NJAS NORDIC JOURNAL OF AFRICAN STUDIES

Pitch Reset in Akan Subordinate Clauses

Olivia Serwaa Oppong
Independent Scholar

Lotta Aunio
University of Helsinki
lotta.aunio@helsinki.fi
https://orcid.org/0000-0003-2261-8699

Martti Vainio
University of Helsinki
https://orcid.org/0000-0003-2570-0196

Abstract

This study explores how lexical tone interacts with intonation in Akan, focusing on pitch reset as a prosodic cue in complex declaratives. Pitch reset – an upward shift in fundamental frequency (f_0) following declination – has been documented in various African languages, and in Akan by Kügler (2016, 2017). However, its role in Akan remains underexamined. Whereas Kügler analysed embedded and coordinated clauses, using a controlled production experiment with five native speakers of Asante Twi, we analysed conditional and temporal subordinate clauses to determine whether pitch reset signals prosodic boundaries. Results show that Akan speakers employ pitch reset *before* subordinate conjunctions, often accompanied by a pause, to mark clause boundaries, while in Kügler's study pitch reset was not attested at all with coordinated clauses and was attested in complementizer sentences *after* the complementizer. Speakers pre-plan utterances by initiating the main clause at a higher pitch to maintain declination, and the degree of reset is influenced by lexical tone: resets are more pronounced on High tones and smaller on Low tones. These findings advance understanding of tone–intonation interaction in Akan and suggest that pitch reset functions as a robust cue to prosodic phrasing in tone languages. Limitations include sample size and potential effects of language attrition; future research should examine other conjunctions, matrix clause types, and dialectal variation.

Keywords: Akan, pitch reset, clausal tonology, subordination

DOI: 10.53228/e2ajy419

1 Introduction

In both tone and non-tone languages, fundamental frequency has the tendency to move downwards or decline during an utterance. This downtrend may be interrupted, causing a reset in the pitch contour, where a declined pitch contour returns to a higher level at the start of a new utterance or intonational unit. Pitch reset occurs at the boundaries between prosodic units and can signal the close relationship between one unit of speech and another; f_0 resets may thus be significant cues for phrasing structure (Gussenhoven 2004). Pitch reset has been explored in non-tonal or marginally tonal languages such as Swedish, Dutch, and Japanese (Swerts et al. 1996; Truckenbrodt 2002; Kawahara and Shinya 2008; Ishihara 2016); in African tone languages, pitch reset has been studied at least in Akan, and in the Bantu languages Ékegusií, Bemba, and Chichewa (Kügler 2016, 2017; Hieber 2016; Kula and Hamann 2017; Downing 2017).[1]

In a tone language like Akan, the phenomenon of pitch reset may not be straightforward since f_0 variations from lexical tones may interact with and shape the intonational contour of utterances. Kügler (2017) observed pitch reset in complex complementizer (embedded) sentences in Akan: the downward movement of the pitch contour was interrupted with a reset occurring at the left edge of embedded clauses; that is, the prosodic boundary of the embedded clause was signalled through pitch register reset. Coordinated sentences, on the other hand, were not separated by pitch reset. Both embedded and coordinated sentences involved a pause, which was realized after the complementizer but before the conjunction, signalling different prosodic phrasing in embedded and conjoint complex declaratives.

Kügler (2016, 2017) are the only studies on pitch reset in the Akan language to our knowledge. Hence, the present study is an extension of Kügler's work, and it concerns how lexical tones interact with pitch reset in the Asante Twi dialect. This study sets out to test pitch reset in different types of complex declaratives than those in Kügler (2017), that is, in conditional and temporal subordinate clauses. In this controlled experiment, we seek to advance insights on the tone and intonation structure of the Akan language, and to find out precisely how prosody is used to mark conditional and temporal subordinate clauses in Akan, as exemplified in (1) and (2). Based on Kügler's (2017) findings on complementizer sentences, our starting hypothesis was that pitch reset, accompanied by a pause, occurs after the clausal marker in complex declarative utterances in Akan.

(1)	*Ámꜜá*	*ǹ-tɔ́*	*sófì*	*gyé*	*sɛ̀*	*Mánsá*	*á-bà.*[2]
	Ama	NEG-buy	shovel	take	if	Mansa	CONS-come
	'Ama won't buy a shovel unless Mansa comes.'						

(2)	*Ámꜜá*	*ǹ-tɔ́*	*sófì*	*ànsá*	*nà*	*Mánsá*	*á-bà.*
	Ama	NEG-buy	shovel	before	FOC	Mansa	CONS-come
	'Ama won't buy a shovel before Mansa comes.'						

[1] We thank the individuals who took part in the experiment for their valuable input, which made this study possible. We are grateful to the anonymous reviewers for their constructive feedback and valuable recommendations that helped refine the analysis and presentation. This paper is based on the MA thesis of the first author at the University of Helsinki.

[2] The Akan orthography is used in the examples. The orthography does not mark tones, but these have been indicated in the examples, as pronounced by the first author.

In the following sections, we first present some background on the Akan language and on tone in Akan. Section 2 describes our data and methods. The results of the experiment are presented in Section 3 and discussed in Section 4. The final section presents some conclusions and further topics for research.

1.1 The Akan language

Akan (ISO 639-3 [twi]) is a Kwa language belonging to the Niger-Congo language family. It is a language widely spoken in Ghana, among over 70 indigenous languages spoken in the country. Akan is the first language of a large percentage of the Ghanaian population, and it is often used as a second language by others throughout the country. Eberhard et al. (2021) indicated that Akan has about 9.1 million speakers, with about 8.1 million first language speakers in Ghana.

Akan is a cluster of several mutually intelligible dialects with two main dialects, Fante and Twi. The Fante dialect is further divided into sub-dialects, including Agona, Ekumfi, Iguae, Breman, Nkusukum, and Gomoa, and the Twi dialect is divided into sub-dialects including Akyem, Akuapem, Asante, Wassa, Bono, and Kwahu. Fante is primarily spoken in the Central, Western North, and Western regions, while Twi is primarily spoken in the Eastern, Ashanti, Bono, Ahafo, and Bono East regions of Ghana. The current study will focus on the Asante Twi dialect, mainly spoken by the Ashanti ethnic group located in the Ashanti region. Asante Twi is spoken by over 3.8 million people in Ghana (Eberhard et al. 2021).

Akan has a subject-verb-object (SVO) word order. In simple transitive and intransitive declarative sentences, the subject argument precedes the predicate, and the object argument follows the predicate within transitive constructions. In noun phrases, determiners and adjectives follow the noun (Boadi 2010; Saah 1994).

1.2 Tone and intonation in Akan

Akan is a level-tone language that uses different pitch levels to distinguish between the meanings and grammatical categories of words. The language makes use of two contrastive level tones, High (H) and Low (L), transcribed as [´] and [`], respectively. The functional load carried by tone in grammar is more significant than in the lexicon (Dolphyne 1986, 1988). Lexical tone contrasts can be found in adverbs, verbs, nouns, and adjectives. Kügler (2017) provides a short summary of the studies on the Akan tone system over the past 50 years.

Abakah (2005) asserted that the tone bearing-unit (TBU) in Akan is the syllable, and he distinguished between three main types of syllables: open syllables (V and CV), as in *á-bà* 'CONS-come', as well as syllables with a sonorant (C) which acts as a syllabic nucleus, as *ǹ* in *ǹ-tɔ́* 'NEG-buy', both shown in context in examples (1) and (2) above. He also reported a third syllable type, CC and suggested that it might be a complex onset of a CCV syllable type in Akan, although this argument is yet to be established. All syllabic consonants and vowels carry tone in Akan (Manyah 2014). Word-final non-vowel sonorants, post-consonantal liquids, and preconsonantal nasal consonants are also tone-bearing units since they are syllabic in Akan, and two contrastive tones which are linked in Akan can produce gliding or contour tones if the tone-bearing units, VV, CV, or VC, are adjacent (Abakah 2015).

Akan has a downstep system where the presence (downdrift/automatic downstep) or absence (non-automatic downstep) of an overt L tone may cause the downward movement of successive H tones (Kügler 2017). In automatic downstep, a H tone preceded by a L tone in a sequence /HLH/ is realized with a lower pitch than the preceding H tone and indicated as !H. Non-automatic downstep presupposes a lexical L tone that has been delinked and is thus floating, causing the following H tone to be realized at a lower pitch than the preceding H tone. In addition to downstep, Kügler (2017) reported that in Akan utterances where only H tones or only L tones occurred, the pitch contour declined throughout the utterances.

