Lexical and post-lexical tones in Akan

Dissertation
zur Erlangung des akademischen Grades
Doctor philosophiae
(Dr.phil.)
an der Humanwissenschaftlichen Fakultät
der Universität Potsdam

eingereicht von
Susanne Genzel

Februar 2013

In memory of my aunt, Sabine Genzel.

Acknowledgements

First of all, I would like to express my very great appreciation to my supervisors, Caroline Féry and Frank Kügler. I am especially grateful to Frank. He hired me as a student assistant in summer 2003 and let me take part in many interesting empirical investigations. Later on, in November 2007, he employed me in the project D5 on "Tonal structure and information structure: Phonetics & Phonology", which is part of the wonderful collaborative research centre on information structure (SFB 632), financed by the Deutsche Forschungsgemeinschaft (DFG). Akan is one of the languages we are investigating in our project. Frank did a great job as supervisor. He was very interested and encouraging and always had time for me to discuss new ideas. Caroline was my Prof. in linguistics when I studied at University of Potsdam. Her seminar "Intonation im typologischen Vergleich", held in winter 2004/2005, impressed me and made me curious for intonation/prosody and typology. Caroline's comments were very helpful and improved the thesis immensely. Discussing with her was really inspiring.

My special thanks are extended to all participants and informants. "*Mo boaa me paa. Medaase!*" Especially, I would like to mention Afua Blay, Reginald Duah, Benjamin Bamfo and Monica Apenten, who made me feel at home in Ghana and were always willing to share their native speaker intuitions with me. Without you, I could not have written the thesis. Further, I would like to thank Akua Appiah-Akuramaa who was my 'Berlin' informant. Further, I am very grateful to Dr. K. Dovlo who very kindly welcomed me to the University of Ghana (Legon), arranged my first talk in front of a native speaker audience and provided me with everything I needed. I also wish to acknowledge the help provided by Prof. Dr. K. Saah and Nana Ama Agyeman. Additionally, I would like to thank Dr. C. Marfo, working at the University of Kumasi (KNUST), for discussing phrasing related and other Akan specific issues with me and for drawing my attention to the conference of the Linguistics Association of Ghana in Winneba (2012).

I have talked to many people during the thesis writing process. First of all, I would like to mention Sabine Zerbian. Since she had her office very close to mine, I went there very often to ask tone language related questions. She always had time for me and also read an early version and provided very helpful comments. Further, I am grateful to Ruben van de Vijver who also gave very useful advices and made me laugh. I would also like to mention Anja Arnhold, Sandra Beyermann, Laura Downing, Ines Fiedler, Anja Gollrad, Markus Greif, Mira Grubic, Robin Hörnig, Shinishiro Ishihara, Sara Myrberg, Annie Rialland, Bernadett Smolibocki and Thomas Westkott….

Finally, I wish to thank my family, Mama, Oma, Onkelchen, Marie und Hannah for their unconditional love and support. Further, I am grateful to all my friends for their patience, for calling me, writing me, cooking for me, giving massages and providing plenty of other sorts of support. Last but not least, I want to thank Himanshu, for keeping the household running, for his irrepressible high spirits, patience, support and goodies.

List of abbreviations and symbols

ATR	advanced tongue root
BT	breathy termination
C	consonant
COMP[1]	complementizer
CP	complementizer phrase
db	decibel
DD	downdrift
DEF	definite
Δ	difference/drop
DIM	diminutive
DET	determiner
DP	determiner phrase
DS	downstep
F	focus
F	formant
F0	fundamental frequency
fin	final
FM	focus marker
h	high register tone
H	high tone
HAB	habitual
Hz	Hertz
IMP	imperative
InflP	inflectional phrase
Int	intensity
Int_{rel}	relative intensity
IP/ι	intonation phrase
l	low register tone
L	low tone
LPC	linear predictive coding
M	mid tone
max	maximal
ms	millisecond
μ	mora
NEG	negative
NOM	nominal
NP	nominal phrase
NSF	non-subject focus
OPT	optative
PL	plural
*p*MCMC	p-value calculated from MONTE CARLO sampling by Markov chain
pP/φ	phonological phrase
PART	particle
Φ	foot
PoD	position of the pitch drop
PRF	perfective
PRO	pronoun

[1] The abbreviations in small capitals basically follow the convention of the Leipzig Glossing Rules (Bickel, Comrie & Haspelmath, 2004).

PROG	progressive
P_{sg}	subglottal air pressure
PST	past
pw/ω	prosodic word
r	baseline value
R^2	coefficient of determination
RTR	retracted tongue root
σ	syllable
s	downdrift quotient
sec	second
SF	subject focus
SG	singular
st	semitone
SPEC	specific
SVO	subject, verb, object
T	tone
T	terminal
T*	pitch accent
T!	downstepped tone
T̥	floating tone
T%	boundary tone
TBU	tone bearing unit
TM	topic marker
TW	target word
U	utterance
V	verb
V	vowel
VP	verbal phrase
x	beat
X^0	lexical word
XP	maximal projection
♀	female
♂	male
↘	local lowering
↗	local raising
↑	global raising

Contents

Acknowledgements……………………………………………………………………...i
List of abbreviations and symbols………………………………………………………… iii

About the book 1

1. Chapter General and theoretical background 7
   1.1 General introduction to Akan 8
   1.2 Phonological representation of tone 12
   1.3 Tone terracing 17
   1.4 Prosodic structure 25
      1.4.1 The prosodic organization of constituents 25
      1.4.2 Prosodic constituents and their phonetic properties 26
   1.5 Tonal and non-tonal intonation 40
      1.5.1 What is intonation? 40
      1.5.2 Sentence type 42
      1.5.3 Focus 45
   1.6 Universal intonational meaning 51
      1.6.1 The Production code 51
      1.6.2 The Frequency code 53
      1.6.3 The Effort code 54
   1.7 Phonetic implementation 55
   1.8 Tonal variations in connected speech 58
   1.9 Summary of the main research questions 60

2. Chapter Material and methods 62
   2.1 The participants 62
   2.2 The corpus 63
   2.3 Recording procedure 70
   2.4 Data pre-processing and statistical analysis 71

3. Chapter The influence of adjacent tones in connected speech 73
   3.1 L raising 73
   3.2 H raising 81

4. Chapter Downtrends 92
   4.1 Types of downstep 93
   4.2 Downstep in initial position 100
   4.3 Declination and downstep 107
      4.3.1 Declination 107
      4.3.2 The implementation of declination 119
      4.3.3 Downstep 124
      4.3.4 The implementation of downstep 143

5. Chapter Anticipatory raising, length and prosodic structure 152
   5.1 Anticipatory raising and downstep 153
   5.2 Anticipatory raising and phrase length 155

6. Chapter The intonation of Yes – No questions 173

7. Chapter The expression of focus 201
7.1 Syntactic focus marking 201
7.2 Prosodic focus marking 209

Conclusion 223

Bibliography 235

List of figures 251
List of tables 258

Statement of Authorship/Selbstständigkeitserklärung 260

About the book

The melodic composition of a sentence is commonly known as intonation. Intonation functions to signal post-lexical pragmatic meanings, such as sentence type and focus (Ladd, 1996). The languages of the world differ with regard to what extent the choice of melodic morphemes is entailed by the choice of words (Pierrehumbert, 2001). In intonation languages, like English and German, the melody of a sentence is assigned post-lexically and serves to inform the listener "…about how the utterance is related to the discourse and to mutual beliefs which interlocutors build up during the course of the discourse." (Pierrehumbert, 2001:11). In tone languages, on the other hand, the melody of a sentence is largely determined by the choice of words and by the choice of grammatical categories, such as tense, aspect or case. The thesis investigates the tone language Akan.

Akan has been first described by Christaller in his 1875 book "The grammar of the Asante and Fante language". He recognized that "Every syllable has, in comparison with neighbouring syllables either *low* or *high* or *middle* tone." (Christaller, 1875:15). Tone in Akan serves lexical and grammatical function (Dolphyne, 1988); a detailed introduction is provided in chapter 1 section 1.1. Akan is an African tone language spoken in Ghana. It has two level tones (low and high), which are part of the lexical entry of (some) morphemes (Hyman, 2001:1367 inspired by Welmers, 1973). Furthermore, it has been described as terraced level tone language (Welmers, 1959; Clements, 1979). Christaller (1875:15) recognized the existence of middle tones and describes it as "…high tones abating by one step or successive steps….". This stepwise lowering of high tones in a sequence of alternating low and high tones has been called downstep. Downstep leads to a terracing surface melody; a detailed introduction is presented in chapter 1 section 1.3. Since its discovery, the nature of downstep has been a thrilling research topic.

The book is about factors that contribute to the surface forms of tones in connected speech. First and foremost, it aims at gaining insights into the nature of downstep in Akan. Controlled experimental data on the issue is presented in chapter 4 section 4.3. It will be argued that the terracing surface pattern in Akan is caused by declination. The term declination describes a gradual pitch lowering phenomenon over the course of an utterance (Cohen & ´t Hart, 1967). Although level tone languages, like Akan, provide the opportunity to study declination directly in sequences containing only high or only low tones, its presence has, until now, not been subject to a controlled empirical investigation. It will be argued that declination is an intonational property of Akan, which serves to signal coherence. A phonological representation using a high (h) and a low (l) register tone, associating to the left and right edge of an intonational phrase (IP) respectively, inspired by Möhler & Mayer (2001), will be proposed. Declination and downstep will be modeled using a (phonetic) pitch implementation algorithm, originally proposed for modeling downstep in English by Liberman & Pierrehumbert (1984) and used by Shih (2003) for modeling declination in Mandarin Chinese. Departing from Liberman & Pierrehumbert's (1984) proposal, an innovative application of the algorithm will be presented which naturally captures the relation between declination and downstep in Akan. Two new procedures will be proposed. First, the pitch implementation algorithm for Akan can differentiate between high and low, if it detects two different tonal entities in sentence, it starts to calculate values for both of them and, second, articulatory goals only supplied to the compatible entity. Stewart (1965) made the significant observation that the lowering of the second high tone in a high-high tone sequence paralleled the lowering of the second high tone in a high-low-high sequence. The former lowering process has been observed to be triggered by an underlying low tone which was not present at the surface, also known as floating tone. He coined the key term automatic downstep to refer to the latter type and non-automatic downstep to refer to the former type. Following Stewart's footsteps a number of researchers have been concerned with the empirical validation of his claim (e.g. Laniran, 1992 and Liberman, Schultz, Hong & Okeke, 1992 for Igbo; Snider, 1998 for Bimoba and Snider, 2007 for Chumburung), largely confirming the phonetic equality of the two types of downstep. For Akan, however, Dolphyne, (1994) has claimed that non-automatic downstep causes a greater degree of lowering than automatic downstep. Controlled experimental data will be presented in chapter 4 section 4.1 (published as Genzel & Kügler, 2011). It will be shown that the two types of downstep are phonetically similar in Akan. On the basis of the result, it will be argued that non-automatic is due to the presence of a non-associated low tone in the tonal string (Stewart, 1965; Abakah, 2000).

Laniran and Clements (2003) report in their seminal paper "Downstep and high raising: interacting factors in Yoruba tone Production" that high tones which are followed by low tones appear raised on the surface. The effect is described in detail in chapter 1 section 1.8. High raising has been observed in many tone languages (e.g. Yu, 2009 for Bole; Gandour, Potisuk & Dechongkit, 1994 for Thai; Xu, 1997 and Wang & Xu, 2011 for Mandarin Chinese). So far, high raising has not been mentioned in the literature on Akan. The thesis provides experimental data, presented in chapter 3 section 3.2, confirming its existence. It will be argued that high raising is a local anticipatory planning effect, employed at the phonetic level, which enhances the perceptual distance between low and high tones (Gandour, Ponglorpisit, Dechongkit, Khunadorn, Boongird & Potisuk, 1993; Potisuk, Gandour & Harper, 1997; Chen, 2012). Furthermore, in chapter 3 section 3.1, data will be presented, showing that low tones are raised if they are wedged between two high tones. L raising will be argued to be a local carry-over effect (co-articulation), following Gandour et al. (1994) and Laniran & Clements (2003).

Christaller (1875:183) made the observation that "… in the *beginning* of a longer sentence or complex sentences, …tones, … are *higher throughout*….", indicating a certain amount of planning at the phonetic level. The scaling of initial tones in relation to sentence/phrase length has not been experimentally studied in Akan. Chapter 5 offers controlled experimental data on global anticipatory raising in simple and complex sentences. It will be shown that Akan speakers anticipate the length of an IP and not the length of the first constituent as in German (Petrone, Fuchs & Krivokapić, 2011) or Whenzou Chinese (Scholz, 2012). Preplanning (anticipatory raising) will be argued to be an important process at the level of pitch implementation. It serves to ensure that declination can be maintained throughout the IP, which prevents pitch resetting.

Apart from tonal variations at the phonetic level, the book is concerned with intonation; a detailed introduction is provided in chapter 1 section 1.5. Following Gussenhoven (2004:58), the phonological surface representation will be constructed "…from the combination of lexical and postlexical tones.". Post-lexical tones are understood as intonational tones, which express discoursal/pragmatic meaning at the phrasal level (Ladd, 1996). In the phonological representation, intonational tones are represented as boundary tones (Pierrehumbert, 1980; Pierrehumbert & Beckman, 1988), which associate with the edges of prosodic constituents (Nespor & Vogel, 1986), like IPs. Two types of discoursal meaning, sentence type and information structure, commonly discussed under the heading of intonation, will be experimentally investigated for Akan. About the expression of the former type of meaning Christaller (1875:97) notes that "The interrogative tone of the sentence also may suffice, …" to distinguish a question without syntactic or morphological marking from a string-identical statement and continues describing the acoustic manifestation of the interrogative tone as "… the last sound of the sentence is lengthened and sinks into the low tone.". Following him, many linguistics described the intonational difference between Yes – No questions and statements in Akan (e.g. Schachter & Fromkin, 1968; Barry & Aidoo, 1975; Dolphyne, 1988; Boadi, 1990; Abakah & Koranteng, 2007). Furthermore, it has been observed that final low tones in Yes – No questions lack final lowering, whereas final high tone exhibit a steep F0 fall (Dolphyne, 1988), that both low and high tones are scaled higher (higher register) in Yes – No questions (Schachter & Fromkin, 1968; Berry & Aidoo, 1975; Dolpyhne, 1988), that automatic downstep is suspended or reduced in Yes – No questions (Hyman, 2001; Gussenhoven, 2004), that the sentence final element in Yes – No questions exhibits extra voicing or glottalization (Boadi, 1990) and that the final vowel is lengthened in Yes – No questions (Christaller, 1875). The thesis provides controlled experimental data on the issue as presented in chapter 6. The data largely confirms the observations above, except the suspension or reduction of automatic downstep. The findings are discussed in the light of the question typology proposed by Rialland (2007, 2009); the discussion unfolds in an extension of Rialland's typology in terms of a new category called "low tense" question prosody. It will be argued that only the low tone at the right edge of a Yes – No question has to be represented phonologically and that all other effects result during the phonetic implementation of the post-lexical tone. A phonological representation of the post-lexical low boundary tone is proposed using autosegmental-metrical theory (Pierrehumbert, 1980). The higher register is claimed to not be part of the phonological representation in Akan. It will be proposed that it emerges at the phonetic level to compensate for the 'unnatural' form of the question morpheme and to satisfy the frequency code (Gussenhoven, 2002; 2004); a detailed introduction is provided in chapter 1 section 1.6. Further, it will be argued that statements are not marked tonally in Akan. An analysis involving a tonally not specified post-lexical boundary tone, (0%), following Grabe (1998), is proposed.

Apart from differentiating questions from statements, speakers may want to express which element(s) in the discourse are of greater importance than others. This is commonly referred to as information structure, which is a cover term for notions such as focus, topic and givenness (Krifka, 2007). Pragmatic uses of focus, which will be of concern here, can be informational, as assumed for elements in answers replacing the wh-element of a preceding question, or corrective, as assumed for elements which replace/correct another element in the preceding context (see Krifka, 2007 for an overview). The latter use is assumed to be more emphatic (e.g. Hartmann, 2007, 2008) or stronger (Féry, 2012). Christaller (1875:146) notes that: "Any part of a sentence may be made emphatic by placing it before the sentence…. The word which is this rendered prominent, is followed … by the conjunction na…". Following him, many researchers in the field of syntax examined this ex-situ 'focus' construction (e.g. Boadi, 1974; Saah, 1988; Drubig, 2001; Marfo & Bodomo, 2005; Kobele & Torrence, 2006; Ermisch, 2006) confirming Christaller's (1875) observation. An element in informational focus, on the other hand, is not replaced i.e. can remain in-situ (Saah, 1988; Ermisch, 2006). The thesis provides semi-spontaneous data (Genzel & Kügler, 2010), presented in chapter 7 section 7.1, on the frequency of the use of the "focus" construction showing that in-situ occurrences are possible with corrective focus. It will be argued that the ex-situ construction does not signal a focus type (e.g. Drubig, 2003) but that its occurrence is related to extra grammatical factors, such as hearer expectation, discourse expectability (Zimmermann, 2007) and emphasis (Hartmann, 2008). Given the fact that in-situ focus is more frequent, it is striking that the prosodic/intonational properties of focus in Akan have received only little attention. The only contribution about tone and in-situ focus is concerned with "emphatic" focus on subjects in Fante, a dialect of Akan. Abakah & Koranteng (2007) report on a tonal change on the verb. If the verb follows a focused subject, it receives "…an extra H and an extra L…" (Abakah & Koranteng, 2007:76). In a recent paper, Kügler & Genzel (2012) investigated, amongst other things, the acoustic characteristics of high and low toned focused in-situ objects in Akan. They show that the verb is not subject to tonal changes, that informational focus is not accompanied by a categorical and/or gradient tonal change and/or durational difference and that corrective focus leads to a lowering of both high and low tones on the object. Kügler & Genzel (2012) concluded that there is no intonational tone marking focus in Akan and that: "In the case of corrective focus, …, an additional pragmatic prominence comes into play which speakers may want to express even prosodically. This additional prominence is correlated with a stronger communicative goal to emphasize a certain part of an utterance, and speakers of Akan draw attention to that kind of information by means of pitch register lowering." (Kügler & Genzel, 2012:353). The thesis provides a re-analysis of the data, presented in chapter 7 section 7.2, showing that objects carrying a high tone are raised under corrective focus. The results are interpreted in the light of the Effort code (Gussenhoven, 2002, 2004). Furthermore, it will be shown with the help of semi-spontaneous data, presented in chapter 7 section 7.2, that focus on objects in Akan is marked by insertion/enhancement of a prosodic boundary to the right of the focus. The phonetic correlates of the boundary are the occurance of pause and glottal stop. The finding supports Féry's (2012) claim that focus prefers to be aligned prosodically.

