceania', Mouton de Gruyter.

Prieto, P., D'Imperio, M. and Fivela, B. G. (2005), 'Pitch accent alignment in Romance: primary and secondary associations with metrical structure', *Language and Speech* **48**, 359–396.
URL: *http://www.ingentaconnect.com/content/king/ls/2005/00000048/00000004/art00003*

Richards, N. (2017), 'Some notes on Tagalog prosody and scrambling', *Glossa* **2**(1), 21.

Sabbagh, J. (2014), 'Word order and prosodic-structure constraints in Tagalog', *Syntax* **17**(1), 40–89.

Sabbagh, J. (2016), Intonation, adjunction, and verb-initial word order in Tagalog. Manuscript, University of Texas Arlington.

Selkirk, E. (2009), 'On clause and intonational phrase in Japanese: the syntactic grounding of prosodic constituent structure', *Gengo Kenkyu* **136**, 35–73.

Selkirk, E. O. (1978/1981), On prosodic structure and its relation to syntactic structure, *in* T. Fretheim, ed., 'Nordic prosody II', TAPIR, Trondheim, Norway, pp. 111–140.

Slifka, J. (2000), Respiratory constraints on speech production at prosodic boundaries, PhD thesis, Massachusetts Institute of Technology.

Swerts, M., Krahmer, E. and Avesani, C. (2002), 'Prosodic marking of information status in Dutch and Italian: a comparative analysis', *Journal of Phonetics* **30**(4), 629–654.
URL: *http://www.sciencedirect.com/science/article/pii/S0095447002901786*

Torreira, F. and Grice, M. (2018), 'Melodic constructions in Spanish: metrical structure determines the associaton properties of intonational tone', *Journal of the International Phonetic Association* **48**, 9–32.

Veilleux, N., Shattuck-Hufnagel, S. and Brugos, A. (2006), 6.911 transcribing prosodic structure of spoken utterances with ToBI. january IAP 2006, Technical report, Massachusetts Institute of Technology: MIT OpenCourseWare.
URL: *https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/lecture-notes/chap2_0and2_1.pdf*

Warren, P. (2016), *Uptalk: the phenomenon of rising intonation*, Cambridge University Press, Cambridge, UK.

Yu, K. M. (2011), The sound of ergativity: morphosyntax-prosody mapping in Samoan, *in* S. Lima, K. Mullin and B. Smith, eds, 'Proceedings of the 39th Annual Meeting of the North East Linguistic Society', Vol. 2, Graduate Student Linguistic Association, Amherst, MA, pp. 825–838.

Yu, K. M. (2018), Prosodically-informed parsing with minimalist grammars, *in* 'Minimalist parsing', Oxford University Press.

Yu, K. M. (2021), 'Tonal marking of absolutive case in Samoan', *Natural Language & Linguistic Theory* **39**, 291–365.

Yu, K. M. and Özyıldız, D. (2016), The absolutive ia particle in Samoan, *in* 'Proceedings of the forty-second annual meeting of the Berkeley Linguistics Society', Berkeley, CA, pp. 387–406.

Yu, K. M. and Stabler, E. P. (2017), '(in)variability in the Samoan syntax/prosody interface and consequences for parsing', *Laboratory Phonology* .

Zuraw, K., Yu, K. M. and Orfitelli, R. (2014), 'The word-level prosody of Samoan', *Phonology* **31**(2), 271–327.