Whether automatic and non-automatic downsteps are the same or two different phonetic processes in Akan has been debated by researchers. Stewart (1965) and Genzel and Kügler (2011) equated automatic downstep with non-automatic downstep, similarly to Snider (2007), who posited that automatic and non-automatic downsteps are the same phonetic processes in Chumburung, a language spoken in the Northern and Volta regions of Ghana. On the other hand, Dolphyne (1994) argued that automatic and non-automatic downstep are two different phonetic processes in Akan.

In addition to the downtrends, Akan H tones spread across word boundaries to the following word's L-toned initial syllable (Kügler 2017). The spread H tone may or may not be downstepped, depending on the context. For a more detailed account of the Akan tone system, see, for example, Kügler (2017), Marfo (2004), and Paster (2010).

There is also a preplanning effect in Akan where speakers start utterances at a higher pitch to maintain declination in speech (Genzel 2013). As in non-tone languages, the pitch contour serves as an indicator of different sentence types in Akan. The intonation of a simple declarative in Akan shows the downward movement of the pitch pattern to a phrase-final L tone.

Kügler's (2016, 2017) studies revealed pitch reset at the beginning of the content word that followed an embedded clause marker. The clausal marker was prosodically phrased with the matrix clause, and a pause followed the clausal marker. Kügler reported that the pause which appeared after the clausal marker indicated the close connection between the clausal marker and the matrix clause. Kügler also asserted that pitch reset was higher when a H tone occurred at the start of the content word and was smaller when a L tone occurred at the beginning of the content word, implying that the degree of pitch reset was dependent on the tone structure.

2 Data and methods

To test the hypothesis – that pitch reset occurs after the clausal marker in complex declarative sentences in Akan – we performed a controlled experiment. This experiment consisted of 64 sentences that were recorded by five speakers.

The sentences in the experiment were first controlled for the conditional subordinate conjunction *gyé sɛ̀* 'unless' (Conjunction 1; C1) and the temporal subordinate conjunction *ànsá nà* 'before' (C2).[3] These conjunctions have different lexical tones and segmental structure but can be used in otherwise identical sentences. Subordinate clauses are understood here as events that are not autonomous, but that relate to the event of the matrix clause (Cristofaro 2003).

In addition, we controlled for the tone before and after the conjunction. The first syllable of the disyllabic word before the conjunction was labelled T1 (Tone 1), and the second syllable T2. The first syllable of the disyllabic word after the conjunction was labelled T3 and the second syllable T4. The four complex sentences were formed from the first conjunction (C1) by giving T2 and T3 all the possible tonal combinations (H&H, H&L, L&H, L&L) to determine the boundary at which the pitch reset occurs. Two trisyllabic words were also used, but only the tone patterns of the second and final syllables were considered. The tones of the subject of the main clause and the verb of the subordinate clause were not controlled for tone as they are not adjacent to the conjunction.

[3] Both conjunctions are written disjunctively in Akan orthography, and therefore glossed separately in the examples; the meaning of the conjunction is reflected in the idiomatic translation. *nà* of 'before' is glossed here as a focus marker following van Dommelen and Beermann (2019), who describe it as giving prominence to *ànsá*.

Two types of sentence frames were used: transitive and intransitive verbs. Two transitive verbs, *ǹ-tɔ́* 'NEG-buy' and *m̀-frɛ́* 'NEG-call', were used to form two sentence frames in which the object noun preceded the conjunction and therefore carried T1 and T2. In example (3), repeated here from (1), the first syllable *só* of *sófì* carries T1 and the second syllable *fì* carries T2. With the four intransitive verbs used, *ń-kàsá* 'NEG-talk', *ǹ-sà* 'NEG-dance', *ń-nìdí* 'NEG-eat', and *ǹ-nà* 'NEG-sleep', the verb itself preceded the conjunction, as in example (4). The first syllable of the disyllabic verbs *ǹ-sà* 'NEG-dance' and *ǹ-nà* 'NEG-sleep' was labelled as T1 (syllable *ǹ* in both verbs) and the second syllable as T2 (syllables *sà* and *nà*, respectively). For the trisyllabic verbs *ń-kàsá* 'NEG-talk' and *ń-nìdí* 'NEG-eat', the second (*nì*) and last syllables (*dí*) were labelled as T1 and T2 respectively (see Table 1).

(3) *Ámꜜá ǹ-tɔ́ sófì gyé sɛ̀ Mánsá á-bà.*
Ama NEG-buy shovel take if Mansa CONS-come
'Ama won't buy a shovel unless Mansa comes.'

(4) *Kòfí ǹ-nà gyé sɛ̀ wɔ̀fà á-dà.*
Kofi NEG-sleep take if uncle CONS-sleep
'Kofi won't sleep unless uncle sleeps.'

The word after the conjunction is the subject of the subordinate clause in all sentences, as with *Mánsá* in example (3) and *wɔ̀fà* in example (4) above. These were all disyllabic words with their syllables labelled as T3 and T4.

The sentences were formed by giving T2 and T3 all the possible tonal combinations. Combining two transitive verbs with both conjunctions gives four sentences. When also controlled for the four combinations of tones before and after the conjunction, the number of sentences is 16. Combining the four intransitive verbs with both conjunctions and with the possible tone variations of the first syllable of the subject of the subordinate clause, 16 sentences were formed. To increase the number of sentences, 64 complex sentences were then produced by varying the subject noun phrases in the matrix clause. Since the tonal patterns of the subjects in the matrix clauses were not controlled for, the subjects were randomly changed, for example *Áꜜmá* in example (3) above was replaced with *Kòfí*.

Table 1 lists the words used before and after the conjunctions in the recorded sentences and shows how they were combined. The first four rows represent the four tonal combinations (LH, LL, HH, HL) of the syllables preceding (T2) and following (T3) with the conjunction *gyé sɛ̀* 'unless' (C1), and the last four rows with conjunction *ànsá nà* 'before' (C2). For example, the first row represents sentences like that in example (3) above, where T2 (the tone preceding the conjunction, that is, the final syllable of *sófì*) is L and T3 (the tone following the conjunction, that is, the first syllable of *Mánsá*) is H.

Table 1: Tonal combinations and words used in the experiment

Preceding word (T1_T2)	Tonal combinations T2 and T3	Following word (T3_T4)
sóf̀ì (shovel), *Ásɔ̀* (female given name), *ǹ-sà* (NEG-dance), *ǹ-nà* (NEG-sleep)	L_C1_H	*Mánsá* (female given name)
sóf̀ì (shovel), *ǹ-sà* (NEG-dance), *ǹ-nà* (NEG-sleep)	L_C1_L	*Yàà* (female given name), *wɔ̀fà* (uncle)
àbɛ́ (palm nut), *Kòfí* (male given name), *ń-nìdí* (NEG-eat), *ń-kàsá* (NEG-talk)	H_C1_H	*Mánsá* (female given name)
àbɛ́ (palm nut), *Kòfí* (male given name), *ń-nìdí* (NEG-eat), *ń-kàsá* (NEG-talk), *Mánsá* (female given name)	H_C1_L	*wɔ̀fà* (uncle)
sóf̀ì (shovel), *ǹ-sà* (NEG-dance), *ǹ-nà* (NEG-sleep)	L_C2_H	*Mánsá* (female given name)
sóf̀ì (shovel), *Ásɔ̀* (female given name), *ǹ-sà* (NEG-dance), *ǹ-nà* (NEG-sleep)	L_C2_L	*Yàà* (female given name), *wɔ̀fà* (uncle)
àbɛ́ (palm nut), *Kòfí* (male given name), *ń-nìdí* (NEG-eat), *ń-kàsá* (NEG-talk)	H_C2_H	*Mánsá* (female given name)
àbɛ́ (palm nut), *Kòfí* (male given name), *ń-nìdí* (NEG-eat), *ń-kàsá* (NEG-talk), *Mánsá* (female given name)	H_C2_L	*wɔ̀fà* (uncle)

In addition to the 64 target sentences, 32 sentences were used as fillers. The fillers were used to prevent participants from creating any pattern when producing the sentences, which might affect the naturalness of the speech data. The fillers were formed with the same syntactic structure as the target sentences, but with different lexical items. The fillers were taken out of the data before analysis.