An essential part of the phonological representation is the prosodic phrase structure. A detailed introduction is provided in chapter 1 section 1.4. The prosodic structure largely mirrors the syntactic structure (e.g. Selkirk, 1978, 1981, 1984; Nespor & Vogel, 1986; Beckman & Pierrehumbert, 1986, Pierrehumbert & Beckman, 1988). The boundaries of prosodic phrases are commonly marked by pre-boundary lengthening and/or ocurrance of pauses (e.g. Vaissière, 1983). Pauses as indicators of prosodic phrases will be utilized in the thesis. Furthermore, prosodic constituents serve as domains for the application of language-specific phonological processes e.g. high tone spread (e.g. Zerbian, 2006). Akan is an SVO language (Boadi, 1974; Saah, 1988; Marfo & Bodomo, 2005; Kobele & Torrence, 2006). Like many African tone languages, Akan exhibits vowel harmony (e.g. Stewart, 1967; Schachter & Fromkin, 1968; Clements, 1981; Obeng, 2000; O'Keefe, 2003). Vowel harmony has been found to apply between verb and nominal prefix of the object (e.g. Dolphyne, 1988; Ballard, 2010) but crucially not between subject and verb (Kügler, 2012); indicating that the subject forms its own phonological phrase, whereas, verb and object are phrased together. Marfo (2003, 2004, 2005), principally, assumes the same prosodic phrase structure on the basis of a 'boundary assimilation' process, which is a tone spreading process. The thesis offers a critical evaluation of Marfo's 'boundary assimilation' account, presented in chapter 1 section 1.4.2. Controlled experimental data, presented in chapter 5 section 5.2, provides insights into the phrasing at the level of the IP. It will be argued that ex-situ 'focus' constructions and embedded complementizer clauses are best represented by a recursive intonation phrase structure, following Selkirk (2009).

The melody of an Akan sentence is largely determined by the choice of words. The lexical entry of words contains a specification for tone (low and/or high tone). When words with lexical tones form a sentence, the overall melodic shape is falling. Besides post-lexical register tones, which trigger declination, post-lexical tones that signal sentence type have global and local effects in Akan. A statement, which is marked by a tonally not specified boundary tone, exhibits final lowering on the final element. A question, which is marked by a low boundary tone, exhibits a sharply falling pitch movement on a final high toned vowel. Furthermore, at the phonetic level the pitch register is raised, which causes low and high tones to be realized higher in questions than in statements. If a speaker of Akan wants to highlight a particular element (focus), the default prosodic structure is modified in such a way that the focused element forms its own phonological phrase. This restructuring is accompanied by an interruption of the otherwise continous melody (pause) and by a glottal stop. Furthermore, at the phonetic level, local and global anticipatory raising processes play a role in determining the surface characteristics of the acoustic signal. At the local level, a sentence initial high tone is realized higher if it is followed by a low tone than if it is followed by another high tone. At the global level, initial low and high tones are realized higher if they occur in a long and/or complex sentence. Hence, many factors that emerge at different levels of the tone production process contribute to the surface form of the acoustic signal in Akan.

1. Chapter
General and theoretical background

Generally, the thesis is concerned with processes attributed to the phonological and phonetic components of the grammar. The figure 1 gives an overview of the division of labour between the two (Gussenhoven, 2004:58).

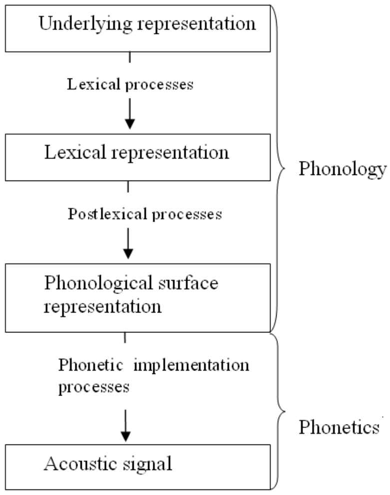


figure 1: Model of Lexical Phonology; reproduced from Gussenhoven (2004:58) with slight modifications (replacement of rules with processes).

Section 1.1 provides a general introduction to Akan with emphasis on lexical tone. Lexical tones are assumed to be part of the underlying representation of morphemes (Pulleyblank, 1986). The phonological representation of tones is the subject of section 1.2. Morphological operations influencing the segmental and tonal structure are introduced and the resulting lexical representations (Pulleyblank, 1986) of selected examples are illustrated. The lexical representations are assumed to be the units that are incorporated into the syntactic structure (Gussenhoven, 2004). At this point, we leave the field of lexical phonology and enter into the area of post-lexical phonology. Section 1.4 will be considered with the mapping of the syntactic structure onto the prosodic structure and with phonetic reflexes accompanying prosodic boundaries generally and Akan-specific. One task of post-lexical phonology is to insert intonational morphemes. Section 1.5 gives an introduction to intonation and its representation. The tone terracing pattern (downstep) of Akan, which has been claimed to be an intonational property (e.g. Hombert, 1974), is topic of section 1.3. The output of the post-lexical phonology is the phonological surface representation with lexical and post-lexical tones and hierarchically organized prosodic constituents (Gussenhoven, 2004). The phonological surface representation is the end-product of the phonology and serves as input to the phonetic component. Processes applying in the phonetics, affecting the pitch range/register and conveying attitudinal and informational meanings are introduced in section 1.6. Background on how these abstract representations are mapped onto concrete articulatory/acoustic goals is provided in section 1.7. Finally, section 1.8 introduces processes that may occur during articulation and affect the actual height of tones in the acoustic signal (F0).

1.1 General introduction to Akan

Akan is a cover term subsuming the dialects Asante, Akyem, Akuapem, Fante, Wasa, Agona, Brong, Kwahu, and Gomua[2] (Dolphyne & Kropp Dakubu, 1988:52). It is a Kwa language belonging to the Niger-Congo family spoken in the central and southern regions of Ghana by about 8.3 million speakers (Lewis, 2009). The data used in this work concentrates on the Asante dialect, which is spoken in the Ashanti province around the cultural centre of Kumasi. I will still use the term Akan throughout.

The basic word order is SVO (e.g. Boadi, 1974; Saah, 1988) with head-initial properties; nouns precede determiners, adjectives, and numerals and adverbial modifiers follow adjectives etc. (Kobele & Torrence, 2006). This is as illustrated in (1).

(1)	Kontromfi	no	tua	dua	kakraa	futufutu.
	monkey	DET	posess.HAB	tail	huge	fluffy
	The monkey has a huge fluffy tail.					

Akan exhibits advanced tongue root (ATR) vowel harmony (e.g. Stewart, 1967; Schachter & Fromkin, 1968; Clements, 1981; Obeng, 2000; O'Keefe, 2003), like many African languages (Hyman, 2003). Further details and examples will be provided in section 1.4.2.

Akan is a tone language with two tones, high (H) and low (L) (Stewart, 1965; Dolphyne, 1988). Tones will be marked acute accent (ˊ) and grave accent (ˋ) respectively, following the Africanist tradition and conventions of the International Phonetic Alphabet. The term tone is an abstraction used to refer phonological distinct categories. An H tone exhibits a relatively higher pitch than an L tone. Pitch is the acoustic correlate of tonal height which in turn is generated through the variation of the speed of the vocal chord vibration, "…the higher the frequency of vibration of the vocal chords, the higher will be the number of periods per second, commonly known as the fundamental frequency, …of the acoustic signal, and higher the resulting pitch." (Gussenhoven, 2004:2). The fundamental frequency, abbreviated F0, is measured in periods per second (Hz) and is the primary variable employed in the thesis. Tone

[2] The dialects differ at the segmental and the tonal level; for a detailed discussion about the differences see Cahill (1985); Dolphyne (1988); Abakah (2005).

in Akan has lexical and grammatical function (e.g. Dolphyne, 1988). It thus contributes to the meaning of words, as illustrated in (2). The lexical meaning of the disyllabic word *papa* changes according to its tonal specification. In (2)a. it carries two lexical H tones, two lexical L tones in (2)b., and an L tone on the first and an H tone on the second syllable in (2)c.

(2) a. pápá – good
b. pàpà – fan
c. pàpá – father
(Dolphyne, 1988:52)

It has been a matter of debate whether there is a third lexical tonal entity in Akan. Dolphyne (1988, 1994) claims that downstepped H tones are part of the lexical entry of words; henceforth marked with an exclamation mark (!). She presents the examples reproduced in (3) amongst others. The crucial comparison is the one between (3)b. and (3)c. The lexical distinction in this case relies on the relatively lower pitch of the third H tone in (3)c. It should be noted that lowered H tones are still higher than L tones (e.g. Christaller, 1875).

(3) a. ɔ̀bɔ̀fó – hunter
b. ɔ̀bɔ́fó – creator
c. ɔ̀bɔ́!fó – messenger
(Dolphyne, 1988:52,55)

Abakah (2000:273), on the contrary, provides evidence that the occurrence of !H "…at least in Akan, is predictable and therefore redundantly automatic.". He reanalyzes Dolphyne's example, see (3)c., and proposes the tonal representation displayed in (4). L̥ refers to floating tone, i.e. a tone which is not associated with a segment; see section 1.2 for further details.

(4) ɔ̀bɔ́ø̀fó (LHL̥H)
(Abakah, 2000:263)

He (2000:264) motivates his assumption with the observation that instances like in (3)c. are etymologically decomposable into primitive elements as illustrated in (5) i.e. these instances are compounds.

(5) ɔ-bɔ́	+	ɔ-fó
NOM-to be in an excessive motion		NOM person

(Abakah, 2000:264)

I left the nominal prefixes (NOM) underlyingly unspecified for tone, following Abakah's (2002, 2005) proposal that "…most Akan nouns, deverbal or non-deverbal, have a nominal prefix and an optional nominal suffix both of which are toneless." (Abakah, 2005:116). Furthermore, Abakah (2002:123) claims that pronominal clitics and tense/aspect affixes are toneless, too. Toneless elements either receive a default L tone during the derivation, in the case of nominal affixes or acquire tone via tone polarization, as in the case of pronouns; see Abakah (2005:123) for details and Arkoh (2011:56ff.) for further influencing factors in Fante. The situation is, however, presumably more complex, as we shall see in section 1.4.2. Under Abakah's analysis, both nominal prefixes in (5) receive an L tone during the derivation. The segmental content of second nominal prefix is deleted due to hiatus resolution when both elements merge. The L tone associated with the second nominal prefix remains in the tonal string as floating tone and causes lowering on the following H tone; see section 1.3 for further details. Unfortunatly, Abakah (2002) does not further explain which elements are involved in the generation of Dolphyne's examples (3)a. and (3)b. In chapter 4 section 4.3.1, I will address the following question:

- Is !H an independent entity in the phonology of Akan?

Beside its lexical function, tone in Akan is used to mark grammatical categories such as tense/aspect/mood. This is illustrated in (6). The verb *bisa* in the habitual form, which is assumed to reflect the underlying form of the verb root (Paster, 2010), surfaces with an L tone on the first syllable and an H tone on the second syllable as illustrated in (6)a. In (6)b. the verb exhibits only L tones. The imperative mood is marked by a grammatical L tone, which replaces the underlying lexical tones. The verb in perfective aspect is realized with only H tones, as shown in (6)c. The perfective aspect comes along with a floating H tone (H̥) and the toneless prefix /a/; see section 1.2 for further details.

(6) a. bìsá
ask.HAB
b. bìsà
ask.IMP
c. a-bísá
PRF-ask

(Paster, 2010:107)

Turning to the pitch realization of tones, Pike (1948) differentiates between level (register) and contour tone systems. Level tone systems are typically found in Africa, tone is articulated with level pitch (Pike, 1948; Hyman, 2001; Yip, 2002). Contour tone systems are commonly found in Asia and exhibit pitch movements (falling/rising pitch) on a syllable (Hyman, 2001; Yip, 2002). Akan is a level tone language (e.g. Clements, 1979; Dolphyne, 1988; Abakah, 2000). Pike (1948:5) characterizes a level tone language as "…one in which, within the limits of perception, the pitch of a syllable does not rise or fall during its production.". However, small changes in pitch may occur (Maddieson, 1978; Yip, 2002). The figure 2 illustrates L and H tone levels with data from a female Akan speaker. The disyllabic proper names *Yàẁ* and *Ésí* are displayed.

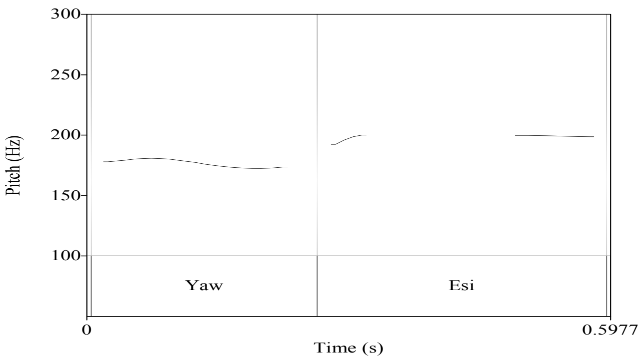


figure 2: Left side: L toned proper name (*Yaw*) taken from the initial position of the sentence *Yaw fi Akyemfo.* 'Yaw comes from Akyemfo.'; Right side: H toned proper name (*Esi*) taken from the initial position of the sentence *Esi yi lɛta*. 'Esi takes a letter away.'; female speaker.

The following section 1.2 provides background on the theoretical framework used in the thesis to represent tones in the phonology.

1.2 Phonological representation of tone

Goldsmith's (1976) proposal that tones have an auto-segmental status will be used to represent tones in the phonolgy. According to his proposal, segments and tones are represented on separate tiers and are linked by association lines (|). This is illustrated in (7) for a disyllabic word with L and H tone. Goldsmith (1979:205) states that these lines indicate the "co-registration" of the different segments in the tiers at the phonetic level. A tone is realized on the surface if it is associated with a tone bearing element (TBU).

```
(7)    L  H          tonal tier
       |  |
      CVCV           segmental tier
```

As Yip (2002:72) points out, in tone languages like Akan "…the associations must be underlying, because they are lexically distinctive.". Associations between tones and segments are subject to conditions of well-formedness; see (8).

(8) a. Every TBU must have a tone.
   b. Every tone must be associated to some TBU.
   c. Association lines must not cross.
(Pulleyblank, 1986:11; Yip, 2002:76)

However, as mentioned in the previous section 1.1, not all lexical elements are underlyingly associated with a tone and vice versa. All association lines, which are introduced by the grammar during the derivation process, are also subject to well-formedness conditions and conventions of association i.e. "…rules of vowel and tone epenthesis, vowel or tone deletion, ect. will automatically be followed by reapplication of the Association Conventions." (Pulleyblank, 1986:11). The association conventions are displayed in (9).

(9) Association proceeds one-to-one; left-to-right.
(Pulleyblank, 1986:11; Yip, 2002:76)

It is important to note that all processes that associate tone to TBUs are considered as part of lexical phonology as opposed to post-lexical processes, which will be further explained in section 1.4 and 1.5, following Pulleyblank (1986).

The nature of the TBU seems to be language-specific and has been a matter of debate; see e.g. Odden (1995) and Yip, (2002:73). It can either be the segment, the syllable (σ) or the mora (μ). For Akan, Abakah (2005:110) observes that "…the discussion is moot, because Akan does not have heavy syllables and, as a result the syllable and the mora, …, overlap.".

Yip (2002:73) states that "If the language has syllabic nasals which bear tone, but onset nasals which do not, we can rule out the segment as TBU.". This is the situation we find in Akan. Hence, I will assume that the syllable is the TBU.