Excluding the filler sentences, the syllables of the subordinate conjunctions and the syllables of the words before and after the conjunctions in all the sentences were manually segmented and labelled in Praat (Boersma and Weenink 2021). The f_0 of the syllables was automatically measured in Praat using the ProsodyPro script (Xu 2013). The script was used to extract the mean f_0 values of the syllables. The f_0 values obtained were analysed in R; the statistical analysis relied on a linear mixed model with speaker and item (eight tonal combinations) as random effects. The difference in the mean f_0 of the syllables was calculated in semitones; this was justified based on the fact that the perception of tone and intonation in terms of f_0 is logarithmic in nature and a relative scale removes the effect of individual pitch level and, more importantly, the differences between speakers' sex, as males tend to have a lower f_0 than females (Simpson 2009).

Five native speakers of Asante Twi participated in the experiment. Three of the speakers were males, and two were females. The age of the speakers ranged from 28 years to 43 years, with a mean age of 34 years. None of the participants reported any speech or hearing-related problems, and they all resided in the Helsinki region in Finland. Two participants were researchers (not in linguistics), and one was a student at the university. For the other two speakers, one was a nurse, and the other worked in a company. All the speakers reported having English as their second language; all the participants also provided written consent to participate in the study.

The audio recordings were made in the homes of the participants in quiet rooms. Each participant was presented with a paper that had 96 sentences in different randomized orders. Before the participants produced the sentences, they were given some time to familiarize themselves with the sentences. The participants were digitally recorded at a sample rate of 44.1 kHz on a Lenovo laptop using Audacity (Version 2.4.2), the Audient iD4 USB audio interface, and a DPA d:fine Headset microphone with omnidirectional sensitivity. The whole experiment lasted about 30 minutes, with participants being given a short break after producing every ten sentences.

3 Results and analysis

This section presents the pitch reset and pause patterns observed in the data, and displays results from the statistical analyses conducted on the mean f_0 values. Pitch reset is represented by "%reset", which is adopted from Peng et al. (2005).

3.1 Descriptive results of pitch reset

While in Kügler's (2017) study the pitch reset occurred *after* the complementizer, in our data the reset occurred *before* the two subordinate conjunctions *gyé sɛ̀* 'unless' and *ànsá nà* 'before'. While there were a few instances where two speakers realized a pause between the matrix clause and the dependent clause, three speakers consistently inserted a pause between the two clauses. When the speakers realized a pause between the clauses, it was consistently *before* the clausal marker, not after; prosodically, the subordinate conjunction was always phrased with the subordinate clause.

Figure 1 depicts the complex sentence in (5) produced by a female speaker. The f_0 contour of the matrix clause *Kòfí ǹ-tɔ́ àbɛ́* 'Kofi will not buy palm nut' moves downwards, and it is followed by a pause (<p>). After the pause, there is a pitch reset in the first syllable of *ànsá nà* 'before'. The tone of *àn* is approximately on the same level as the H tone of the syllable *bɛ́* preceding the pause. However, the reset is not total as the L of *àn* is not on the same level as the initial L tone of the matrix clause. Note that the final H of the noun *Kòfí* spreads to the L-initial first syllable of the following verb *ǹ-tɔ́* 'NEG-buy' ; the L is dissociated from its segment, and the H tone of the verb surfaces as a downstepped H (!H). The figures also illustrate a higher pitch range in the main clause than in the subordinate clause in example (5).

(5)	*Kòfí*	*ǹ-tɔ́*	*àbɛ́*	*ànsá*	*nà*	*Mánsá*	*á-bà*
	Kofi	NEG-buy	palm_nut	before	FOC	Mansa	CONS-come

'Kofi will not buy palm nut before Mansa comes.'

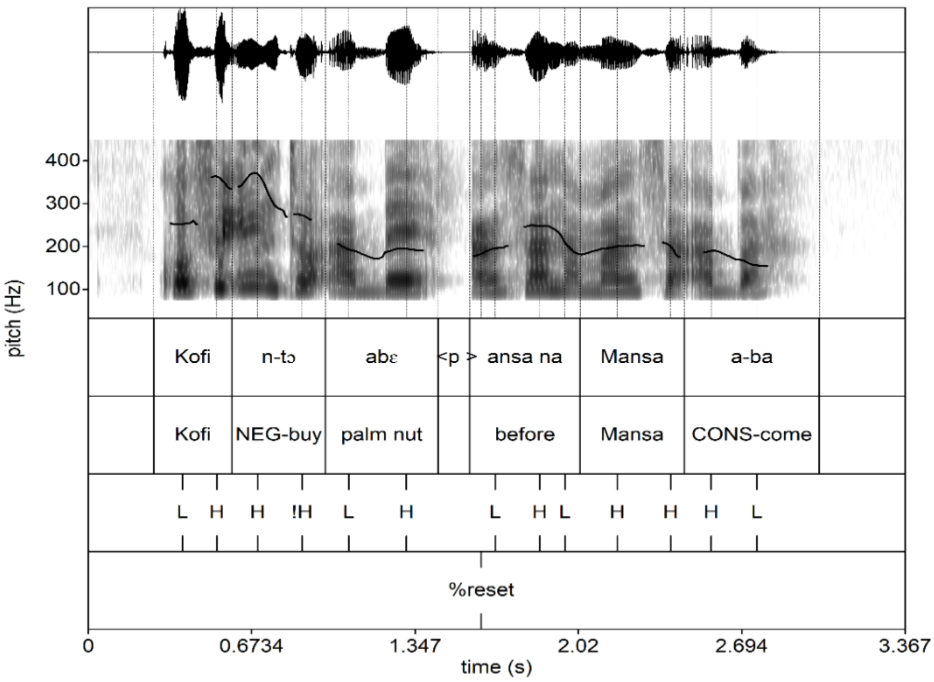

Figure 1: Waveform, spectrogram, and smoothed f_0 contour of sentence (5)

After the pause in Figure 2, there is a reset in the initial syllable tone of *ànsá nà* 'before'. Since both the tones preceding and following the pause are L, the reset is seen more clearly in the H-toned second syllable of the conjunction, which is realized approximately on the same level as the H on the word *sófì* (T1) in the matrix sentence. Within the clause all subsequent H tones are realized with a lower pitch after an intervening L tone. The downstepped H final syllable of the noun *Á'má* also spreads to the L-initial syllable of the following verb *ǹ-tɔ́* 'NEG-buy'. The underlying initial L becomes a floating tone, causing the following H tone to downstep (!H). The figure also shows that the pitch range of the main clause is higher than the pitch range of the subordinate clause in example (6).

(6) *Ám'á ǹ-tɔ́ sófì ànsá nà Mánsá á-bà.*
Ama NEG-buy shovel before FOC Mansa CONS-come
Ama will not buy a shovel before Mansa comes.'

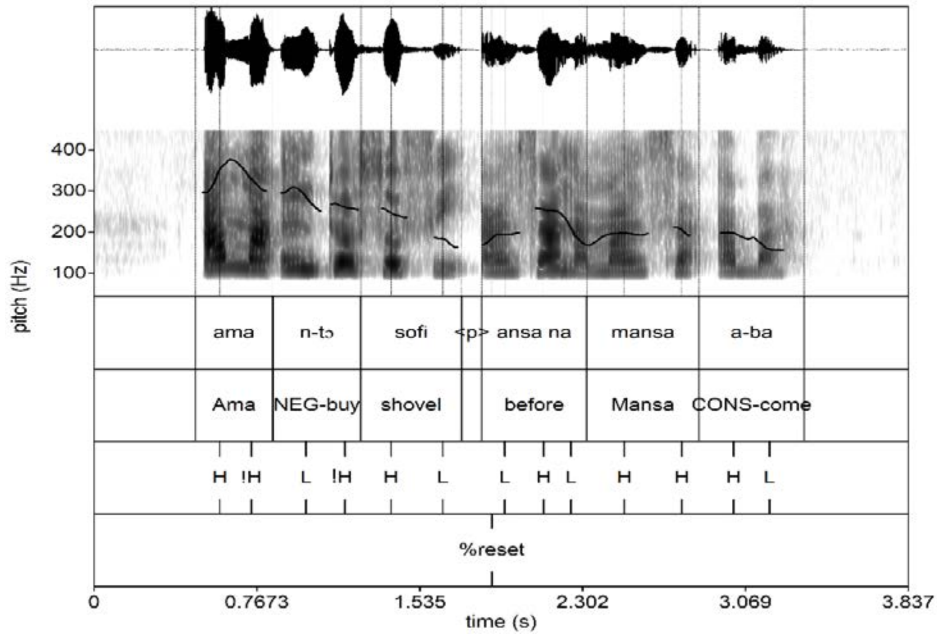

Figure 2: Waveform, spectrogram, and smoothed f_0 contour of example (6)

Figures 3 and 4 illustrate a pitch reset in *gyé sɛ̀* ‘unless’ in the complex sentences in examples (7) and (8). In Figure 3, which depicts example (7), also produced by a female speaker, the speaker did not insert a pause after the main clause *Áꜝmá ǹtɔ́ àbɛ́* ‘Ama will not buy palm nut’. The pitch reset occurs at the start of the conjunction *gyé sɛ̀* ‘unless’ and the initial H of the conjunction is much higher than the H of the preceding word.

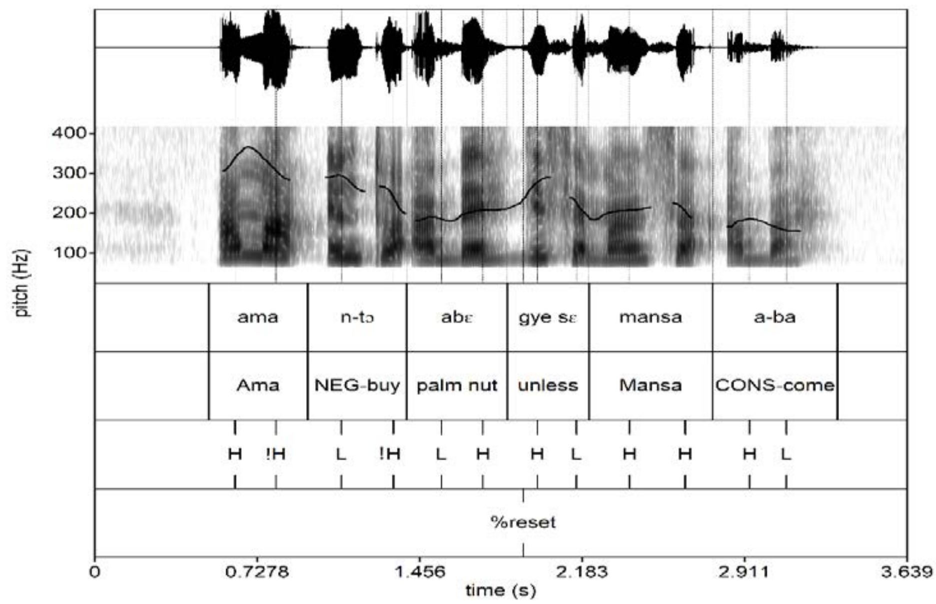

Figure 3: Waveform, spectrogram, and smoothed f_0 contour of example (7)

(7) *Ám!à ǹ-tɔ́ àbɛ́ gyé sɛ̀ Mánsá á-bà.*
Ama NEG-buy palm_nut take if Mansa CONS-come
'Ama will not buy palm nut unless Mansa comes.'

In Figure 4, the f_0 contour of the main clause, *Kòfí ǹtɔ́ sófì* 'Kofi will not buy a shovel' (example 8), declines to the lower end of the speaker's pitch range, and it is followed by a reset at the start of *gyé sɛ̀* 'unless', as the H on the first syllable is somewhat higher than the preceding H on the first syllable of *sófì* (T1). The pitch contour of *Yàà ábà* 'Yaa comes' also glides down to a lower level.

(8) *Kòfí ǹ-tɔ́ sófì gyé sɛ̀ Yàà á-bà.*
Kofi NEG-buy shovel take if Yaa CONS-come
'Kofi will not buy a shovel unless Yaa comes.'

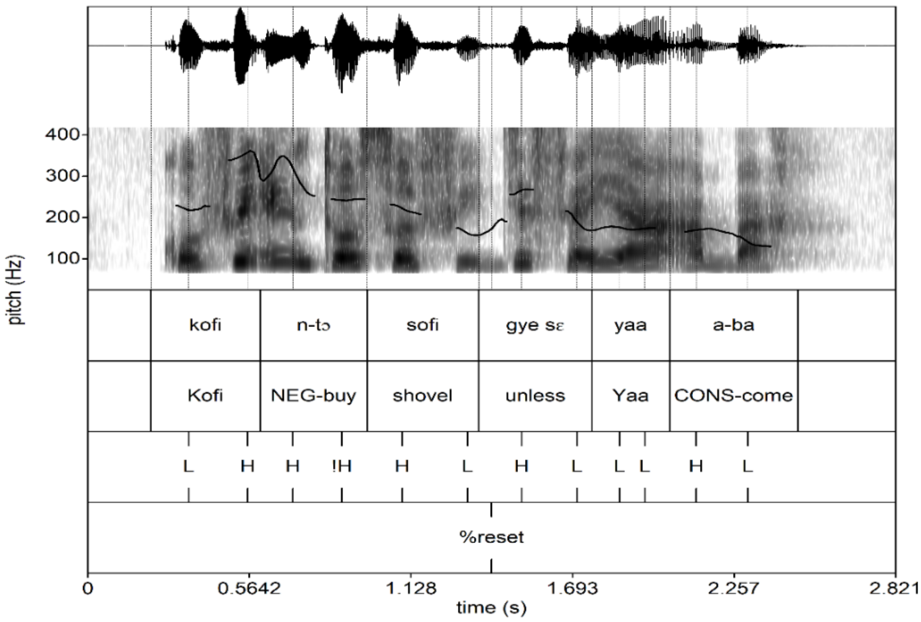

Figure 4: Waveform, spectrogram, and smoothed f_0 contour of sentence (8)

3.2 Statistical analysis of the tonal combinations

Figure 5 illustrates the mean f_0 values of the eight tonal combinations used to form the complex sentences. The numbers represent the consecutive syllables of the word that precedes the conjunction (syllables 1 and 2, i.e., T1 and T2), the conjunction (syllables 3, 4, and 5), and the word that follows the conjunction (syllables 6 and 7, i.e., T3 and T4). Syllable 3 has no value in the tonal combinations of the first conjunction (C1) *gyé sɛ̀* 'unless' since the conjunction is disyllabic, causing the discontinuity in the figure. The male speakers are on the lower level, and the female speakers are on the upper level. Sentences with the first conjunction (C1) are marked with black, while the red lines represent sentences with the second conjunction (C2). Across both the male and the female speakers, Figure 5 shows that the H tone of both C1 and C2 is higher than any syllable of the following word, demonstrating that the reset occurs before the subordinate conjunction.

Figure 5 also demonstrates that the initial H of the first conjunction (C1) *gyé sɛ̀* 'unless' (syllable 4 in black) is always higher than the tone of the preceding syllable (T2), regardless of its lexical tone. However, the difference is not notable with the second conjunction (C2) *ànsá nà* 'before', where the initial tone of the conjunction is L: syllable 3 (in red) appears to be on the same or a slightly lower level than the preceding syllable (T2), but again the lexical tone of the preceding syllable does not seem to play a role.

The lexical tone specification of syllable 2 (T2) seems to be neutralized in the context of the pitch reset. On the other hand, after the conjunctions, the H and L tones of syllable 6 (T3) appear distinctive. Female speakers pronounce the H of C1 somewhat higher than the H of C2 (syllable 4), possibly due to the lowering effect of the first L syllable of C2, but the male speakers do not make this difference.