Before continuing with the relevant arguments in favor of the auto-segmental theory, I will present a side note on the syllable structure of Akan. Structurally Akan prefers unmarked C(onsonant)V(owel) syllable structure (e.g. Dolphyne, 1988; Abakah, 2005; Adomako, 2008), like most of the Kwa languages (Manfredi, 1993), as in *sà.mí.nà* – 'soap'. Additionally, bare V syllables do occur, usually as word- or morpheme-initial vowel as in e.g. *ɔ̀.-dá.ń* – 'NOM-house'. Surface occurrences of CVV[3], which would count as heavy syllables, are reanalyzed as disyllabic CV.V; the vowel may be identical or not. Dolphyne (1988:5) analyzes identical VV sequences in e.g. *dá* – 'day' vs. *dá.á* – 'daily' and *kɔ̀* – 'to go' vs. *kɔ̀.ɔ̀* – 'very red' as phonemic vowel length. However, I will assume here in accordance with Abakah (2005:112) that vowel length is not categorical. Single C syllables do occur. They are made up of sonorants only. A syllabic consonant can occur word-initially as nominal prefix e.g. *ǹ-tò.má* – 'NOM-cloth' (Dolphyne 1988:58) and word-finally e.g. *nó.m̀* – 'to drink', *dà.ǹ* – 'turn it over' Dolphyne (1988:53). The syllabification is motivated by the observation that syllabic consonants carry their own tone. Therefore, Dolphyne (1988:53) and Abakah (2005:111) analyze surface instances of e.g. *dà.ǹ* as reduced forms of underlying CV.CV syllables in which the final vowel got lost. CCV sequences as in *ɔ̀.br̀.á* – 'life' are also reanalyzed as CV.CV syllables (Dolphyne, 1988:103, Marfo & Yankson, 2008).

Arguments in favor of an autonomous representation are presented in Yip (2002:74). In the following, I will briefly outline two of them, tonal mobility and tonal stability. Tonal mobility, also known under the term spreading[4] (Gussenhoven & Jacobs, 1998:138), refers to the situation in which a tone moves away from its point of origin. Paster (2010:104) presents a proposal of the derivation of the tonal surface pattern of a SVO sentence with H toned subject

[3] It is not clear whether Akan has stress, since no comprehensive analysis has been undertaken yet. Three recent studies (Purvis, 2009; Anderson, 2009, 2011) examine the rhythmic structure of Akan. Anderson (2009:140) concludes that Akan is not syllable-timed. He further states that L and H toned sentences behave rhythmically different, which may be caused by a durational difference. According to Manyah (2006), vowels that carry an L tone are significantly shorter (between 80 to 100 ms) than vowels that carry an H tone (against a universal tendency see e.g. Faytak & Yu, 2011). According to Christaller (1933), stress exists. He defines it as "emphasis put on a syllable" (Christaller, 1933, p. XXVIII) and differentiates between stress on verbs and on nouns. The stress in verbs can either fall on the prefix or on the stem. In nouns stress can either fall on the first H toned syllable, or on the preceding L toned syllable. In nouns containing only L tones, the first syllable carries the stress, as in wɔfa – 'uncle'. Since tone and rhythm might trigger the impression of stress, there are so far no convincing reasons to assume that Akan has stress.

[4] Tonal mobility could also be tone shifting, as in the South African Nguni languages (Zerbian, p.c.).

and verb in perfective aspect, involving H tone spreading. An example whose derivation involves H tone spread is displayed in (10).

(10) Wá-!káé Kòfí.
PRO.PRF-remember proper name
'You have remembered Kofi'.
(Paster, 2010:104)

The tonal processes and tonal associations applying during the derivation are reproduced in (11). Contrary to Abakah (2002), Paster assumes that the pronoun is underlyingly associated with an H tone. The perfective aspect is marked by a toneless morpheme /a/ and a floating H tone, as already mentioned in section 1.1. The verb is associated with LH tones. The underlying tones are represented in (11)a. The H tone of the subject spreads to the right onto the toneless aspect prefix. As Paster (2010:96) points out "… it is crucial to indicate that the target of spreading is toneless, because tone spreading does not apply to any mora that already bears a tone.", further examples of tone spreading on toneless TBUs will be provided in section 1.4.2 which deals with the so called 'boundary assimilation' process (Marfo, 2003, 2004, 2005). It becomes clear from her statement that she assumes the mora as TBU in Akan. However, this is not further motivated in the paper. (11)b. shows that the grammatical H̥ tone associates to the final syllable of the verb. After that perfect polarity, see (11)c., applies. It leads to an insertion of a tone of the opposite identity to that of the second syllable of the verb on the first syllable of the verb. The process of tone plateauing is illustrated in (11)d., Paster (2010:86) notes that "… a L-toned mora surfaces with a downstepped H tone between two H-toned moras." under specific conditions, which are not relevant here.". The L tone which was previously associated with the first syllable of the verb is now associated with an H tone. The L tone is set afloat and causes the following H tones to be realized lower on the surface (downstep): this is further exemplified in (12) and introduced in detail in section 1.3. Finally, the final vowel of the pronoun and the perfective prefix fuse, as illustrated in (11)e. However, it remains unclear whether the vowels fuse or whether the vowel of the pronoun is deleted due to hiatus resolution; see again (12) for an example. It should be noted that the materials presented in the thesis are constructed to avoid tonal processes in the lexical phonology as far as possible.

(11) a. H-toned subject

wo a- kae
| ||
H H̥ LH

b. tone spreading and grammatical tone association

wo a- kae

H H̥ H

c. perfect polarity

wo a- kae
/ |
H L H

d. tone plateauing

wo a- kae
\|
H L̥ H

e. vowel fusion

w a- kae
\|
H L̥ H

(Paster, 2010:104)

Another argument known as tonal stability (Yip, 2002:74), already touched above, refers to the situation in which an element on the segmental tier is deleted, e.g. for reasons of hiatus resolution. The tone of the deleted vowel remains unassociated (floating) on the tonal tier. The process is illustrated with the help of an associative construction (Abakah, 2000:197); see (12). The underlying tones are given in (12)a. A default L tone is assigned to the toneless nominal prefix. This is illustrated in (12)b. Two adjacent vowels at a morpheme boundary are dispreferred in Akan; see also Marfo (2004). This situation (hiatus) is resolved by deletion of the segmental content of the nominal prefix, as displayed in (12)c.

(12) a. underlying representation

L H		HH	tonal tier
| |		||	
Kofi	+	ɔ-dan	segmental tier
proper name		NOM-house	

b. default tone assignment

L H		L HH	tonal tier
| |		| ||	
Kofi	+	ɔ-dan	segmental tier

c. hiatus resolution

```
    L H            L̥  HH          tonal tier
    | |                ||
   Kofi      +        dan         segmental tier
```

(Abakah, 2000:197)

As pointed out above, there are good reasons to assume that the floating L tone (L̥) just remains in the tonal string since it has a lowering effect on the following H tone(s). The following section 1.3 provides background and details on tone terracing and chapter 4 section 4.1 offers phonetic details on the influence of L̥ during phonetic implementation.

1.3 Tone terracing

Akan has been classified as a terraced-level tone language (Christaller, 1875; Welmers, 1959, 1973; Clements, 1979; Dolphyne, 1988; Abakah, 2000). Stewart (1971:184) characterizes it as follows: “Where a high tone is followed by a Low tone which is followed in turn by a high tone in this language the second high tone is normally lower in pitch than the first, so that when the high tones of a sentence are interrupted by L tones at a number of points, the high normally descend in pitch by a series of steps from the beginning to the end of the sentence.”. The lowering of an H tone in the previously described environment is commonly termed downstep. Additionally, the termini automatic downstep (Stewart, 1965) and downdrift (Hombert, 1974) have been used to refer to the lowering of H tones. The term downstep will be used throughout the thesis. If downstep is caused by a floating L tone, I will use the term non-automatic downstep; see Connell (2002a), (2011) for a terminological overview. The figure 3 illustrates downstep in Akan with a sentence exhibiting alternating HL tones; see (13).

(13) Ánàné bìsá.
proper name ask.HAB
‘Anane asks.’

The second H tone on the third syllable of the subject *Anane* is lower than the first H tone on the first syllable. The final H tone on the last syllable of the verb is lower than that of the second H tone.

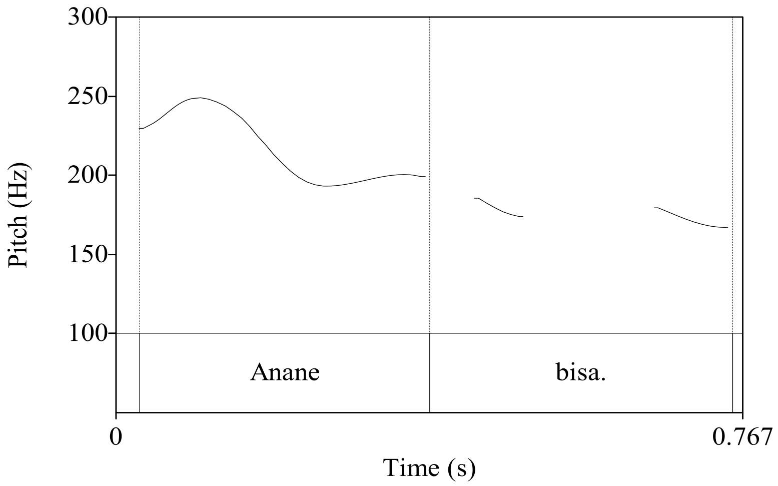


figure 3: Sentence with alternating HL tones; uttered by the female speaker.

According to Clements (1979:537), the terracing property can also be found in e.g. Igbo, Ga, Yoruba, Izi, Tiv, Efik, Mbemebe, Shambala, Sotho, Kikuyu and Zulu. An outstanding issue in the study of tone languages is the status of downstep in the grammar. Xu & Sun (2002) and Wang & Xu (2011) classify downstep as a phonetic effect. Downstep in Mandarin Chinese, also sometimes called carry-over lowering (Gandour et al., 1994; Xu, 1999), is analyzed as tonal co-articulation, which appears because more time is needed to implement a raising pitch movement from L to H (target undershoot). However, a purely co-articulatory explanation seems not be appropriate for terraced level tone languages like Akan for the following reasons:

- H tones are lowered after L̥.
  (e.g. Stewart, 1965; Dolphyne, 1994; Genzel & Kügler, 2011)
- Downstep does not apply sentence-initially.
  (e.g. Huang, 1985)
- A lowered H tone establishes a “ceiling” for following H tones.
  (e.g. Stewart, 1965; Huang, 1985; Clements, 1979)
- The intervals between (automatic) downstepped H tones are fairly stable.
  (e.g. Rialland & Somé, 2000)
- The second H tone in a HLH sequence is lowered in relation to the first.
  (e.g. Welmers, 1973; Huang, 1985)
- Downstep does not occur in all languages with two level tones.
  (e.g. Connell, 2002a, 2011)

The lowering of an H tone following a non-associated L tone is known under the heading non-automatic downstep (Stewart, 1965), often referred to as downstep, and is usually represented as !H. Non-automatic downstep is illustrated in figure 4. The sentence is displayed in (14).

(14)	Àgyèmàǹ	rè-síésìè	Kòfí	L̥!dáń	mú.
	proper name	PROG-clean	proper name	house	in
	‘Agyeman is cleaning Kofi’s house.’				

(14) contains an associative constructing in object position exhibiting non-automatic downstep; see (12) for details. Non-automatic downstep can be observed on the H tones of the word *dan*, which are realized lower than the preceding H tone on the second syllable of the proper name *Kofi*.

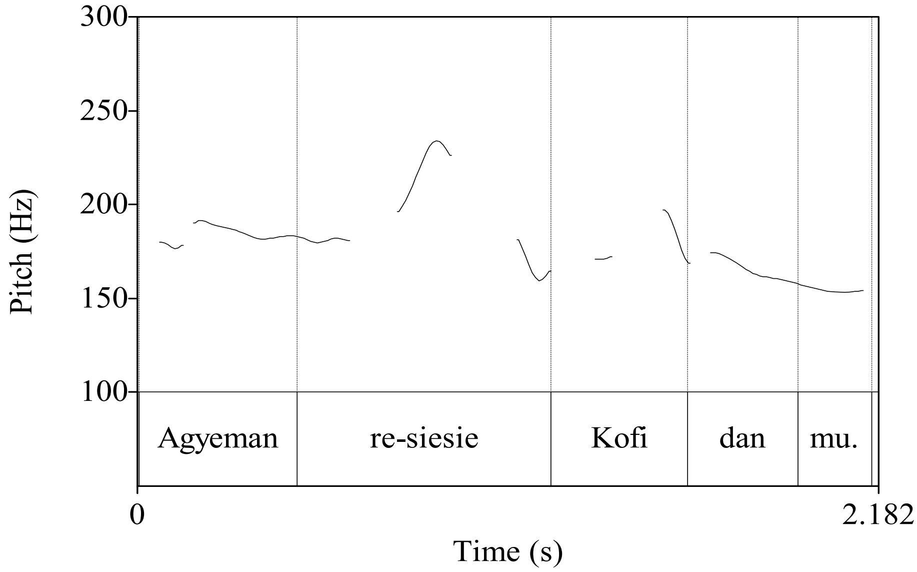


figure 4: Sentence involving non-automatic downstep; female speaker.

Since L̥ is not associated with a TBU, it is phonetically not realized. Hence, the lowering of the following H tone cannot be attributed to articulatory constraints in this case. Most African tone languages exhibit automatic and non-automatic downstep, such as Akan (e.g. Stewart, 1965; Dolphyne, 1988, 1994; Abakah, 2000, 2002; Genzel & Kügler, 2011), Igbo (Liberman, Schultz, Hong & Okeke, 1992; Laniran, 1992), Baule (Ahoua, 1996), Bimoba (Snider, 1998) and Chumburung (Snider, 2007).[5] The majority of studies providing experimental phonetic data on the similarity/difference between automatic and non-automatic downstep have shown, that the lowering of an H tone following L̥ is comparable to the lowering of an H tone following an associated L tone (e.g. Snider, 1998; 2007). However, the empirical investigation of the issue in Akan has revealed that the lowering in the case of non-automatic downstep is greater than for automatic downstep (Dolphyne, 1994); see chapter 4 section 4.1 for further details – the following question will be investigated:

- Is automatic downstep phonetically similar to non-automatic downstep?

Huang (1985) impressionistically observes for Akan that downstep does not apply sentence-initially when an H tone is preceded by an L tone. This is a crucial observation because if the lowering of H would be co-articulatory, it should also apply in initial position since it constitutes the critical environment (LH) for the application of downstep. Connell & Ladd

[5] Connell (2011:830) points out that the two do not necessarily co-occur; see also Hyman (1975:227). However, instances of non-automatic downstep in the absence of downstep are rare. Connell (2011) mentions Dschang, Kikuyu (Clements & Ford, 1979) and Ikaan (Salffner, 2010).

(1990) and Laniran & Clements (2003) present controlled experimental data for Yoruba confirming Huang's claim. They compared the height of the first H tone (H1) in an HL environment, see (15)a., to the height of H1 in an LH environment, see (15)b. Downstep is assumed to be absent if H1 in (15)b. is not realized lower than H1 in (15)a.

```
(15)    a.  H1 L   H   L          b. L   H1 L   H
              |   |   |   |             |    |   |   |
            CVCVCVCV             CVCVCVCV
```

In chapter 4 section 4.2, Huang's claim will be examined empirically and will address the following question:

- Does downstep apply to initial H tones, which are preceded by an L tone?

The observation that H tones following a 'downstepped' H tone are not scaled higher than the 'downstepped' H tone is one of the major motivations to model (automatic) downstep phonologically (e.g. Clements, 1979, 1983, 1990; Huang, 1985; Snider, 1999); see below. Under a co-articulatory account, this behavior is not necessarily expected. The figure 5 presents the pitch track of a Mandarin Chinese sentence with alternating HL tones (dotted line), reproduced from Wang & Xu (2011:604), together with a sentence containing only H tones (solid line). The second H tone of the alternating sequence is realized lower than the first but the pitch of the following third H tones is higher than that of the second. However, lowering still takes place since none of the non-initial H tones is realized as high as the H tones of the only H tones sequence.

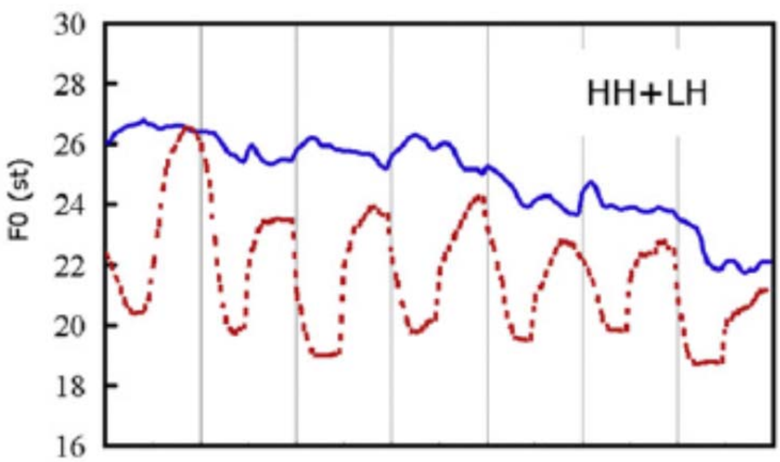


figure 5: Mandarin Chinese sentence containing only H tones (solid line), Mandarin Chinese sentence with alternating H and L tones (dotted line); reproduced from Wang & Xu (2011:604).

Let us consider the Akan sentence displayed in (16). The final syllable of the word *adamfo* is associated with an L tone. All following words carry H tones.

(16) Mé pàpá á-dáḿfò téńtéń nó bá.
PRO father NOM-friend tall DET child
‘The child of my father’s tall friend.’

The figure 6 shows the surface F0 pattern of the sentence in (16). The H tone following the L toned syllable (*fo*) is realized lower than the preceding H tone and, in opposition to the Mandarin Chinese example, all following H tones are realized even lower than the ‘downstepped’ H tone, this also know as ‘ceiling effect’ (e.g. Clements, 1979).

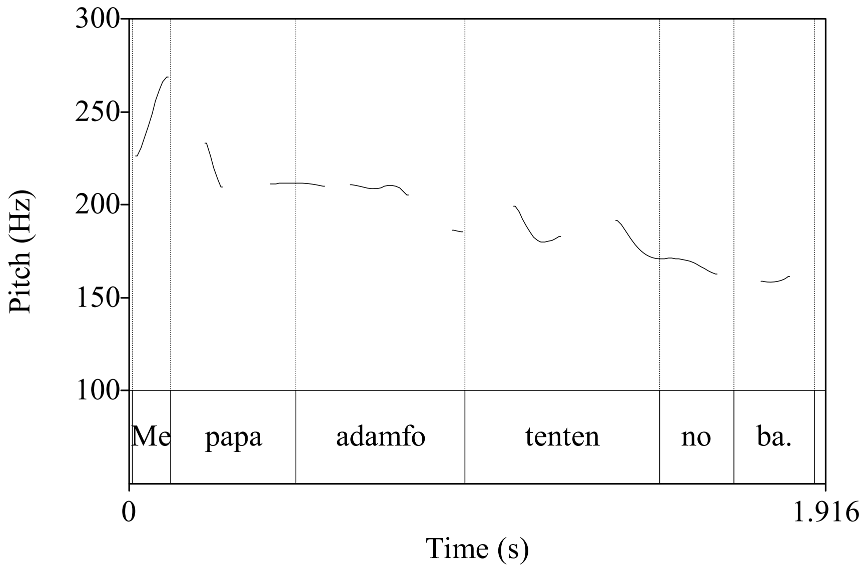


figure 6: Sentence illustrating the ‘ceiling effect’; female speaker.

The ‘ceiling effect’ has inspired many phonological proposals involving register tones (Huang, 1980, 1985; Clements, 1979, 1983, 1990; Hyman, 1985; Inkelas & Leben, 1990; Snider & van der Hulst, 1993; Snider, 1999). Register refers to a frequency band internal to the speaker’s range, “which determines the highest and lowest frequency within which tones can be realized at any given point in the utterance.” (Clements, 1990:59). All these approaches, despite their feature geometric differences and theoretical procedures, have in common that a “downstepped” H tone is associated (directly or indirectly) with a low register tone (l), which is represented on a register tier. The phonetic component interprets the l tones on the register tier as progressively lowering of the H tones on the tonal tier (e.g. Hyman, 1985); see chapter 4 section 4.2 for details and arguments against the register tone approach. Clements (1990) points out that register is a useful concept but it is unclear whether and how it must be expressed in phonological representation. I will follow that idea but rather use it as a holistic concept, which means that there is mental representation of the register as defined above, though not with utterance internal register tones, but rather with a starting point and an end point only, as suggested by Möhler & Mayer (2001); see chapter 4 section 4.3.1 for further details.

The stability of downstep intervals is not expected under a co-articulatory account since more gradience (e.g. Xu, 1993) is likely to occur, see also figure 5. Furthermore, it has been observed that the second H tone in an HLH sequence is lowered in relation to the first (e.g. Welmers, 1973; Huang, 1985). This observation clearly speaks against the assumption that H is lowered because more time is needed to implement a raising pitch movement from L to H (Xu & Sun, 2002). The presence of the intervening L tone plays a crucial role in the generation of downstep in all models, be it phonological (Clements, 1983; Huang, 1985) or phonetic (Liberman & Pierrehumbert, 1984); see chapter 4 section 4.2 and 4.3.3 for further details.

Turning to the last argument against a purely co-articulatory account of downstep, it has been noted that the presence of downstep constitutes a difference in the phonology of a language since not all languages with two level tones like Akan exhibit downstep (Childs, 2003; Connell, 2002a, 2011). Examples found in the literature are Loma (Welmers, 1973), Jukun (Courtenay, 1971; Anderson, 1978), Ikaan (Salffner, 2010) and Embosi (Rialland & Aborobongui, 2010; Downing & Rialland, 2012). A closer look into the database of the examples reveals that the situation is more complex. Welmers (1973:93) reports for Loma: "…I do not recall any temptation to transcribe a higher-lower-higher sequence as tentatively "high-low-mid", which I probably would have done if there were an appreciable lowering of high after low. Throughout an utterance, Loma … seems to show little if any lowering….". If we take the latter statement more generally, it might be possible that Loma does also not show declination. The data basis for Jukun is not very extensive. Courtenay (1971) cites Welmers (1959) who presents Jukun as a prime example of a "discrete-level" language. In "discrete-level" languages the pitch ranges of tones do not overlap (Welmers, 1973:81) i.e. H tones are not lowered in reference to each other. Courtenay (1971:242) remarks the following in a footnote: "Welmers states now [personal communication] that this particular Jukun sentence is grammatically not very well-formed, though perhaps acceptable. At any rate, this information has no bearing on the discussion of the Jukun tone system.". Additionally, Hyman (1975) states that Junkun is a three-tone system (L, M, H).

Salffner (2010:21) notes that Ikaan is a two tone language with two allotones "…a downstepped H and an extra low tone…." (!H only occurs after L̥). The figure 7 presents a sequence of only L tones (left hand side) and alternating LH tones (right hand side) in Ikaan; a sentence with only H tones was not available. However, all instances involving a series of H tones, presented in her work, rather show an upsweep in F0. The sentence containing only L tones, see figure 7, does not show a clear declination effect (gradual lowering over the course of the utterance, see section 1.6 for further details) and the sentence with alternating LH tones lacks downstep.

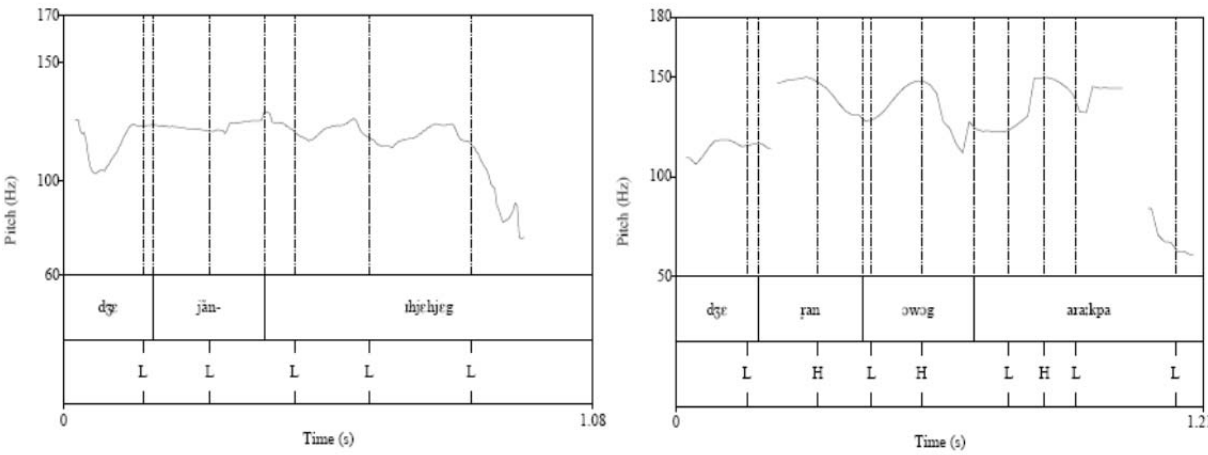


figure 7: Left pitch track Ikaan sentences with only L tones Dʒɛ̀jã̀nìhjɛ̀hjɛ̀g 'I bought sponges.'; Right pitch track sentences with alternating LH tones Dʒɛ̀ɽ̥ã́n ɔ̀wɔ́g àrákpà. 'I made bean soup.'; reproduced from Salffner (2010:137, 143); male speaker.

Finally, Embosi may not totally lack downstep. A pitch track is presented in figure 8. The two H tones in the beginning of the second prosodic word (ω), see following section 1.4 for details on prosodic constituents, are realized lower than the previous H tone. However, downstep may be masked by the fact that intonational tones; see section 1.5 for more information, are realized as superimposed register tones. A high boundary tone (H%), used to mark continuation, is not realized at the end of the topic but shows up on top of the last H tone of the topic.

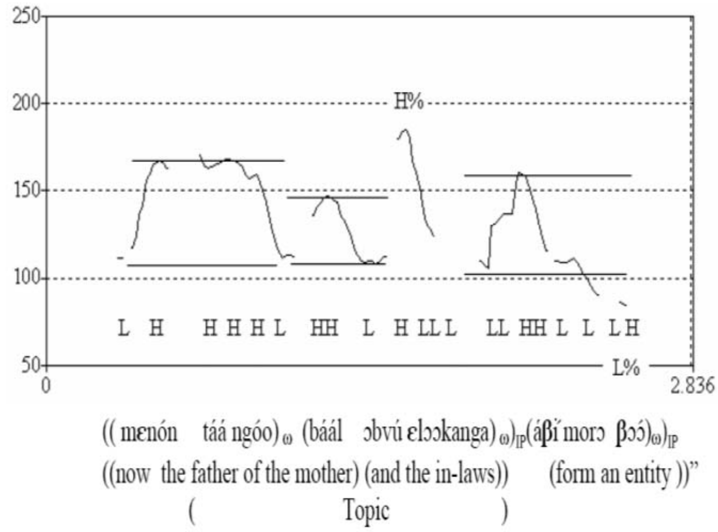


figure 8: Embosi sentence with superimposed high boundary tone (H%); reproduced from Rialland & Aborobongui (2010:45).

Summarizing the observations from languages without downstep, we have seen that for Loma and Jukun, the empirical evidence is not satisfactory. In Ikaan declination may be absent and in Embosi superimposed intonational register tones may prevent downstep. It looks as if downstep is the “unmarked” pattern in African tone languages with two level tones but its presence seems to depend on the existence of declination and the phonological system of the language (intonational system, number and nature of tonal contrasts).

The presence of downstep can be investigated experimentally by comparing the height of H3 in a sentence with only H tones, see (17)a., to H2 in a sentence with alternating tones (HL/LH), see (17)b. (e.g. Lindau, 1986; Beckman & Pierrehumbert, 1992; Connell, 2002a; Laniran & Clements, 2003) and by comparing the amount of pitch drop between H1 and H3 (H1-H3) in (17)a. to the amount of pitch drop between H1 and H2 (H1-H2) in (17)b. If the H2 tone in (17)b. is realized lower (↘) than H3 in (17)a. and the amount of pitch drop is greater in (17)b. than in (17)a., downstep is at work.

```
(17)   a. H1 H   H3 H        b. H1 L↘H2 L
          |  |   |  |           |  |  |  |
         CVCVCVCV              CVCVCVCV
```

In chapter 4 section 4.3.3, the following questions will be investigated.

- Is downstep to be regarded as independent effect i.e. is downstep different from other pitch lowering processes (declination)? If yes, is downstep to be represented in the phonology or in phonetics?
- How can downstep (automatic and non-automatic) be modeled?

1.4 Prosodic structure

This subchapter provides background information on the prosodic structure. Section 1.4.1 presents a brief introduction to the basic prosodic units (constituents) and their relation to syntactic constituents. The following section 1.4.2 is concerned with universal and language-specific phonetic reflexes that accompany higher level prosodic constituents such as the phonological phrase and the intonation phrase.

1.4.1 The prosodic organization of constituents

Besides lexical and lexical/grammatical tones, a crucial part of the phonological representation is the prosodic structure. It is generally assumed that the syntactic structure of a sentence is reflected in the prosodic structure i.e. the syntactic phrase structure is mapped onto a prosodic phrase structure; see Elordieta (2008) for an overview on different theoretical approaches of the syntax-prosody interface. The basic prosodic units[6] are the syllable (σ), the prosodic word (pw, ω), the phonological phrase (pP, φ), the intonational phrase (IP, ι) and the utterance (U)[7] (Nespor & Vogel, 1986). The units are hierarchically structured. The σs represent the lowest level. They are dominated by the pws. The pws are dominated by the pPs. One or more pP(s) form an IP. The IP is dominated by the highest level of the prosodic hierarchy, the U.

In what follows, I will present a very minimalistic introduction to the mapping of the syntactic and the prosodic structure, following Büring (2010). Only pPs and IPs will be considered. According to Büring (2010), 'normal' (default/unmarked) mapping from syntactic structure to prosody proceeds roughly as follows: each lexical word (X^0) is mapped onto a pw, each maximal projection (XP) onto a pP and each sentence onto an IP. Languages differ with regard to the mapping strategy. Büring (2010:181) distinguishes between radical splitting, moderate and radical mapping[8]. In a language, which favors radical splitting, see (18)a., each XP and any remaining non-phrasal elements are mapped onto their own pP's. A radical mapping language maps all XPs onto one big pP, see (18)b., and a moderate mapping language would map each XP onto its own pP and non-phrasal elements together with the structurally closest phrase, as illustrated in (18)c. This is the case for Akan, as we will see later in the following section 1.4.2.

[6] Ladd (1996:10) refers to the prosodic units as phonological constituents; see section 1.5 for further details.
[7] Since all sentences used in this thesis are single sentences, IP and U overlap.
[8] Büring (2010) uses the term wrapping instead of mapping following Truckenbrodt (1995).

(18) a.

```
(                  )ι
(   )φ  (   )φ ( )φ
 XP     YP     Z
```

b.

```
(                  )ι
(                  )φ
 XP     YP     Z
```

c.

```
(                  )ι
(   )φ  (          )φ
 XP     YP     Z
```

(Inspired by Büring, 2010:181)

The notion of headedness plays a crucial role in Büring's original proposal. Theories of the syntax-prosody interface are largely based on intonation languages, like German and English. All tonal events (pitch-accents) in these languages are considered as post-lexical i.e. "…the choice of melody is not entailed by the choice of words, but rather functions independently to convey pragmatic meaning." (Pierrehumbert, 2000:11). According to Büring (2010), pPs and IPs are identified with the help of accents (prosodic heads). The notion of prosodic headedness crucially relies on the occurrence of pitch accents, which associate with metrically strong elements such as stressed syllables[9] at the level of pP and IP. This concept is problematic in the context of tone languages like Akan. A growing body of research on tone languages is concerned with the question of phrasing and phonetic cues to identify prosodic domains on different levels of the prosodic hierarchy; see e.g. Hyman (2003) and Zerbian (2004) for an overview on phrasing in Bantu languages and references therein.

1.4.2 Prosodic constituents and their phonetic properties

Phonetic cues of prosodic phrasing fall into two groups: universal and language-specific. Among the universal cues are pauses[10], final lengthening, reset (Vaissière, 1983) and possibly final lowering (Arvaniti & Godjevac, 2003). Pauses are related to the respiratory system (inhalation phase). A number of researchers have shown that pauses occur at major syntactic boundaries (Fletcher, 2010:574 and references therein). Since the syntactic structure is mapped onto the prosodic structure, as outlined above, pauses also occur at boundaries of higher prosodic domains such as pPs (Horne, Strangert & Heldner, 1995) and/or IPs (Ferreira,

---

[9] Prominence relations (strong/weak; stressed/unstressed) in Akan are already problematic at the very basic prosodic levels (foot). Anderson (2011:1) observes that: "… register tone languages do not usually exhibit stress, nor is there much evidence of phonological rules applying to the metrical structure….".

[10] I refer here to silent pauses, which according to Campione & Véronis (2002) can be grouped into three categories: brief pause (<200ms), medium pause (200 – 1,000 ms) and long pause (>1,000ms).

1993; Horne et al., 1995; Krivokapić, 2007). Furthermore, a relation between prosodic boundary strength and pause duration has been observed, roughly speaking: the higher the prosodic domain the longer the pause (Fletcher, 2010:574 and references therein). Pauses as indicator of prosodic boundaries in Akan will be employed later in this section.

Final lengthening might occur in connection with pauses. Lengthening of the final vowel[11] before a pause, also known as pre-pausal lengthening, has been reported for a number of typologically unrelated languages (Vaissière, 1983:60). It may accompany pPs and IPs (e.g. Féry, Hörnig & Pahaut, 2010). Like final lengthening, resetting of the F0 might occur in connection with pauses. F0 values in declaratives are usually lower at the end than at the beginning (e.g. Gussenhoven, 2004), the resulting downward trend in F0 is known as declination (Cohen & ´t Hart, 1967). Vaissière (1983:57) notes that "… resetting is used as a boundary marker, and the degree of resetting indicates the importance of the boundary." and concludes that: "The physiological basis of the relations among pauses, breathing, declination and resetting is difficult to establish, since speakers may pause without breathing, or reset the baseline without pausing….". Declination reset rather occurs at IP boundaries (Terken, 1993), but see Laniran & Clements (2003) for resetting in downstepping sequences within the IP in Yoruba.

Final lowering refers to an abrupt fall in F0 in the last centi-seconds of an utterance/IP (Herman, Beckman & Honda, 1996). It has been observed in numerous languages with different prosodic systems (e.g. Pike, 1948; Japanese, Poser, 1984; Pierrehumbert & Beckman, 1988; Dutch, Gussenhoven & Rietveld, 1988; Yoruba, Connell & Ladd, 1990; Kipare, Herman, 1996). Herman et al. (1996) and Herman (2000) argue that final lowering can be attributed to a decrease of articulatory effort towards the end of an utterance/IP.