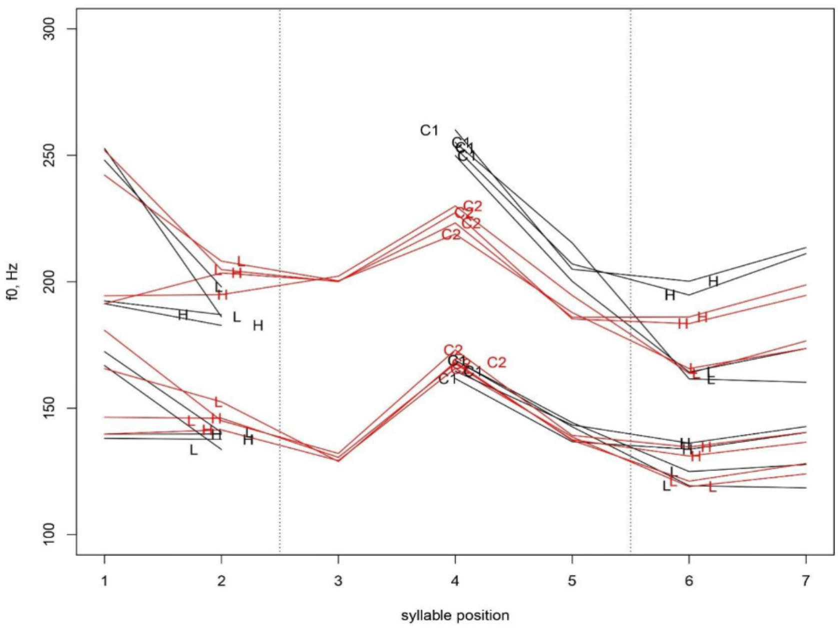

Figure 5: Tonal combinations for the male and female speakers

3.2.1 Pitch difference between the conjunction and the following word

The means and standard deviations of the tonal difference between the final syllable of the conjunction (L tone in both conjunctions) and the initial syllable of the following content word are illustrated in Tables 2a and 2b, as well as 3a and 3b. The tonal combinations are those of T2 and T3, that is, the tones preceding and following the conjunctions.

Table 2: a. Means of C1

Conjunction 1		
	H	L
H	0.5606826	3.825365
L	0.6165494	2.812227

b. Means of C2

Conjunction 2		
	H	L
H	0.3784224	2.626949
L	0.2924390	2.051545

Table 3: a. Standard deviations of C1

Conjunction 1		
	H	L
H	1.508593	1.602018
L	1.918468	1.777229

b. Standard deviations of C2

Conjunction 2		
	H	L
H	1.573614	1.508601
L	1.601063	1.515964

Figure 6 depicts the boxplot for the tonal difference. It shows that the final L tone of the conjunction is always higher than the initial tone of the following content word (T3), indicating that no pitch reset occurs after the conjunction. For the disyllabic conjunction, the difference is small but non-zero when the initial tone of the following word is H (first two boxes), but is larger when the initial tone of the following word (T3) is L (3rd and 4th boxes). The differences are similar when the conjunction is trisyllabic. In the trisyllabic conjunction, the difference is smaller when the initial tone of the following word (T3) is H (5th and 6th boxes) but relatively larger in the case of a following L tone (7th and 8th boxes). Regardless, the final tone of the conjunction is never lower than the following tone (T3).

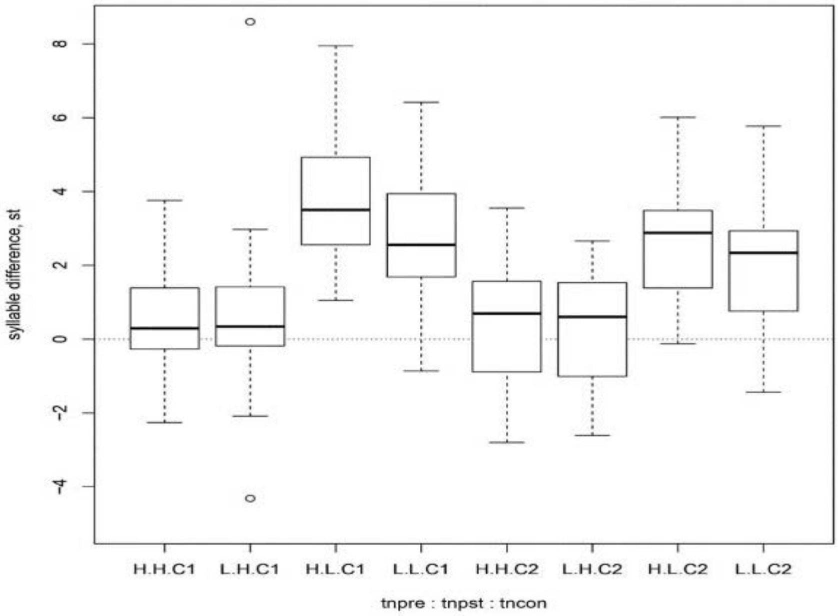

Figure 6: Tonal difference in semitones between the final syllable of the conjunction and the initial syllable of the following word (*tnpre* = tone of the preceding syllable, *tnpst* = tone of the following syllable, *tncon* = tone of the conjunction)

In order to assess the statistical significance of the effects of different tonal contexts on the pitch difference between the last syllable of the conjunction (always L) and the first syllable of the following word (H or L depending on the word), a linear model was fitted. More specifically, a linear mixed model was fitted using the *lme4* package in R (Bates et al. 2015). The models were fitted to predict syllable difference in semitones with the tone of the preceding syllable (H or L), conjunction type (tri- vs disyllabic), and tone of the following syllable (H or L) as fixed factors. The model included speaker and item (eight tonal combinations) as random effects.

The model's total explanatory power was substantial (conditional R^2 = 0.69), and the part related to the fixed effects alone (marginal R^2) was at 0.35. The model's intercept, corresponding to *tnpre* = High, *tncon* = C1, and *tnpst* = High, was at 0.97 (95% CI [0.19, 1.76], t(314) = 2.42, p <.05). Within this model:

a. The effect of *tnpre* [Low] was non-significantly negative (beta = -0.41, 95% CI [-1.18, 0.37], t(314) = -1.03, p = 0.304; Std. beta = -0.20, 95% CI [-0.58, 0.18]).
b. The effect of *tncon* [C2] was non-significantly negative (beta = -0.61, 95% CI [-1.39, 0.16], t(314) = -1.55, p = 0.122; Std. beta = -0.30, 95% CI [-0.68, 0.08]).
c. The effect of *tnpst* [Low] was significantly positive (beta = 2.36, 95% CI [1.59, 3.14], t(314) = 5.96, p < .001; Std. beta = 1.15, 95% CI [0.77, 1.53]).

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% confidence intervals (CIs) and p-values were computed using the Wald approximation. In summary, the pitch level of the last syllable of the conjunction (always L) was the only tonal feature that significantly increased the difference between the syllables. More importantly, the difference was always positive, as shown by the estimated intercept. Regardless of the tone type of the following syllable, the last syllable of the conjunction was consistently higher in pitch than the following syllable.

3.2.3 Pitch difference between the conjunction and the preceding word

The means and standard deviations of the tonal difference between the last syllable of the preceding word and the H tone of the conjunctions (initial syllable of the first conjunction *gyé sὲ* 'unless' and medial syllable of the second conjunction *ànsá nà* 'before') are indicated in Tables 4a, 4b and 5a, 5b. Note that the L tone of the trisyllabic conjunction *ànsá nà* 'before' was not considered.

Table 4: a. Means of C1

Conjunction 1		
	H	L
H	-3.839128	-4.393892
L	-4.144874	-4.025599

b. Means of C2

Conjunction 2		
	H	L
H	-2.118915	-2.765566
L	-2.537841	-2.151170

Table 5: a. Standard deviations of C1

Conjunction 1		
	H	L
H	2.097084	2.636278
L	2.798901	2.077119

b. Standard deviations of C2

Conjunction 2		
	H	L
H	2.230436	1.89631
L	1.898981	2.78828

The tonal differences between the final syllable of the preceding word and the H tone of the conjunctions are represented as a boxplot in Figure 7. The first four boxes illustrate the difference in the female group, and the last four boxes indicate the difference in the male group.

Figure 7 demonstrates that the final tone of the preceding word (T2) is always lower than the H tone of the conjunctions (i.e., the difference is always positive), whether the preceding tone is H (1st and 3rd boxes in the groups) or L (2nd and 4th boxes in the groups). The H tone of the conjunctions is systematically higher by about four semitones when the conjunction is disyllabic *gyé sὲ* 'unless' (H-L) but over 1.5 semitones lower when the conjunction is trisyllabic *ànsá nà* 'before' (L-H-L). Interestingly, the differences are only determined by the syllable count of the conjunction, as there are no differences within the two groups determined by the tones around the conjunction.