Turning to language-specific cues, prosodic domains[12] can be identified with the help of language-specific phonological processes, which apply in certain domains or are blocked at specific prosodic boundaries (Nespor & Vogel, 1986). In the following, I will outline which prosodic domains have been established in Akan with reference to phonological processes, such as vowel harmony and tone spreading. The pw as a prosodic unit in Akan has been found to be useful for the description of the vowel harmony process (Ballard, 2010). Before going into the details a short introduction to the vowel system of Akan is provided in table 1.

[11] Note that the domain of final lengthening seems to be language-specific (e.g. Nakai, Kunnari, Turk, Suomi & Ylitalo, 2009).

[12] Recent research (Bickel, Hildebrandt & Schiering, 2009) suggests that the prosodic hierarchy may not be universal, i.e. languages may differ in the choice of smaller prosodic domains e.g. no evidence for pws in Vietnamese.

According to Dolphyne (1988:17), Akan uses five basic vowels with advanced tongue root (ATR) /i/, /e/, /a/, /o/, and /u/, each of which has an restricted tongue root (RTR) counterpart /ɪ/, /ɛ/, /æ/, /ɔ/, and /ʊ/ (Stewart, 1967, 1970; Schachter & Fromkin, 1968; Dolphyne, 1988); see table 1. /æ/ has a limited distribution (before /i/ and /u/) and is assumed to be a variant of /a/ by some authors (e.g. Dolphyne, 1988:7; Baković, 2003:11). According to Dolphyne (1988:7), the low vowels are not specified for backness and all back vowels are rounded, whereas all front vowels and the central /a/ are unrounded. Akan behaves thus as the majority of languages with regard to the frontness/roundness relation (Ladefoged & Maddieson, 1996).

	-BACK		**+BACK**	
	+ATR	**-ATR**	**+ATR**	**-ATR**
+HIGH, -LOW	i	ɪ	u	ʊ
-HIGH, -LOW	e	ɛ	o	ɔ
-HIGH, +LOW	æ a			

table 1: Vowels of Akan and their phonetic description.

Some of these oral vowels have nasalized counterparts, these are: /ĩ/, /ã/, /ẽ/, /ũ/ and /õ/ (Schachter & Fromkin, 1968:25), as in e.g. *fĩ* – 'dirt' vs. *fi* – 'to go out' and *kã* – 'to say' vs. *ka* – 'to be left behind' (Dolphyne 1988:4). Nasal vowels are regarded as phonemes since they occur also with non-nasal consonants (Westermann & Ward, 1990:43).

Tongue root harmony can be observed between verb root and pronominal prefixes, tense/aspect affixes and nominal affixes (Dolphyne, 1988; O'Keefe, 2003). Following Clements (1981), the above mentioned affixes are regarded as underspecified for tongue root, which will be marked by capital letter in the transcription. In (19)a. the harmony process is exemplified with a pronominal prefix (Dolphyne, 1988:15; Ballard, 2010:13). The underspecified vowel of the pronominal prefix is realized as +ATR on the surface harmonizing with the tongue root specification of the vowel of the verb, which is specified as +ATR. The pronominal prefix and the verb form a pw, following Nespor & Vogel (1986:110f.) who state that a pw may be composed of a stem plus all affixes. (19)b. shows that the harmony is bidirectional with an example involving a +ATR root vowel, an underspecified pronominal prefix and an underspecified tense suffix (Ballard, 2010:13). The spreading of an –ATR specification of the vowel root into the pronominal prefix is exemplified in (19)c.

(19)	a.	/wU-	di/		→	[wu-di]ω
		2.SG	eat			
		'You eat.'				
	b.	/O-	hũ -	I/	→	[o-hũ-i]ω
		3.SG	see	PST		
		'She saw.'				
	c.	/wU-	bE-tʊ/		→	[wʊ-bɛ-tʊ]ω
		2.SG	FUT.throw			
		'You will throw.'				

(Ballard, 2010:13)

Kügler (2012) presents evidence for a pP boundary between subject and verb. Vowel harmony applies between verb, which is underlyingly specified for (-ATR) and object, which is specified for (+ATR), as illustrated in (20)a. The process is blocked by a pP boundary between subject (-ATR) and verb (+ATR); see (20)b. The prosodic structure indicated in (20) is reminiscent of Büring's (2010) moderate wrapping language.

(20)	a.	Sukuni	to (/ʊ/)	kube.
		(([sukuni])φ	([tu]	[kube])φ)ι
		student	throw.HAB	coconut
		'A student throws a coconut.'		
	b.	A-dokɔ	di	a-buro.
		(([a-dʊkɔ])φ	([di]	[æ-buro])φ)ι
		NOM-chicken	eat.HAB	NOM-maize
		'A chicken eats maize.'		

(Kügler, 2012)

Marfo (2003, 2004, 2005) also suggests that Akan is basically of the moderate mapping type. This is illustrated in (21)a. (IP marking was added to the original prosodic structure). Furthermore, he claims that Akan shows characteristics of a radical splitting type language, if the verb (root) is made up of more than one syllable, as illustrated in (21)b.; see Jun (1998) for length related phrasing differences in Korean[13].

(21) Syntactic structure: $[_{\text{InflP}}$ [NP] $[_{\text{VP}}$ *V* [NP]]]

a. Prosodic structure: ((NP)φ ($V_{\text{mono-syllabic}}$ NP)φ)ι

b. Prosodic structure: ((NP)φ ($V_{\text{di-syllabic}}$)φ (NP)φ)ι

(Marfo, 2003:130; with slight modifications of the syntactic structure)

[13] I am grateful to Sabine Zerbian for directing my attention to this paper.

Marfo (2003, 2004, 2005) presents typologically surprising evidence of a tonal spreading process, which he calls ‘boundary assimilation’ applying at the right edge of pPs in Akan. At the pP juncture “…a final tone of a word spreads …” (Marfo, 2003:130) on the first syllable of the following word. Marfo presents data of ‘boundary assimilation’ between subject NP and perfective prefix as evidence for a pP boundary. The ‘boundary assimilation’ process is illustrated in (22), based on Marfo (2003:130). (22)a. illustrates H tone spreading and (22)b. L tone spreading. Note that he assumes that the perfective prefix is L toned. I left it unspecified following Abakah’s (2002, 2005) and Paster’s (2010) claims.

(22) a. Syntactic structure: (with lexical tones)

LH H L L
[$_{\text{InflP}}$ [$_{\text{NP}}$Kofi] [$_{\text{VP}}$ a-bo [$_{\text{NP}}$Ado]]]
proper name PRF-beat proper name
‘Aku has beaten Ado.’

Prosodic structure:

L H H H L L
((Kofi)$_{\varphi}$ (a-bo Ado)$_{\varphi}$)$_{\iota}$

b. Syntactic structure: (with lexical tones)

LL H H H
[$_{\text{InflP}}$ [$_{\text{NP}}$Yaw] [$_{\text{VP}}$ a-bo [$_{\text{NP}}$ Esi]]]
proper name PRF-beat proper name
‘Yaw has beaten Esi.’

Prosodic structure:

LL L H H H
((Yaw)$_{\varphi}$ (a-bo Esi)$_{\varphi}$)$_{\iota}$

(Based on Marfo, 2003:130)

Apart from the syntactic structure of a sentence, prosodic phrasing is influenced by additional factors, such as the length and speech rate (e.g. Gee & Grosjean, 1983; Jun, 1998). More phrases are expected with increasing length of a phrase (e.g. Jun, 1998, 2003; Marfo, 2003, 2004, 2005) and more phrases are expected at slower rates and vice versa (e.g. Vaissière, 1983; Fougeron & Jun, 1998). Marfo (2003:130), referring to the phrasing in (21)b., remarks in a footnote that: “This mapping in Akan is, however, subject to phonetic condition of speech rate. Thus, in a slow speech a disyllabic verb stem, like a mono-syllabic one, may be mapped into a common φ with its (non-branching) complement.”. Marfo predicts less phrasing under slow speech rate, which contradicts the relation established from the literature. In the following, I will provide self-recorded acoustic data based on Marfo’s examples. In a first step, we will have a look at ‘boundary assimilation’ for L and H tones in examples involving a mono-syllabic verb in perfective aspect, uttered at different speech rates (moderate & slow). Additionally, data containing verbs with progressive aspect prefix uttered at moderate speech rate will be evaluated. After that, we will have a look at ‘boundary assimilation’ between verb and object.

The figure 9 presents pitch tracks of the data presented in (22)a., at moderate (left hand side), and slow speech rate (right hand side). Generally, more pauses occurred at slow speech rate, also between the mono-syllabic, underlyingly H toned verb, and the object. At moderate speech rate no pauses occur. According to Marfo’s ‘boundary assimilation’ account, the perfective prefix is expected to surface with an H tone, because it is immediately preceded by an H toned syllable. The expectation is borne out; see figure 9. Additionally, it can be observed that the tone on the verb root is subject to tone polarization. It surfaces with an L tone at both speech rates.

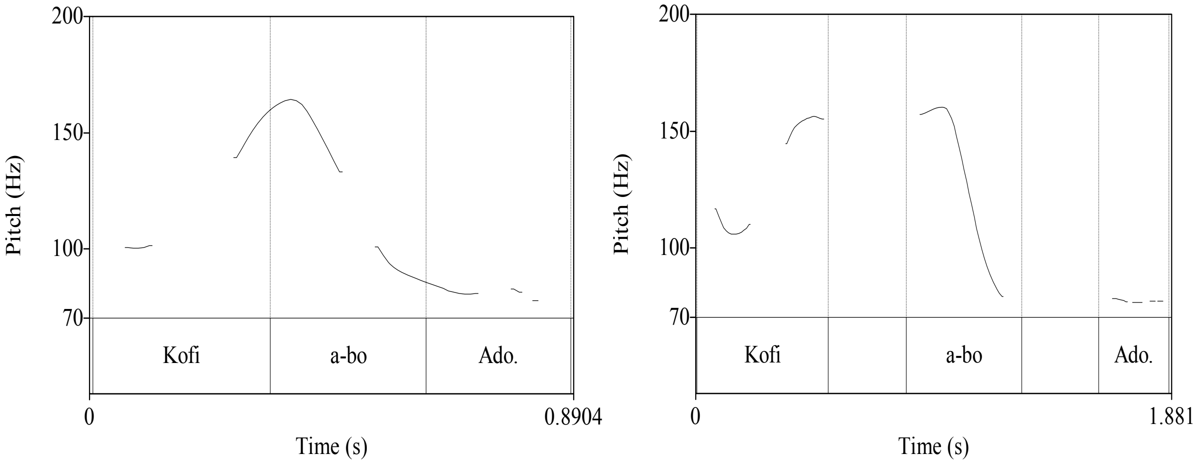


figure 9: Left hand side, pitch track of a sentence with LH toned subject NP and perfective aspect uttered at normal speech rate; Right hand side, pitch track of a sentence with LH toned subject NP and perfective aspect uttered at low speech rate; male speaker.

Turning to the examples involving an L toned subject represented in figure 10, see (22)b., uttered at moderate (left hand side) and slow speech rate (right hand side), Marfo’s account predicts that the perfective prefix should be realized with an L tone because it is immediately preceded by an L toned subject. However, the prediction is only borne out at normal speech rate and crucially not at slow speech rate, which shows a clear boundary marked by a pause. In this case the perfective prefix appears with an H tone on the surface. Furthermore, if the perfective prefix is realized with an H tone as in the slow speech rate, the tone on the verb (root) polarizes.

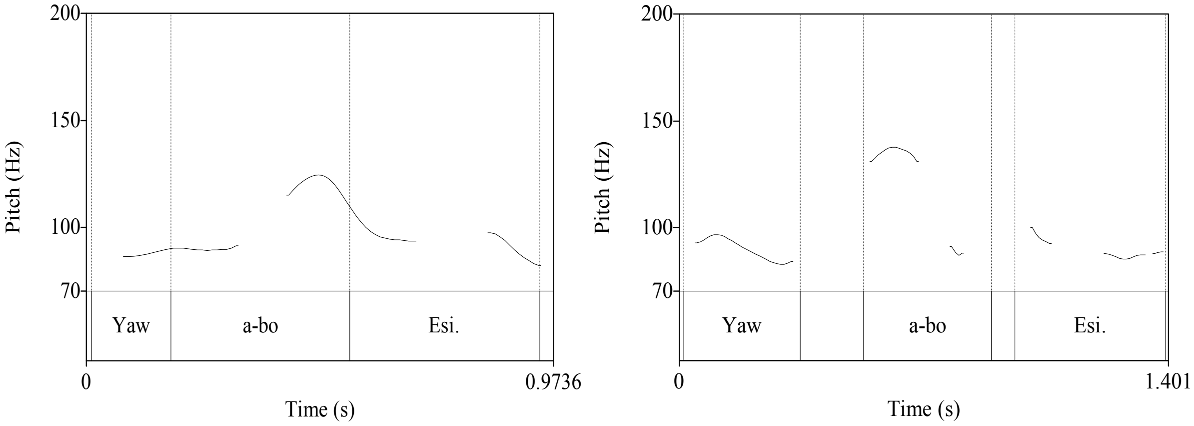


figure 10: Left hand side, pitch track of a sentence with L toned subject NP and perfective aspect uttered at normal speech rate; Right hand side, pitch track of a sentence with L toned subject NP and perfective aspect uttered at slow speech rate; male speaker.

Paster (2010:103) presents an alternative proposal to account for the tonal changes on verbs in perfective aspect; see section 1.2, (11). Under her analysis, the perfective prefix associates with a grammatical floating H tone, irrespective of the tonal specification of the subject NP. If the verb root carries a lexical H tone it is subject to tone polarisation. Hence, it appears with an L tone on the surface, as in the sample pitch tracks above; except for figure 9 right hand side. For this case, she also observes that the perfective prefix surfaces with the same tone as the final syllable of the preceding subject NP. Paster (2010) analyzes the process as tonal spreading like Marfo, but crucially without reference to the prosodic structure.

Marfo (2003) presents further indirect evidence for tone spreading as boundary marker involving progressive aspect. Marfo (2003:130) claims that 'boundary assimilation' is blocked if the tense/aspect prefix of the verb exhibits a consonantal onset. This is illustrated in (23). The H tone of the final syllable of the subject does not spread onto the progressive prefix.

(23) Syntactic structure: (with lexical tones)

```
                 H H          H         L  L
                 | |          |         |  |
[InflP [NPAku]  [VPre-bo     [NPAdo]]]
proper name      PROG-beat   proper name
'Aku is beating Ado.'
```

Prosodic structure:

```
 H H     L  H  L L
 | |     |  |  | |
((Aku)φ (re-bo Ado)φ)ι
```

(Based on Marfo, 2003:130)

It is important to note that the progressive aspect prefix is usually not pronounced as *re* in Asante but rather in Akuapem. In Asante the prefix "…is realised as lengthening of the vowel of the preceding syllable”, as in Kòfí ì-kó. ‘Kofi is going.’ (Dolphyne, 1988:92). It thus does not constitute an optimal test case for the blocking of spreading as proposed by Marfo, but for ‘boundary assimilation’. According to Marfo (2003:130), the tone of the final syllable of the subject should spread onto the progressive prefix of the verb. Hence, it should appear with surface H tone if it is preceded by an H toned subject NP, as in the examples in figure 11; see (23) and (24) for details.

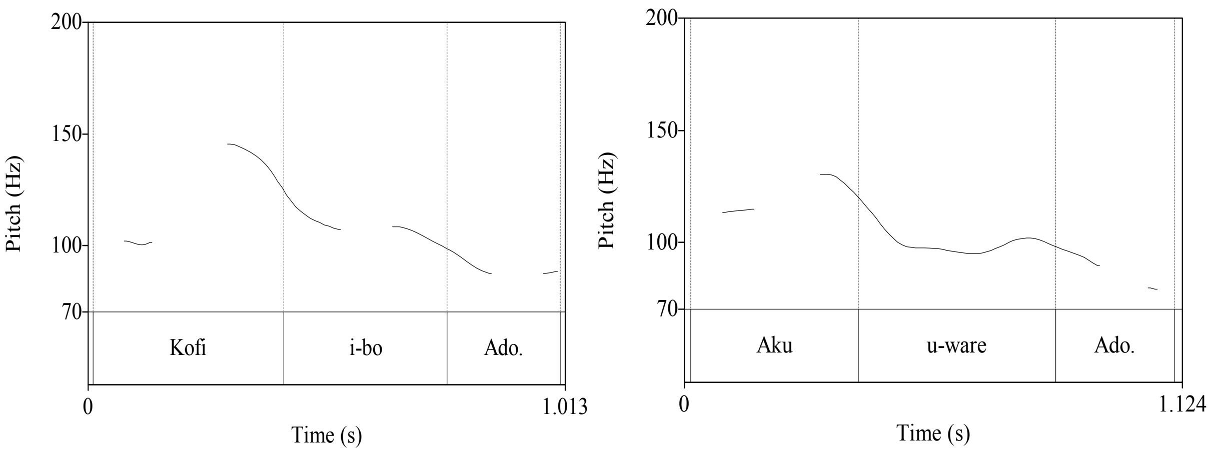


figure 11: Left hand side, pitch track of a sentence with LH toned subject and mono-syllabic verb in progressive aspect, uttered at moderate speech rate; Right hand side, pitch track of a sentence with H toned subject and di-syllabic verb in progressive aspect, uttered at moderate speech rate; male speaker.

However, the data shows that the progressive prefix occurs with an L tone at the surface irrespective of the tonal specification of the preceding syllable. The verb root retains its underlying tonal specification. This pattern fits in with Paster’s (2010:98) observation that the progressive prefix is a single L toned mora which is underlyingly not associated with any segmental features and that it takes on the quality of the preceding segment.

Turning to the tonal events at the vicinity of verb and object NP, a phrasing difference between mono- and di-syllabic verbs is expected, following Marfo’s claim; see (21). ‘Boundary assimilation’ should apply between verb (root) and object if the verb is disyllabic and speech rate is moderate, as illustrated in (24) with an example from Marfo (2003:130). The H tone of the final syllable of the disyllabic verb spreads onto the first syllable of the object. The underlying L tone is dislodged and set afloat.

(24) Syntactic structure: (with lexical tones)

```
                 H H          LH        L   L
                 | |          | |       |   |
[InflP [NP Aku]   [VP re-ware    [NP Ado]]]
proper name      PROG-marry   proper name
'Aku is marrying Ado.'
```

Prosodic structure:

```
  H H     L  L H     L̥  L
  | |     |  |  |\    |
((Aku)φ (re-ware)φ (Ado)φ)ι
```

(Based on Marfo, 2003:130)

The data presented in figure 11 constitutes an ideal test case to evaluate whether 'boundary assimilation' takes place between disyllabic verb and following object. In the track on the left hand side, the verb (root) is mono-syllabic. Hence, no tonal spreading should appear. The first syllable of the object NP, which is lexically L toned, should retain its tone. In the track on the right hand side, we can observe a disyllabic verb (root). Following Marfo's claim, the first syllable of the object NP should surface with an H tone due to tonal spreading at the pP boundary from the H toned verb root, as illustrated in (24). However, in both cases the pitch on the first syllable of the object falls towards an L target; see also chapter 7 section 7.2, crucially no difference in the tonal specification can be observed.

In sum, we have seen that, contrary to Marfo (2003, 2004, 2005), pP boundaries in Akan are not marked by 'boundary assimilation'. Pauses, especially at slow speech rate, and blocking of vowel harmony (Kügler, 2012) may serve as indicator of pPs in Akan. Furthermore, the acoustic data presented here, provided no evidence in favor of a phrasing difference due to the length of the verb (root). Observed tonal shifts on the tense/aspect prefixes of the verb are rather due to floating/associated grammatical tones and tonal spreading on toneless affixes, as proposed by Paster (2010). This may also apply to nominal prefixes; see below. I will thus assume that a simple SVO sentence in Akan exhibits the prosodic structure displayed in (25), which involves recursion (see e.g. Ladd, 1986, 1996; Selkirk, 2009; Féry, 2010 and Féry & Schubö, 2010 for the arguments in favor of recursion in the prosodic structure).

(25) Syntactic structure: $[_{\text{InflP}}$ [NP] $[_{\text{VP}}$ *V* [NP]]]
Prosodic structure: ((NP)φ (*V* (NP)φ) φ)ι

The prosodic structure in (25) will become relevant for the evaluation of the size of the processing window available to the phonetic implementation component, explored under the heading of anticipatory raising in chapter 5 section 5.2. Pauses, as indicators of phrasing, will be utilized in chapter 7 section 7.2, which is concerned with the prosodic marking of focus.

Turning to the phrasing and its phonetic cues in more complex sentences, Marfo (2003) provides evidence for phrasing at the level of the IP in Akan on the basis of the so called ‘focus’ or ex-situ construction which exhibits the characteristics listed (26).

(26) a. left-peripheral dislocation of the focused element
b. introduction of a clitic morpheme after the focused element
c. pronoun resumption in a canonical clause position
(Marfo & Bodomo, 2005:180)

The notion of focus as one of the key concepts of information structure will be defined in the introduction to section 1.5.3. Let us assume for now that elements that are of greater importance in a discourse (Hirschberg, 2002) are in focus and that the ‘focus’ construction resembles the structure of a cleft; see (27) for a schematic representation.

(27) (ɛ-yɛ) [Focused XP] na [S*V*O] (no)
(Kobele & Torrence, 2006:165)

The copula *ɛ-yɛ* ‘it is’ may optionally precede the element under focus. As a precision of (26)a., it has been observed that all elements (subject, object, verb, adverb) can be fronted in Akan (Boadi, 1974; Kobele & Torrence, 2006). The fronted element is followed by the particle *na*, (26)b. I will use the neutral term PART (particle) in the glosses since its function is not agreed upon. *Na* is analyzed as focus marker (FM) by Boadi (1975), Saah (1988, 1994) and Amfo (2010a). Boadi (1974:7) refers to *na* as exclusive focus marker and notes that ‘‘…na narrows down the referential range of the constituent to which it is attached and places it in an exclusive class by itself, thus bringing this constituent into sharp contrast with all other members of the paradigm to which it belongs’’. Fiedler & Schwarz (2005), however, refer to *na* as conjunction[14]. Complementary to (26)c., it has been observed that a resumptive pronoun may appear in the base position of the element under focus, if the fronted element is animate (e.g. Saah, 1988; Ermisch, 2006; but see Fiedler & Schwarz, 2005). At the right

[14] They base their analysis on the observation that “Akan has a clausal sequential conjunction *nà* with the meaning ‘and (then)’ (Bearth 2002) which is identical with the FM.” (Fiedler & Schwarz, 2005:127). Christaller (1875:90) as well as Amfo (2007:3) note that the conjunction *na* is L toned and e.g. Marfo & Bodomo (2005:185) observe that *na* is also L toned. However, Ofori (2011:260) supposes that *na* has developed by merge of *ne* “…the basic focus verb/marker/copula” and *a* “…‘the relative marker.’”.

periphery *no* may occur which adds "...some type of "emphasis"." (Kobele & Torrence, 2006:165)[15].

Turning to the prosodic structure of the 'focus' construction, Marfo (2003:131f.) only presents data with animate fronted XPs and observes that they form their own IP together with *na* and the following element. This is illustrated in (28). If the element following *na* is pronominal, in cases of focus on the subject, it forms its own pP together with the fronted element as in (28)a. In cases of focus on the object, the fronted element is phrased separately, as in (28)b.

(28) Syntactic structure: [$_{\text{InflP1}}$[NP] [$_{\text{CP}}$na [$_{\text{InflP2}}$[NP] [$_{\text{VP}}$ *V* [NP]]]]]
a. Prosodic structure: ((NP$_i$ na PRO$_i$)φ)ι ((*V*)φ (NP)φ)ι
b. Prosodic structure: ((NP$_i$)φ (na NP)φ)ι ((*V* PRO$_i$)φ)ι

(Deduced from Marfo, 2003:131; with slight modifications of the syntactic structure based on Boadi, 2005:159)

The crucial empirical observation in favor of the phrasing outlined in (28), presented by Marfo (2003, 2004, 2005), is that the verb appears H throughout at the surface, irrespective of tonal specification the element preceding it; see (29) for focused subject (Marfo, 2003:131) and (30) for focused object (Marfo, 2003:131).

(29) Syntactic structure (with lexical tones)

LH L HL LHH

[$_{\text{InflP1}}$ [$_{\text{NP}}$Kofi$_i$] [$_{\text{CP}}$na [$_{\text{InflP2}}$ [$_{\text{NP}}$ɔ$_i$] [$_{\text{VP}}$re-boa [$_{\text{NP}}$Abena]]]]]
proper name PART PRO PROG-help proper name
'It is Kofi who has helped Abena.'

Prosodic structure

L H L H H HH L̥!HH

((Kofi na ɔ)φ)ι ((re-boa)φ (Abena)φ)ι

(Based on Marfo, 2003:131)

Marfo (2003:131) suggests that: "…the phonetic H tone we observe on the verbs…is an inserted H tone…associated with focus constructions, but specifically induced by the L-toned FM. This inserted-H prefers to dock on a constituent at the left-edge of a succeeding *I*…." (FM stands for the focus marker *na* and *I* stands for IP). Boadi (1974:19) also states that "If

[15] Kobele & Torrence (2006:165) further note that: "…it looks somewhat similar to the "clausal" or "event" determiners found in other Kwa languages like Fongbe and in Haitian Creole (Lefebvre 1998). In Asante Twi, this element only seems to occur when a +human DP undergoes A'-extraction.", Kropp Dakubu (2005) suggests that is a defininiteness marker which affirms the reality of the event expressed and Ameka (2010) suggests that its use depends on discourse pragmatic factors which need further investigation.

the sentence is *na*-focused the predicate is raised one step higher tonally, if it was not high originally.” and presents data involving a fronted pronoun; see discussion of chapter 7 section 7.1 for further details. However, it should be noted that insertion of an H tone on the verb does not depend on the occurrence of *na*. Fiedler & Schwarz (2005:122) observe that in “Akan relative constructions … the verb in the relative clause changes its tone pattern in adopting an H tone (cf. Schachter 1973, the so-called “link tone” in Bearth 2002)”. Further research is needed to understand the process of H insertion and its function.

Marfo’s account, outlined in (28)a., predicts that the resumptive pronoun ɔ is phrased together with the fronted subject NP (*Kofi*) and the particle *na*. Furthermore, the first syllable of the proper name *Abena* should show up with an H tone. The figure 12 presents self-recorded data illustrating the realization of the sentence shown in (29), uttered at moderate speech rate. The particle *na* is strongly co-articulated with the resumptive pronoun (see also Dolphyne, 1988:90), which indicates that they may be phrased together as Marfo predicts. The pitch on the resumptive pronoun itself is rising. In line with the prediction (e.g. Boadi, 1974; Marfo, 2003, 2004, 2005), the progressive prefix which is realized with an L tone in canonical word order, see figure 11, exhibits a high pitch and the verb which is underlyingly associated with HL tones appears H toned throughout at the surface; but see figure 88 presented in the discussion of chapter 7 section 7.1 for evidence that the pronoun may retain its tonal specification. The first syllable of the object is also H toned. However, it seems that we are dealing with a toneless nominal prefix in this case which receives its tone via tone spreading.

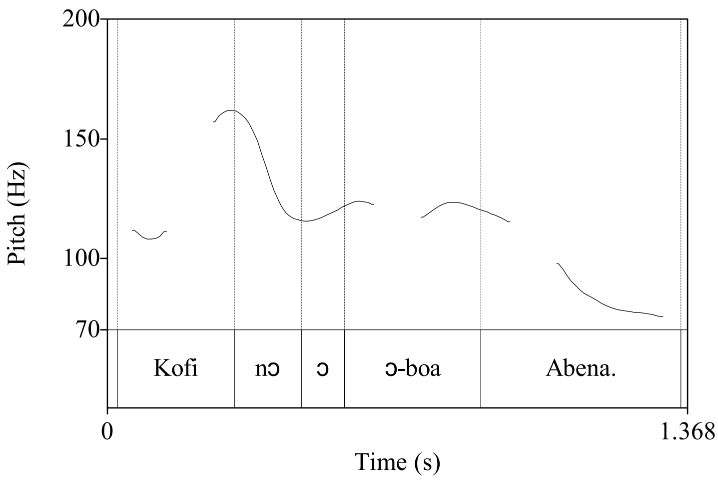


figure 12: Pitch track of a sentence with progressive aspect and fronted subject NP, uttered at moderate speech rate; male speaker.

(30) illustrates the prosodic phrasing with tonal specification for a sentence involving a fronted H toned object and an H toned verb in perfective aspect, schematically represented in (28)b.

(30) Syntactic structure (with lexical tones)

H H L L L H

[$_{\text{InflP1}}$ [$_{\text{NP}}$Esi$_i$] [$_{\text{CP}}$na [$_{\text{InflP2}}$ [$_{\text{NP}}$Yaw] [$_{\text{VP}}$ a-bo [$_{\text{NP}}$no$_i$]]]]]
proper name PART proper name PRF-beat PRO
'It is Esi who Yaw has beaten.'

Prosodic structure

HH L LL H H H

((Esi)φ (na Yaw)φ)ι ((a-bo no)φ)ι)

(Based on Marfo, 2003:131)

The figure 13 presents the pitch track of data displayed in (30), uttered at moderate speech rate. It can be compared to the pitch track in figure 10 showing the structure with canonical word order. The pitch on the perfective prefix exhibits a raising movement towards the H tone on the verb (root), which means that is does not bear an H tone itself, calling into question the association with an IP boundary proposed by Marfo (2003, 2004, 2005).

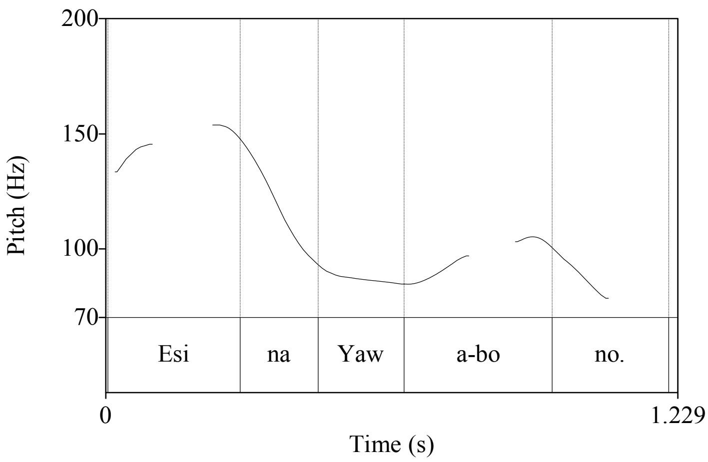


figure 13: Pitch track of a sentence with perfective aspect and fronted object NP, Esi na Yaw abo no. 'It is Esi who Yaw has beaten.' uttered at moderate speech rate; male speaker.

In sum, we have seen once more that the tonal processes applying in the 'focus' construction may not be due to the prosodic structure. The evidence on the role of H tone insertion on the verb (complex) as an indicator of an IP boundary preceding the verb (complex) of the matrix clause was overall not convincing. For the time being, I will assume a recursive prosodic, inspired by Selkirk (2009), to represent the prosodic phrasing of the 'focus' construction. The prosodic structure is displayed in (31). (31)a. represents the prosodic structure of a fronted subject NP and (31)b. for a fronted object NP.

(31) Syntactic structure: [$_{\text{InflP1}}$[NP] [$_{\text{CP}}$na [$_{\text{InflP2}}$[NP] [$_{\text{VP}}$ *V* [NP]]]]]
a. Prosodic structure: ((NP$_i$ (na PRO$_i$)φ)φ ((*V* (NP)φ)φ)ι)ι
b. Prosodic structure: ((NP$_i$ (na)φ)φ ((NP)φ (*V* PRO$_i$)φ)ι)ι

Evidence for the structure in (31)b. is presented in figure 14. The sentence, which was elicited with a picture description task (Genzel & Kügler, 2010), is displayed in (32); see chapter 7 section 7.1 for details.

(32) (À-) mángò$_i$ nà Ànúm̀ tɔ́-ɔ̀ nò$_i$ à-nɔ́pà yì.
NOM-mango PART proper name buy-PST PRO NOM-morning this
'It is a mango that Anum bought this morning.'

The speaker (female) shows a clear pause after the particle *na*, which is taken as an indicator of phrasing. Furthermore, the nominal prefix of the object was dropped in her spontaneous utterance and although the fronted object is inanimate, we can observe a resumptive pronoun in its base position, contrary to Saah's (1998) and Ermisch's (2006) observations. The tone on the verb (root) changed from L to H because it is located in an embedded sentence (Boadi, 1974, Ameka, 2010).

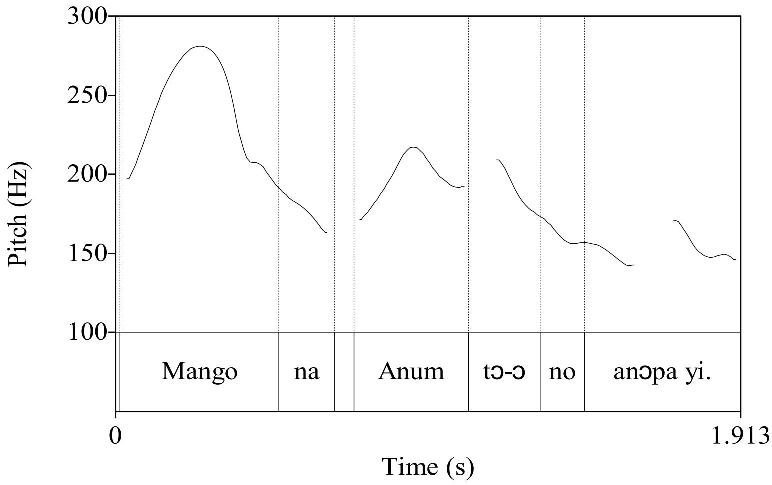


figure 14: Pitch track of a sentence with past tense and fronted object NP, *Mango na Anum tɔɔ no anɔpa yi.* 'It is mango that Anum bought this morning.', spontaneously uttered; female speaker.

In chapter 5 section 5.2 about anticipatory raising, sentence length and prosodic structure the representation in (31) will become relevant. It will be used to describe the phonological representation of complex structures (complementizer clauses). Evidence from pause distribution, co-articulatory segmental effects and preplanning will be provided to empirically support the assumed recursive structure.

1.5 Tonal and non-tonal intonation

This chapter provides background on intonation and prosody. In section 1.5.1 intonation will be defined on the basis of Ladd (1996) with some slight additions. The following section 1.5.2 is concerned with the phonological representation of intonational tones marking sentence type. General phonetic reflexes of intonational morphemes will be outlined and Akan-specific information will be presented which cumulates in the formulation of the main research questions concerning the intonational marking of Yes – No questions in Akan. Section 1.5.3 contains a definition of focus as well as theoretical and empirical background on the matter.