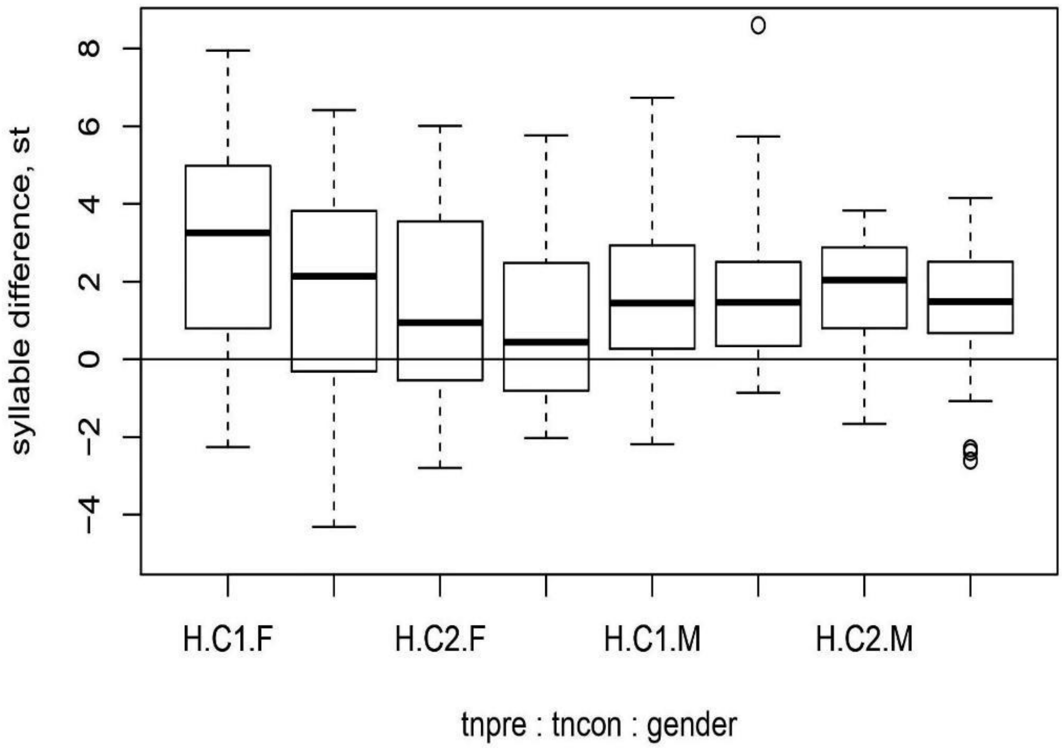

Figure 7: Tonal difference between the final syllable of the preceding word and the High tone of the first conjunction (C1) and the second conjunction (C2). The first four boxes illustrate the difference in the female group, and the last four boxes depict the difference in the male group. Within the groups, the 2nd and 4th boxes represent a preceding L tone.

Figure 8 shows the boxplot for the tonal difference within the syllables of the word that preceded the conjunction. The figure demonstrates that the H tone within the preceding word is H, and the L tone is also L. The difference between the H tone and the L tone is about two semitones.

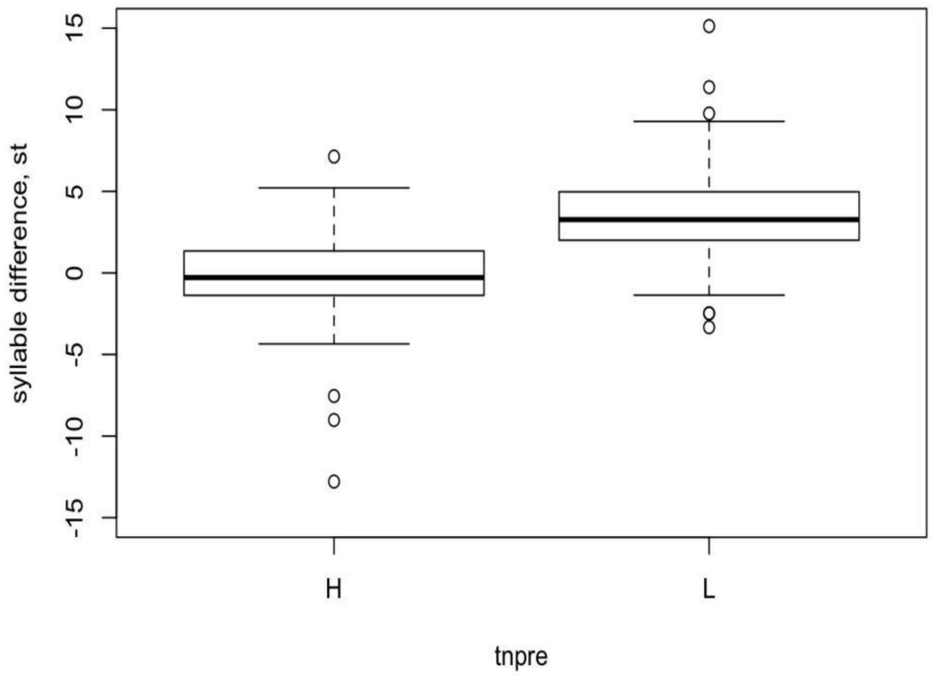

Figure 8: Tonal difference within the syllables of the preceding word

3.3 Distribution of pause patterns across utterances

Overall, utterances with pauses occur more frequently than those without pauses. In total, 56% of the recorded utterances have a pause before the conjunction. Clear speaker-specific variation is evident in Figure 9: three speakers (SP02, SP03, and SP04) produce pauses in the majority of their utterances, with SP03 showing the strongest tendency to pause. In contrast, SP01 and SP05 rarely insert pauses, indicating a distinctly different pausing style.

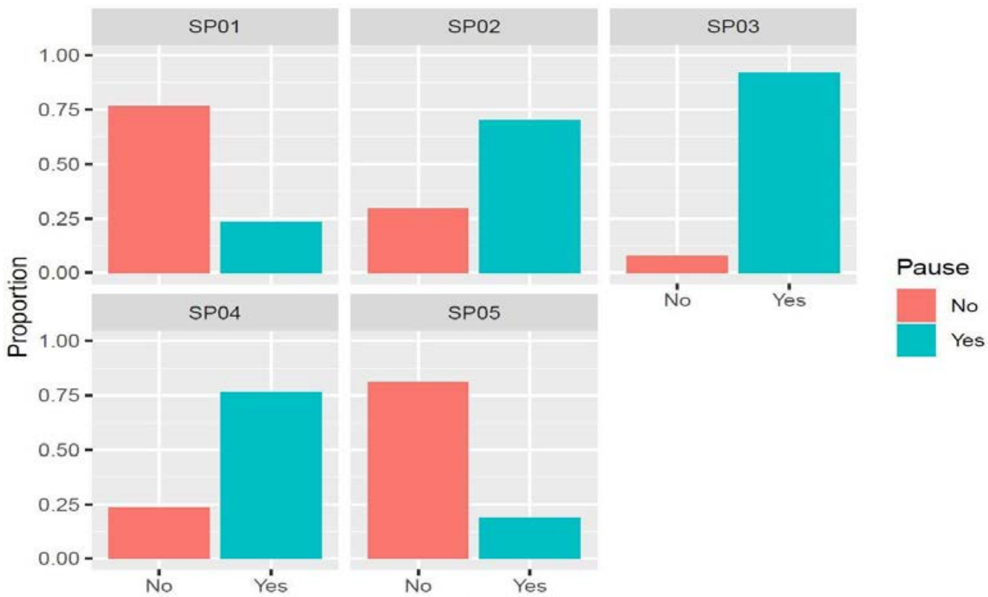

Figure 9: Speaker-specific proportions of utterances with and without pauses

4 Discussion

The hypothesis at the beginning of this experiment was that pitch reset and pause in complex declarative utterances in Akan occur after the clausal marker. The results do indeed corroborate the assumption that subordinate phrases make use of pitch reset as the downward movement of the pitch contour is always interrupted. However, with the conjunctions tested in this experiment, the pitch pattern of the conjunction moves to a higher level, indicating the start of a new intonational unit already *before* the conjunction. This result contrasts with the study conducted by Kügler (2017), where, within the complex declarative utterances, pitch reset occurred *after* the complementizer, that is, at the beginning of the following content word. However, Kügler's study tested complementizers and coordinating sentences while subordinate conjunctions were used in this study.