1.5.1 What is intonation?

From a typological point of view, every language is assumed to have intonation (Bolinger, 1962; Hockett, 1963; Gussenhoven, 2004, Grice, 2006). As a working definition, I basically adopt Ladd's (1996:6f.) proposal that "Intonation, …, refers to the use of suprasegmental phonetic features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured* way." with some slight modifications. Ladd restricts the definition of suprasegmental features to F0, intensity and duration. However, research of intonation and its acoustic correlates especially in tone languages has revealed that laryngeal voice quality features, such as glottalization, breathiness, devoicing, and vowel quality should be added to the set of suprasegmental features (e.g. Rialland, 2007, 2009; Hyman & Monaka, 2008). Hyman & Monaka (2008:285) suggest that an intonational feature should be defined independently of its acoustic correlates. The definition should rather be archieved according to the sphere of action of the feature in the grammar: "…a feature which has to be present earlier in the phonology would not be intonational, nor would a particle which has to be present in the syntax.".

Ladd uses the term sentence-level or post-lexical to specify the place in the grammar at which intonational morphemes originate. He (1996:7) states that "…intonation conveys meanings that apply to phrases or utterances as a whole, such as sentence type or speech act, or focus and information structure.". Importantly, this part of the definition excludes lexically determined tones. Ladd basically refers to pragmatic uses of intonation in his statement. Some authors (e.g. Zerbian, 2010; Beckman & Venditti, 2010) remark that: "…intonation invokes other functions, such as mirroring the syntactic structure of an utterance…." (Beckman & Venditti, 2010:603). In the previous section 1.4, this function has been elaborated under the heading of prosody. Ladd (1996:10) argues in favor of a division between prosody and intonation and notes: "…I assume that utterances have a phonological constituent structure (or prosodic structure), and that the prosodic constituents have various phonetic properties, both segmental and suprasegmental. Intonation has no privileged status in signaling prosodic structure….". Others (e.g. Selkirk, 1995; Büring, 2012), however, use prosody as a general term covering intonation and prosodic phrasing (and its correlates including pauses). Grice (2006:778) remarks that the "…distinction between prosody and intonation is rather artificial, since the terms are often used interchangeably – not only in more traditional phonetic models such as the British School (Crystal, 1969; Cruttenden, 1997), but also within phonological models of intonation which embrace the autosegmental-metrical framework.". A very brief introduction to the autosegmental-metrical framework (Pierrehumbert, 1980) will be provided below. I agree with Ladd on the point that intonational events that appear on the (right) edges of IPs do not serve to signal the prosodic structure but serve to convey sentence level pragmatic meanings. However, the distinction between intonation and prosody will get blurred when it comes to the realization of focus, as we will see below.

Turning to the final part of the definition, Ladd proposed that intonation is linguistically structured. He (1996:8) states that: "…intonational features are organised in terms of categorically distinct entities…. They exclude 'paralinguistic' features, in which continuously variable physical parameters (e.g. tempo and loudness) directly signal continuously variable states of the speaker (e.g. degree of involvement or arousal).". Following Ladd, I will describe intonation as an abstract phonological representation in terms of intonational morphemes (Gussenhoven, 2004). As far as possible, linguistic and paralinguistic phenomena will be kept apart in the thesis. Controlled experimental settings will be used to avoid fluctuations in tempo and loudness. However, variations in the degree of involvement or arousal are not totally avertable, especially when it comes to the exploration of the marking of focus whose elicitation involves context. The terms focus, as linguistic notion, and emphasis, as paralinguistic notion, will be kept apart. It should be noted that it is assumed that F0 related paralinguistic and linguistic forms of intonation are underlyingly intertwined (Gussenhoven, 2002, 2004). The relation will be further explored in section 1.6, which deals with the universal aspects of intonational form and meaning. The following section 1.5.2 is concerned with the phonological representation of intonational tones, which signal sentence type, and their phonetic manifestation.

1.5.2 Sentence type

Gussenhoven (2004:45) points out that "Most tone languages will have some form of structural intonation. Frequently even tonally quite dense tone languages have intonational boundary tones, causing questions to end at higher F0 than statements.". The low F0 in statements is represented theoretically as L tone and high(er) F0 in questions as H tone. Following the autosegmental-metrical approach to intonation, proposed by Pierrehumbert (1980) for English, these intonational morphemes are additionally represented with the diacritic %, which marks them as boundary tones. They frequently occur at the edgemost syllable of an IP (e.g. Myers, 2004; but see Michaud (2008) and Rialland & Aborobongui (2010) for intonational tones which are superimposed on lexical tones). Moreover, in the majority of languages the intonational boundary tone(s) appears at the right periphery of the IP (Karvonen, 2008) but boundaries may also be unmarked (0%), i.e. not tonally marked (Grabe, 1998).

The mental construct of a speaker uttering a question in Japanese, which uses H% to signal question meaning, is illustrated in figure 15; reproduced from Pierrehumbert & Beckman (1988:21). The segments are represented on the phoneme tier on the bottom of the figure. The phoneme tier is dominated by the tone tier, on which the lexical and post-lexical reside. Lexical tones associate with their TBUs. Association is illustrated by association lines. The TBU in Japanese is the mora, which is represented on the mora tier. Moras form syllables and are therefore dominated by them in the phonological surface representation. Syllables in turn are dominated by prosodic words. These are resented on the word tier. Prosodic words in Japanese form accentual phrases, which are associated with post-lexical phrase tones. Accentual phrases are dominated by intermediate phrases. The highest node of the prosodic tree represents the utterance, which is formed by two intermediate phrases. The utterance is also associated with post-lexical tones. The final high boundary tone associates to the highest utterance. This representation assures that H% can potentially influence the scaling of utterance internal tones (Pierrehumbert, 2000:31) and expresses that its meaning applies to the whole utterance.

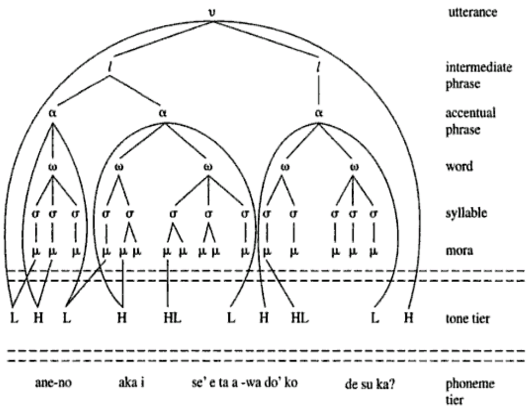


figure 15: Surface representation for the sentence *Ane-no akai se'etaa-wa do'ko desu ka?* 'Where is big sister`s sweater?', reproduced from Pierrehumbert & Beckman (1988:21).

Turning to the phonetic manifestation of boundary tones, it is important to note that there are language-specific differences. However, it is widely acknowledged that (right edge) boundary tones have local F0 effects (e.g. Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990; Rialland, 2007; see Hyman & Monaka, 2008 for further possible interactions between boundary tones and lexical tones in tone languages). Boundary tones are the final tones on the tonal tier i.e. in the tonal string, see figure 15, and are assumed to be phonetically distinguishable from final (lexical) tones by their scaling. Possible scaling effects of H% and L% in a hypothetical language are schematically presented in (33), inspired by Pierrehumbert & Hirschberg (1990:281). The dotted line signals interpolation between the tones (Pierrehumbert, 1980:52). In (33)a. H% is following an H tone. It is assumed that the boundary tone is scaled higher than H. This results in a raising F0 on the final syllable of the utterance. If H% is preceded by an L tone, see (33)b., F0 is also raising but since L is scaled relatively lower than H, the raising movement is presumably steeper than in (33)a. No scaling difference is expected when L% follows L, as illustrated in (33)c. However, the final syllable might be longer (Rialland, 2007). Finally, the F0 movement on the final syllable should be falling if a final H tone is followed by L%, as in (33)d.

(33)

a.	b.	c.	d.
H H%	L H%	L L%	H L%
|	|	|	|
CV	CV	CV	CV

Apart from local scaling effects, it has been observed in a number of languages that boundary tones, especially H%, affect the scaling of earlier utterance internal tones. Tones are realized higher in questions than in statements in e.g. Dutch, Swedish, Russian, Mandarin Chinese, Jita, Kikuyu, Hausa and Chichewa (Myers, 1996, 2004) and/or declination and/or downstep is reduced or suspended in questions compared to statements in e.g. English, Swedish, Russian, Kikuyu, Zulu, Hausa and Chichewa (Myers, 1996, 2004). For some of theses languages, it has been shown that differences in the scaling of utterance internal tones leads to earlier identification of the sentence as questions (Myers, 2004 and van Heuven & Haan, 2000 for Dutch). It has been a matter of debate whether early effects of H% should be modeled phonologically or not. Van Heuven & Haan (2000) suggest that an early adaption of F0 to H% may not be intended but rather a by-product of economization of articulatory effort; but see Inkelas & Leben, 1990 for a phonological account of question intonation in Hausa.

Abstracting away from question intonation which involves high F0, Rialland (2007, 2009) presents an impressive collection of intonational markers of Yes – No questions in African (tone) languages; see chapter 6 for details. Rialland (2007:51) concludes that: "Question prosodies without any high-pitched correlates are not just exceptions…. Markers without any type of high pitched correlates are diverse, including falling intonation (or final Low tones), lengthening, breathy termination, open vowels, polar tones, and cancellation of penultimate lengthening.". Akan is classified as one of these in the citation mentioned languages which use L or L% to distinguish Yes – No questions from statements (Rialland, 2007:58). In a later paper, Rialland (2009) presents Akan as a member of the group of "lax" question prosody languages which utilize at least one of the following phonetic cues to mark sentence type: a falling intonation, lengthening, breathy termination and insertion of an open vowel e.g. [a] as question marker. Crucially, Rialland (2009:929) interprets these phonetic cues as resulting from laryngeal relaxation. In the literature on the intonation of Yes – No questions in Akan, the final falling F0 movement is frequently mentioned (e.g. Dolphyne, 1988; Boadi, 1990; Abakah & Koranteng, 2007). Dolphyne (1988:69) reports on an interesting implementational difference depending on the tone of the final TBU. Final H tones in Yes – No questions are realized as fall to the bottom line, whereas final L tones lack final lowering. Furthermore, Berry & Aidoo (1975) and Dolphyne (1988) observe that Yes – No questions in Akan are uttered in a higher register (tones are realized higher throughout the utterance) than statements. Hyman (2001) tentatively states that Yes – No questions in Akan might show a suspension or reduction of declination/downstep. It is important to note that Rialland (2007) classifies a higher register, the lack of final lowering and reduction/cancelation of declination/downstep as H-pitched question markers. Moreover, Christaller (1875:97) and Boadi (1990:72) observe that the last element of the Yes – No question is lengthened. I will refer to it as extra final lengthening from now on. Reports on non-tonal cues can also be found in the literature. Boadi (1990:72) observes that the final TBU in Yes – No questions is accompanied by extra voicing and glottalization. Dolphyne (1988:50) explicitly points out that the glottal stop does not occur after interrogatives. It is, however, not clear to me, what extra voicing means. It might be connected to a higher amplitude/higher intensity or go together with more pronounced spectral cues/formants (F1/F2 vowel space expansion); see e.g. Zhenglai, Hiroki & Hideki, (2003) and Baumann, Becker, Grice & Mücke (2007).

The literature overview on question intonation in Akan revealed that several different tonal and non-tonal phonetic cues have been observed. Akan seems to combine phonetic features from H and L-pitched question markers (Rialland, 2007), which makes it an interesting test case also from a typological point of view. The following main questions will be addressed in chapter 6:

- What is the intonational morpheme marking Yes – No questions in Akan and which effects are a by-product of the phonetic implementation of it?
- How can Akan be classified along the lines of Rialland's (2007, 2009) typology?

1.5.3 Focus

Information structure refers to the structuring of linguistic information to optimize the transfer of information within the discourse. Krifka (2007) identifies focus, topic and givenness as the basic notions. We will be basically concerned with the concept of focus here, which is assumed to be a universal category. Givenness will be shortly mentioned. For the definition of focus I take as basis an Alternative Semantic approach to focus, based on Rooth (1992), which considers that a focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions (Krifka, 2007:18). Following Féry & Krifka (2008:126), it is assumed that "…the meaning of a question identifies a set of alternative propositions, the answer picks out one of these, and the focus within the answer signals the alternative propositions inherent in the question.". If the sentence in (34) is uttered in an out-of-the-blue context (Ladd, 1980) or as an answer to a very general question of the type "What happened?", it can be considered as a neutral statement i.e. as broad or wide informational focus (Ladd, 1980; Büring, 1997; Krifka, 2007), I will use the term wide informational focus here. In the case of wide informational focus the whole sentence is considered as focused (e.g.

Büring, 2012 but see Selkirk, 2007 for an alternative view) and usually assumed to be accompanied by default prosody (e.g. Weber, Grice & Crocker, 2005).

(34)	Kofi	di-i	a-dua	no.
	proper name	eat-PST	NOM-beans	DET
	'Kofi ate the beans.'			

Informational focus commonly refers to focus in answers to wh-questions. In the data presented here all instances of narrow informational focus, narrow refers a focus which is smaller than the whole utterance (Ladd, 1980), are elicited with a wh-question of the type in (35)a. The focus in the answer to that question, see (35)b., signals the alternative propositions inherent in the question. Since apart from *the beans*, all other elements were previously mentioned in the question, we can be sure that they are not focused but given (e.g. Baumann, 2006). It should be noted that it is possible to front wh-words in Akan as illustrated in (35)a., which resembles the focus construction. Saah (1988) points out that those instances are not due to wh-movement but due to focus. Fronted wh-words can also be optionally preceded by ɛ-yɛ ('it is'). Saah (1988:19) reports for ex-situ wh-construction that: "It seems to me that sentence with clause-initial wh-words/phrases are more emphatic than this in which the wh-word does not occur in initial position.". The interested reader is referred to Saah (1988), Drubig & Schaffar (2001), Drubig (2003), Marfo & Bodomo (2005), Ermisch (2006) and Kobele & Torrence (2006) for further details.

Ermisch (2006) makes the important claim that focus in answers to wh-questions (narrow informational focus) is not (morpho) syntactically marked as illustrated in (35).

(35)	a.	Deɛbɛn	na	Kofi	di-i?
		what	PART	proper name	eat-PST
		'What did Kofi eat?'			
	b.	Kofi	di-i	[a-dua	no]$_F$.
		proper name	eat-PST	NOM-beans	DET
		'Kofi ate the beans.'			

Besides narrow informational focus, the notion of (narrow) corrective focus (Krifka, 2007) will be relevant. The elicitation of narrow corrective focus requires an antecedent in the previous discourse (question) that the focused element of the sentence would correct; see (36)a. for an example. Like in the case of narrow informational focus, all instances of narrow corrective focus used in the thesis involve, apart from the focus, only given information.

(36)	a.	Kofi	di-i	kobi?
		proper name	eat-PST	salty fish
		'Kofi ate salty fish?'		

b.	(ɛ-yɛ)	[a-dua$_i$	no]$_F$	na	Kofi	di-i	no$_i$.
	PRO-be.HAB	NOM-beans	DET	PART	proper name	eat-PST	PRO
	'It is the beans that Kofi ate.'						

Complementary to Saah's (1988) observation about fronted wh-words, Dolphyne (1988:50) notes that the fronted elements are more emphatic. Numerous papers (e.g. Boadi, 1974; Saah, 1988; Drubig, 2003; Marfo & Bodomo, 2005; Fiedler & Schwarz, 2005; Ermisch, 2006; Kobele & Torrence, 2006; Amfo, 2010a) are concerned with the ex-situ or 'focus' construction; see (36)b. for illustration. What is interesting to note is that its use has been reported to be limited to contexts of (narrow) corrective focus (e.g. Marfo & Bodomo, 2005; Ermisch, 2006). Marfo and Bodomo (2005:187) explicitly state that: "… a constituent cannot be contrastively focused *in situ* in Akan.".

It is a subject of discussion whether informational focus and corrective (contrastive) focus are different focus types or whether they reflect different degrees of emphasis (Dolphyne, 1988; Saah, 1988; Hartmann, 2008). Drubig & Schaffar (2001:1079) suppose that "…two types of focus must be distinguished in terms of form and interpretation: presentational focus (or information focus, focus of assertion, rheme, usually wide focus) refers to a constituent which must be interpreted as new, or context-incrementing information, whereas contrastive focus (or identificational focus, operator focus, usually narrow focus) denotes a constituent that identifies a subset within a set of contextually given alternatives."; see also e.g. Selkirk (2007). Other researchers, however, assume that the difference in form corresponds to different degrees of emphasis (e.g. Dolphyne, 1988; Saah, 1988; Hartmann, 2008) or strength of the focus (Féry, 2012). Hartmann (2008:407) notes that: "The grammatical differentiation between the two focus interpretations is a reflex of a *difference in emphasis*. If a speaker uses a contrastive focus she makes an unexpected or important discourse move which may be accompanied by an emphatic rise in the basic frequency or dislocation of the focus constituent to an ex situ position.". Féry (2012:7) assumes the following strength scale: wide informational focus < narrow informational focus < narrow corrective focus, and states that "The probability that a focus is realized with a marked syntactic or prosodic structure increases with the strength of the focus.".

A contribution to the discussion will be made by investigation of the following questions, in chapter 7 section 7.1:

- How frequently is the ex-situ construction used in general and specifically with corrective focus?
- Do our distributional findings speak in favor of a focus type or emphasis analysis?