The current study supports Kügler's (2017) findings that the degree of pitch reset is dependent on the tonal structure. Findings from the present study indicate that pitch reset is more noticeable when the first tone after the reset is H and less so when it is a L tone.

Interestingly, the tonal behaviour of the subordinate conjunction remains the same regardless of the tonal structure of the word before and after the conjunction. Regardless of whether the content word that follows the conjunction has a L or H tone at the beginning, the final L tone of the conjunction is higher than the initial tone of the following word. If the final L tone is higher than the H tone of the initial syllable of the following word, it could be explained as a local case of automatic downstep (downdrift) at that specific juncture. Probably for declination to take place, the following L tone also needs to become lower than the last syllable L tone of the conjunction; nevertheless, it is not always straightforward to decide which phonetic or phonological effect is causing an observed lowering in some cases since the process could be complicated in tone languages.

The subordinate conjunction being prosodically phrased with the subordinate clause, with the pause consistently appearing before the conjunction, contrasts with Kügler's (2017) study. Kügler's findings revealed that the conjunction was phrased with the matrix clause, with the pause occurring after conjunction. The fact that, in the present study, the subordinate conjunction is consistently phrased with the subordinate clause signals the close connection between the conjunction and the subordinate clause. The reset and the pause mark the boundary between the two intonational phrases, the main clause and the subordinate clause, since pauses are frequently found at intonational phrase boundaries (Truckenbrodt 2005). Similarly, in Bemba (ISO 639-3 [bem]), where pitch reset occurs at the beginning of the main clause when a dependent clause was fronted, a pause follows the dependent clause, defining the boundary between the two intonational units (Kula and Hamann 2017). In Kɔnni (ISO 639-3 [kma]), a Gur language spoken in the Northern part of Ghana, speakers insert a pause before embedded clauses in complex utterances (Cahill 2016). However, a pause may be optional in complex sentences in Tswana (ISO 639-3 [tsn]), a southern Bantu tone language; the occurrence of a pause depends on the linear order of the clauses (Zerbian 2017). Although a pause may accompany a pitch reset within a complex sentence to define the boundary between the main clause and the dependent clause, the results suggest that pause realization in Akan is not a necessary feature of pitch reset and is optional, since two speakers rarely inserted a pause between the two units. In such cases, pitch reset is the only consistent cue for a dependent clause.

Additionally, the downtrend observed in the matrix and dependent clauses was caused by the expected local interactions between L and H tones. The final H and L tones of the intonational phrases moving to a lower level exemplify the Final Lowering rule pointed out by Paster

(2010). Paster maintained that, in Akan, there is a Final Lowering rule, which modifies H tones in the final position in an utterance to L. Kügler (2017) explained this lowering of the final tone in Akan utterances as tonal neutralization appearing at the end of an intonational phrase. Kügler stated that a sentence-final H tone neutralizes to L even when there is no L tone to cause downstep. Similarly, the f_0 contour of a sentence with only H tones may glide downwards throughout the sentence without a L tone to initiate the lowering, and the pitch pattern of sentences with only L tones may gradually move downwards to a final L tone (see Genzel 2013). Sentence-final lowering is also found, for example, in Mambila, Moro, Chimwiini, Tumbuka, and Embosi, although Tumbuka and Embosi have no downdrift, which initiates the downward movement of the f_0 contour (Downing and Rialland 2017).

Moreover, the difference between the pitch range of the dependent clause and the matrix clause seen in the utterances appears to demonstrate the dependency of the subordinate clause on the matrix clause. The difference in the pitch range could also explain how Akan speakers preplan longer utterances, adapting the pitch register according to the sentence's length to maintain declination throughout the utterance (see Genzel 2013). As posited by Truckenbrodt (2002), the pitch range difference suggests that pitch reset does not always reach the peak of the initial utterance, especially in a register lowering language like Akan, where a dependent clause may also be downstepped relative to the main clause. Also, compared to the more significant tonal difference observed within the syllables of the word that precedes the subordinate conjunction, the slight difference reflected in the word that follows the conjunction is probably because the following word is found towards the end of the complex sentence.

The effect of the final H syllable of the preceding noun on the initial L syllable of the following verb reflected in the data is accounted for by H tone spreading in Akan. Kügler (2017) asserted that H tone spreading is found across word boundaries, where the initial L tone of a following word is deleted, and the final H tone of a preceding word spreads onto the initial L syllable. On the other hand, Marfo (2004) identified H tone spreading as boundary assimilation and explained that boundary assimilation could be expressed through deletion or dislodging. With assimilation by deletion, the H tone deletes the L tone and causes no downstep in the following H tone, while with assimilation by dislodging, the L tone is dislodged and causes downstep in the following word-initial H tone. Marfo (2003) also mentioned that boundary assimilation is conditioned by factors such as speech rate. Whether a case of H tone spreading or boundary assimilation, the results demonstrate that a word-final H tone could influence a following word-initial L tone in Akan. The H tone spreading was particularly dominant in the utterances produced by one male speaker; this speaker's speech rate was especially high compared to the other speakers.

The interaction between lexical tones and intonation is a significant facet of linguistic systems since both phenomena implicate f_0 but perform distinct functions, and the occurrence of lexical tones together with structural intonation or intonational tones could be complex. Michaud and Vaissière (2015, 59) described intonational tones as when "in some tonal languages, tone can serve as a marker for functions at the phrasal level". For instance, tone languages may use tonal means to differentiate interrogative sentences from declarative ones (for Akan, see Genzel and Kügler 2020). Additionally, in some tone languages, certain junctures within an utterance could be defined by boundary tones. Boundary tones are found at the edges of intonational phrases, and phonetically, they are differentiated from final lexical tones by their scaling (Genzel 2013). For example, in Chichewa, boundary tones affect the f_0 realization of lexical tones, which causes the f_0 values of lexical tones in questions to be higher than in declarative statements (Myers 1996).

Hyman and Monaka (2011) identified three types of interactions between lexical tones and intonation in tone languages: accommodation, submission, and avoidance. In languages that exhibit accommodation, the system is divided so that lexical tones and intonation minimally interact, and they are both realized. In languages with submission, intonation overrides lexical tones, while intonation is minimized in languages which reflect avoidance. Hyman and Monaka explained that avoidance can be incomplete or complete in languages that exhibit avoidance. In incomplete avoidance, one or more lexical tones may override intonation, while in complete avoidance, the language does not allow any form of structural intonation.

Kügler (2017) explained the interaction between lexical tones and intonation in Akan as avoidance, as lexical tones are dissociated from intonational tones. Observations made from the present study signal that the interaction between lexical tones and pitch reset in Akan can be interpreted as avoidance, as disclosed by Kügler; it is clear that reset is larger when it is a H tone but smaller when it is a L tone, and lexical tones after pitch reset are also maintained.

In Akan, intonation is minimized, and lexical tones generally shape the f_0 contour of utterances. However, Kügler pointed out that the appearance of tonal neutralization at the end of declarative utterances could support accommodation, and this implies that Akan exhibits accommodation in addition to avoidance.

5 Conclusion

The present study has aimed to increase the knowledge of lexical tones as they occur alongside intonation in the Akan language. The study has validated the claim that Akan speakers employ pitch reset as a prosodic cue for dependent clauses. As speakers reach the lower end of their pitch register in the main clause within a complex utterance, they reset the f_0 contour within the subordinate conjunction to signal the direct relationship between the conjunction and the subordinate clause; simultaneously, speakers pre-plan complex utterances by pitching the initial clause higher to maintain the declination contour. Since lexical tones typically shape the f_0 contour of Akan utterances, lexical tones also determine the degree of f_0 reset in Akan, as reset is more noticeable when it is a H tone but smaller when it is L.

Future studies could concentrate on the tonal behaviour of other conjunctions in post-lexical intonation, as well as different types of matrix clauses. The small sample size, only five native speakers of Asante Twi, is a limitation of this study. While the study focused on speakers in the Helsinki region, the study did not consider how long the subjects have lived in Finland since a longer time living outside one's home country could cause first language attrition. Future studies could control for this factor, or more participants could be sampled from Kumasi, where the Asante Twi dialect is originally spoken. In the same vein, the number of participants does not allow us to establish systematic differences between gender-specific patterns.

Although the main dialects of Akan, including Twi and Fante, are mutually intelligible, and there are minor variations in their prosodic patterns and sound inventory, our results could be generalized across the dialects. Nevertheless, future investigations could focus on the other dialects to ascertain any dialectal differences.

Abbreviations

C conjunction; H high tone; L low tone; T tone; CONS consecutive; FOC focus; NEG negation

References

Abakah, Emmanuel N. 2005. “Tone Rules in Akan.” *Journal of West African Languages* 32 (1–2): 109–134.

Abakah, Emmanuel N. 2015. “On Tone and Morphophonology of the Akan Reduplication Construction.” *Journal of Universal Language* 16 (1): 1–47.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software* 67 (1): 1–48.

Boersma, Paul, and David J. M. Weenink. 2021. *Praat: Doing Phonetics by Computer* (version 6.1.39). www.praat.org.

Boadi, Lawrence Addai. 2010. *The Akan Noun Phrase: Its Structure and Meaning*. Black Mask.

Cahill, Michael. 2016. “Kɔnni Intonation.” In *Intonation in African Tone Languages*, edited by Laura J. Downing and Annie Rialland, 53–88. De Gruyter.

Cristofaro, Sonia. 2003. *Subordination*. Oxford University Press.

Dolphyne, Florence. 1986. “Tone and Grammar in Akan. The Tone of Possessive Constructions in the Asante Dialect.” In *The Phonological Representation of Suprasegmentals*, edited by Koen Bogers, Harry van der Hulst, and Marten Mous, 35–50. De Gruyter.

Dolphyne, Florence A. 1988. *The Akan (Twi-Fante) Language: Its Sound Systems and Tonal Structure*. Ghana Universities Press.

Dolphyne, Florence. 1994. “A Phonetic and Phonological Study of Downdrift and Downstep in Akan.” Paper presented at the 25th Annual Conference of African Linguistics, 27, Rutgers University.

van Dommelen, Wim A., and Dorothee Beermann. 2019. “Tonal Properties of the Akan Particle *Na*.” In *Proceedings of the 19th International Congress of Phonetic Sciences*, edited by Sasha Calhoun, Paola Escudero, Marija Tabain, and Paul Warren, 3438–3442. Australasian Speech Science and Technology Association, and International Phonetic Association.

Downing, Laura J. 2017. “Tone and Intonation in Chichewa and Tumbuka.” In *Intonation in African Tone Languages*, edited by Laura J. Downing and Annie Rialland, 365–392. De Gruyter.

Downing, Laura J., and Annie Rialland, eds. 2017. *Intonation in African Tone Languages*. De Gruyter.

Eberhard, David M., Gary F. Simons, and Charles D. Fenning, eds. 2021. *Ethnologue: Languages of the World*. 24th edition. SIL International.

Genzel, Susanne. 2013. “Lexical and Post-Lexical Tones in Akan.” PhD diss., University of Potsdam.

Genzel, Susanne, and Frank Kügler. 2011. “Phonetic Realization of Automatic (Downdrift) Non-Automatic Downstep in Akan.” *ICPhS* XVII: 735–738.

Genzel, Susanne, and Frank Kügler. 2020. Production and Perception of Question Prosody in Akan. *Journal of the IPA* 50 (1): 61–92.

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge University Press.

Hieber, Daniel William. 2016. "The Cohesive Function of Prosody in Ékegusií (Kisii) Narratives: A Functional-Typological Approach." MA thesis, University of California, Santa Barbara.

Hyman, Larry M., and Kemmonye C. Monaka. 2011. "Tonal and Non-Tonal Intonation in Shekgalagari." In *Prosodic Categories: Production, Perception and Comprehension*, edited by Sonia Frota, Gorka Elordieta, and Pilar Prieto, 267–289. Springer.

Ishihara, Shinichiro. 2016. "Japanese Downstep Revisited." *Natural Language & Linguistic Theory* 34 (4): 1389–1443.

Kawahara, Shigeto, and Takahito Shinya. 2008. "The Intonation of Gapping and Coordination in Japanese: Evidence for Intonational Phrase and Utterance." *Phonetica* 65 (1–2): 62–105.

Kügler, Frank. 2016. "Embedded Clauses and Recursive Prosodic Phrasing in Akan." *5th International Symposium on Tonal Aspects of Languages (TAL 2016)*, 38–42.

Kügler, Frank. 2017. "Tone and Intonation in Akan." In *Intonation in African Tone Languages*, edited by Laura J. Downing and Annie Rialland, 89–130. De Gruyter.

Kügler, Frank. 2024. "Topicalization and Prosodic Phrasing in Akan." In *To the Left, to the Right, and Much in Between: A Festschrift for Katharina Hartmann*, edited by Anke Himmelreich, Daniel Hole, and Johannes Mursell, 241–257. Goethe University, Frankfurt.

Kula, Nancy C., and Silke Hamann. 2017. "Intonation in Bemba." In *Intonation in African Tone Languages*, edited by Laura J. Downing and Annie Rialland, 321–364. De Gruyter.

Manyah, Kofi Adu. 2014. "Relationship between Lexical Tone Contrasts and Vowel Quality." *European Scientific Journal* 10 (17): 418–428.

Marfo, Charles O. 2003. "The Prosody-Syntax Interface in Akan Focus and Topic Constructions." In *Proceedings of the 4th Postgraduate Research Forum on Linguistics*, 125–133. University of Hong Kong.

Marfo, Charles O. 2004. "On Tone and Segmental Processes in Akan Phrasal Verbs: A Prosodic Account." *Linguistic Online* 18 (1): 93–110.

Michaud, Alexis, and Jacqueline Vaissière. 2015. "Tone and Intonation: Introductory Notes and Practical Recommendations." In "Theoretical and Empirical Foundations of Experimental Phonetics," special issue, *KALIPHO – Kieler Arbeiten zur Linguistik und Phonetik* 3: 43–80.

Myers, Scott. 1996. "Boundary Tones and the Phonetic Implementation of Tone in Chichewa." *Studies in African Linguistics* 25 (1): 29–60.

Paster, Mary. 2010. "The Verbal Morphology and Phonology of Asante Twi." *Studies in African Linguistics* 39 (1): 77–120.

Peng, Shu-Hui, Marjorie Chan, Chiu-Yu Tseng, Tsang Huang, Ok Joo Lee, and Mary E. Beckman. 2005. "Towards a Pan-Mandarin System for Prosodic Transcription." In *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by Sun-Ah Jun, 230–270. Oxford University Press.

Rialland, Annie. 2009. "The African Lax Question Prosody: Its Realisation and Geographical Distribution." *Lingua* 119 (6): 928–949.

Saah, Kofi. 1994. "Studies in Akan Syntax, Acquisition, and Sentence Processing." PhD diss., University of Ottawa.

Simpson, Adrian P. 2009. "Phonetic Differences Between Male and Female Speech." *Language and Linguistics Compass* 3: 621–640.

Snider, Keith. 2007. "Automatic and Nonautomatic Downstep in Chumburung: An Instrumental Comparison." *Journal of West African Languages* 34 (1): 105–114.

Stewart, John Massie. 1965. "The Typology of the Twi Tone System." *Bulletin of the Institute of African Studies* 1: 1–27.

Swerts, Marc, Eva Strangert, and Mattias Heldner. 1996. "f_0 Declination in Spontaneous and Read-Aloud Speech." *Proceedings of Fourth International Conference on Spoken Language Processing* (*ICSLP '96*), 1501–1504.

Truckenbrodt, Hubert. 2002. "Upstep and Embedded Register Levels." *Phonology* 19 (01): 77–120.

Truckenbrodt, Hubert. 2005. "A short report on intonation phrase boundaries in German." *Linguistische Berichte*, 203: 273–296.

Xu, Yi. 2013. "ProsodyPro: A Tool for Large-Scale Systematic Prosody Analysis." In *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (*TRASP 2013*), 7–10.

Zerbian, Sabine. 2017. "Sentence Intonation in Tswana (Sotho-Tswana Group)." In *Intonation in African Tone Languages*, edited by Laura J. Downing and Annie Rialland, 393–434. De Gruyter.