Departing from the observation (e.g. Ermisch, 2006) that an element in focus, especially narrow informational focus, remains in its base position (in-situ) in Akan, the question arises if and how focus is expressed prosodically. Focus on a constituent has been observed to be accompanied by high(er) pitch, intensity increase and longer duration, signaling higher effort, in many languages, tonal e.g. Mandarin Chinese (Xu, 1999), Beaver (Schwiertz, 2009), Thai (Pan, 2007) and non-tonal e.g. German (Féry & Kügler, 2008), English (Eady & Cooper 1986; Eady, Cooper, Kloouda, Mueller & Lotts, 1986). Tuwuli, a Ghana-Togo-Mountain language, has been observed to mark focus by a raised tonal contour particularly on H toned syllables, higher intensity and compression of the surrounding elements (Harley, 2005 cited in Ameka, 2010). In sum, the distinction between narrow informational and narrow corrective focus seems to be rather gradient with narrow corrective focus being realized with higher F0 than narrow informational focus (see e.g. Baumann, Grice & Steindamm, 2006 for German and Hartmann, 2008 for an overview).

The correlation of high(er) pitch, intensity increase and longer duration and focus has been theoretically formalized as Focus Prominence (Büring, 2010) or Stress-Focus constraint (e.g. Selkirk, 1984, 1995; Truckenbrodt, 1995; Szendröi, 2003) demanding focus to be maximally prominent. Stress is understood as highest metrical prominence on the IP level usually corresponding to the use of a nuclear pitch accent on the focus-marked element in intonation languages like German and English. Other languages, however, employ a different strategy to express focus, which refers to prosodic phrase structure, i.e. an insertion of a phrase (break) before or after a focused constituent (e.g. Frajzyngier, 1989 for Pero; Kidda, 1993; Hartmann, 2007; for Tangale; Karneva, 1990; Downing, Mtenje & Pompino-Marschall, 2004; Downing, 2008 for Chichewa; Karlsson, House, Svantesson & Tayanin, 2007 for Kammu). Boundary insertion has been observed in some Kwa languages spoken in Côte d'Ivoire e.g. Baule (Leben & Ahoua, 2006). It can be detected by non-application of hiatus resolution, interruption of upsweep and interruption of L tone spreading (Leben & Ahoua, 2006). It should, however, be noted that Downing & Pompino-Marschall (to appear) and Leben & Ahoua (2006) observe that boundary insertion is due to emphasis and not due to focus per se. Besides phrasal means, many other languages have been found to not use prosodic means for the expression of focus at all (e.g. Northern Sotho, Zerbian, 2006; Yucatec Maya, Kügler & Skopeteas, 2006, 2007; Kügler, Skopeteas & Verhoeven, 2007; Gussenhoven & Teeuw, 2008; Navajo, McDonough, 2002). Hartmann & Zimmermann (2007) present evidence from production data and perception tests that also Hausa does not use prosodic means for the encoding of focus contra observations expressed by Leben, Inkelas & Cobler (1989) and Inkelas & Leben (1990).

Féry (2012) presents an alternative approach to Focus Prominence. In her Focus Alignment model, the empirical fact that the one to one correspondence of prominence and focus is far from absolute is acknowledged. The Focus Alignment model predicts that prosodic alignment, and not prominence, is the universal reflex of focus: "A focused constituent is preferably aligned prosodically with the right or left edge of a prosodic domain of the size of a prosodic phrase or of an intonation phrase. Languages have different strategies to fulfill alignment, … syntactic movement, cleft-constructions, insertion of a prosodic boundary, and enhancement of existing boundaries. Additionally, morpheme insertion and pitch accent plus deaccenting can also be understood as ways of achieving alignment." (Féry, 2012:1). Under the Focus Alignment approach focus marking is not obligatory, e.g. in cases in which the focus is already aligned per default or when higher ranked grammatical constraints block alignment, and is expected to rather appear with narrow corrective focus.

The prosodic realization of in-situ focus in Akan has not attracted much attention in the literature. The only work I'm aware of is that of Abakah & Koranteng (2007:75f.) who propose that focus/emphasis can be marked prosodically. All examples they show, see (37), come from the Fante dialect of Akan and involve past tense, which is expressed by a grammatical HL tone in Fante (L in Asante). We cannot be sure to what extent Abakah & Koranteng's observation holds for Asante. They state that the effect appears "…especially in the Fante dialect when the emphasis is placed on the agent NP…." (Abakah & Koranteng, 2007:75). The prosodic effect is not located on the focused constituent itself but on the verb. In non-final position, it is realized with an extra H and an extra L, which would correspond to pitch register expansion. Further, they observe that the verb is stressed, which would correspond to a higher intensity. If the verb is final, it undergoes tonal changes. The tones of the verb in neutral condition are presented in (37)a. (HLL) and the tonal change on the verb, which is triggered by focus/emphasis on the subject is presented in (37)b. (HLH).

(37) a. Kòdʑɥó súmà-ì.
proper name send-PST
'Kodwo sent (someone).'

b. [Kòdʑɥó]$_F$ súmà-í.
proper name send-PST
'It was Kodwo who sent (someone).'

(Abakah & Koranteng, 2007:75; glosses were slightly changed)

Abakah & Koranteng (2007) analyze the tonal change on the verb as superimposition of a grammatical HLH tone onto the underlying HL tone on the verb (root). If the verb is followed by an affix, the H tone is mapped onto it. In case of a polysyllabic verb root, the initial TBU receives the H tone and all following TBUs are realized low except the final one which carries the H tone. The description is interesting but rather inconclusive. First, they relate the effect to emphasis than to focus per se and second, it is rather surprising that the effect is not located on the focus-marked element but on the verb. The question arises how a focus on the verb should be signaled. It seems thus that the above reported effect does not reflect a prosodic strategy to mark focus. In chapter 7 section 7.2, I will address the following main questions:

- Is focus prosodically marked?
- Does the prosodic marking of narrow informational focus differ from that of narrow corrective focus?

1.6 Universal intonational meaning

In this chapter, I will outline Gussenhoven's (1999, 2002, 2004) account of ethologically determined, universal meaning of intonation. Universal tendencies in the use of F0 are explained by the proposal of three biological codes: the Production code, also known as respiratory code (Nolan, 2006), to be introduced in section 1.6.1, the Frequency code, which is presented in section 1.6.2, and the Effort code, which is further explained in section 1.6.3. These codes are based on universal interpretations of the variation of three inherent features, which affect the vocal chord vibration rate: phases of the source signal, size of the larynx and articulatory precision. The codes are exploited at the phonetic level to convey particular affective (paralinguistic) and informational (linguistic) meanings. Importantly, the phonetic form which is associated to a specific meaning can become grammaticalized and thus lead to the evolution of intonational morphemes which may mirror the universal form. As Gussenhoven (2004:80) notes "…languages may possess forms-meaning relations in their grammars which go against the universal, biological codes….". Haan (2002:150) suggests that if a language exhibits intonational morphemes with 'unnatural' forms, the phonetic implementation component may compensate for it; see also Gussenhoven (2004:83) and Gussenhoven & Chen (2000). The consideration of universal tendencies of form and function in intonation research is crucial in determining whether certain measurable effects should be regarded as language-specific and therefore analyzed as phonological, or whether they are universal aspects of intonation, which are expressed in the phonetic component in all languages.

1.6.1 The Production code

Air pressure is the energy (source signal) that makes the vocal chords vibrate. The rate of vibration is related on the exhalation phase of breathing. Utterances correlate with breath groups (Lieberman, 1967). Since subglottal air pressure is higher and pitch is higher at the beginning of a breath group/utterance than at the end, pitch lowers gradually. This is referred to as declination (e.g. Collier, 1975; Becker, 1979). The term is most commonly used with regard to European languages and less used in work on African tone languages (Connell, 2002a). Although declination is assumed to be present regardless of the tonal specification of a sentence, the term is most commonly used, in the context of tone languages, to describe the F0 downward trend in sentences consisting of tones of the same phonological entity; see (38)a. for a sentence with only L tones and (38)b. for a sentence with only H tones.

(38) a. Yàẁ fì Àkyèm̀fò.
proper name originate.HAB proper name
‘Yaw comes from Akyemfo.’

b. Ábá twá sérɛ́.
proper name cut.HAB grass
‘Aba cuts grass.’

The figure 16 illustrates the downward trend in F0 (declination) in Akan. The F0 realization of a female speaker of the sentence with only L tones, (38)a., is displayed on the left hand side and the sentence with only H tones, (38)b., on the right hand side.

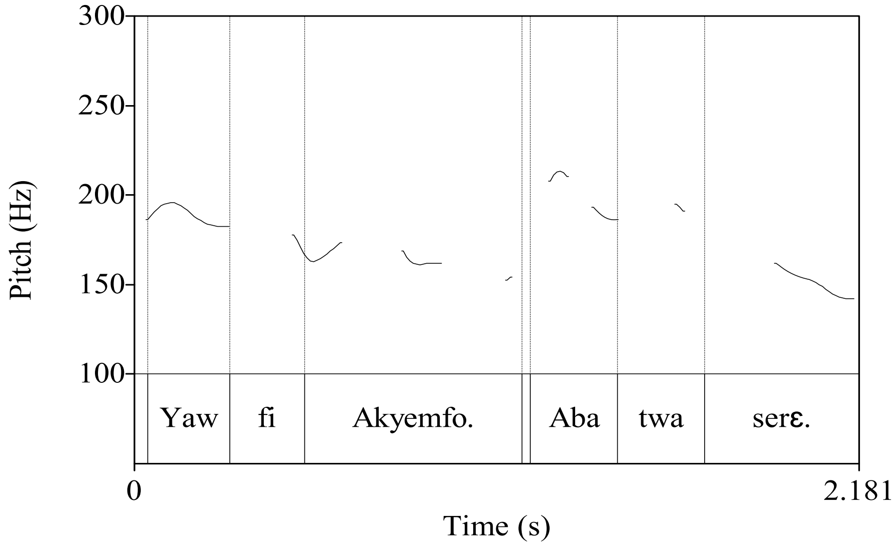


figure 16: Left hand side, pitch track of a sentence with only L tones, *Yaw fi Akyemfo.* ‘Yaw is from Akyemfo.’; Right hand side, pitch track of a sentence with only H tones, *Aba twa serɛ.* ‘Aba cuts grass.’; female speaker.

The correlation of higher beginnings and lower endings is captured by the Production code. According to Gussenhoven (2004:89) the ”… significance of declination … does not lie in it’s slope, …, but rather in the variation at the utterance edges.”. Thus grammaticalization may be limited to initial and final position of an utterance. Utterance initial high pitch is functionally associated with new topics. Utterance final high pitch is associated with continuation, whereas, utterance initial low pitch signals end of topic and utterance final low pitch finality. However, Gussenhoven (2002, 2004) also mentions that declination may be grammaticalized as downstep.

In tone languages, however, an interesting relationship with declination and tonal contrast has been observed. L tones are likely to show declination in general whereas H tones do not decline if they contrast with either M or !H in a language (Hyman, 1975:228). I will take advantage of this interaction to come to a conclusion regarding the existence of !H in the grammar of Akan. Furthermore, it has been found that declination rate interacts with sentence length i.e. the shorter the utterance the steeper/faster the declination slope/rate (e.g. Cooper & Sorensen, 1981 for English; Lindau, 1986 for Hausa; Swerts, Strangert & Heldner, 1996 for Swedish; Connell, 2003, 2004 for Mambila; Shih, 2000, Yuan & Liberman, 2010, for Mandarin Chinese). Ohala, Dunn & Sprouse (2004:163) conclude "…declination is purposeful and thus phonological, not phonetic, since the rate of declination varies with utterance length…. If declination were purely a function of lung volume one might expect much the same rate of F0 declination no matter how long the utterance.". In chapter 4 section 4.3.1, we are going to explore the following questions for Akan:

- Do declarative sentences show declination?
- Does the degree of declination differ as a function of tone?
- Is declination phonological or phonetic?
- How can declination be modeled?

1.6.2 The Frequency code

The vocal chords are smaller and lighter in a smaller larynx, which results in faster vibrations and higher pitch than in a bigger larynx. Following Ohala (1984, 1994), who introduced the term Frequency code, and Gussenhoven (2002, 2004:79) "…the correlation between larynx size and rate of vocal chord vibration is exploited for the expression of power relations.". Hence, the form (higher vs. lower pitch) has been associated with a certain function. Higher pitch is associated with submission and lower pitch of dominance, which is said to be the universal affective meaning of the Frequency code. According to Gussenhoven (2004:82), "'Informational' interpretation of the Frequency code are 'uncertainty' (for higher pitch) vs. 'certainty', and hence 'questioning' vs. 'assertive'.". Gussenhoven (2004:84), citing Hyman (2001), presents Akan as reference example of a language, which suspends downstep in questions as illustration of grammaticalized informational interpretation of the Frequency code. However, as already touched in section 1.5.2, the most frequently mentioned phonetic effect appearing with Yes – No questions in Akan is the terminal falling F0, at least if the final vowel is associated with an H tone (Dolphyne, 1988). Hence, Akan seems to be one of the languages which "…have intonation patterns with meanings that go against Ohala's "frequency code", notably falls signalling interrogatives…" Gussenhoven (1999:302). In chapter 6, the following question will be investigated.

- Does Akan show phonetic compensation for the 'unnatural' intonational question morpheme?

1.6.3 The Effort code

The stiffness of the vocal chords is regulated by muscle control (cricothyroid, vocalis…). Contraction of the muscles leads to stiffer vocal chords, which in turn cause pitch elevation. Lower pitch is accompanied by relaxation of these muscles plus active sternohyoid muscle (Ohala, 1972). Following Gussenhoven (2004), the Effort code associates increased articulatory precision or effort, i.e. the amount of energy (laryngeal and respiratory muscle tension) put into speech production, which results in a wider overall pitch range/register (higher F0) to affective meanings such as surprise or agitation and the informational meaning of emphasis/significance/prominence. According to Gussenhoven (2004:86), the meaning of the Effort code is commonly grammaticalized as focus; see section 1.5.3 for further details. Since it is assumed here that focus is marked by alignment (Féry, 2012), the following question, answered in chapter 7 section 7.2, will be of interest:

- Is emphasis expressed in terms of higher F0?

1.7 Phonetic implementation

The previous section 1.6 has already shown that “A rich communicative world appears to exist in this phonetic implementation component”. (Gussenhoven, 2004:49). Apart from gradient modifications of F0, which are under speaker control and signal attitudinal and informational meaning, one central part of the phonetic implementation is the mapping of abstract phonological representations (lexical & post-lexical tones) onto F0 targets. As Pierrehumbert (2001:140) points out “The phonetic implementation component computes the articulatory and/or acoustic goals ….The phonetic implementation component applies in exactly the same way to all surface phonological representations, and the outcome depends solely on the categories and prosodic structures displayed in those representations.”. Building on work by Bruce (1977), who proposed a phonetic target-interpolation model for Swedish, Pierrehumbert (1980) developed her tonal sequence model in which tones are mapped onto F0 targets by implementation rules which “…apply in a “running window” over the phonological description, mimicking the process whereby speakers transform their abstract intentions for an utterance into actual phonetic outcomes with particular physical characteristics.” (Pierrehumbert, 2000:21).

Since Pierrehumbert’s pitch implementation rules have been developed for English and are thus language-specific, they will not be mentioned further. But her downstep rule that reduces the F0 value of an H tone relative to a preceding H tone (Pierrehumbert 1980:146), further developed by Liberman & Pierrehumbert (1984), is of relevance here. The downstep rule mirrors the pattern observed for terraced level tone languages like Akan; see section 1.3. In English, the downstep rule is triggered by the L tone, which is analyzed as part of a bi-tonal accent (H*+L, L*+H etc.)[16]. Downstep is modeled as an exponential decay towards a speaker-specific baseline value (non-zero asymptote), which will be further described in chapter 4 section 4.3. The initial H tone value serves as reference (input) value for the calculation and is equated with pitch range (Liberman & Pierrehumbert, 1984:191), register in the terminology used here (Clements, 1990); but see Pierrehumbert & Beckman (1988) for a revised version involving an abstract phrase level parameter. Pierrehumbert (2000:29) notes that the model “…has not been…successful in describing the scaling of L tones…. A unified treatment of the phonetics of tone is therefore likely to require innovations….”. The model is attractive because it has been successfully applied to model pitch implementation in tone languages (Myers, 1996 for Chichewa; Shih, 2000 for Mandarin Chinese; Laniran &

[16] H* is a local peak, which aligns with the stressed syllable and is followed by a falling pitch movement, represented by L (Pierrehumbert, 1980; Ladd, 1996).

Clements, 2003 for Yoruba) and it produces F0 values, which can be empirically evaluated. In chapter 4 section 4.3 the following question will be investigated for Akan:

- Which innovations are required to map abstract phonological entities (tones) onto F0 targets?

The implementation algorithm performs from left-to-right under strict locality with one tone look-ahead (phonetic look-ahead). Phonetic look-ahead has to be distinguished from phonological look-ahead which is introduced into the model by Pierrehumbert & Beckman, (1988); see also Truckenbrodt (2004). Phonological look-ahead allows for "… upward search in the tree structure." (Pierrehumbert, 2000:31). Considering the tree structure represented in section 1.5.2 figure 15, the H%, which is associated to the highest node "…would be accessible as an influence on the realization of any tone within the phrase." (Pierrehumbert, 2000:31).

As Rialland (2001) points out "…one of the main issues in phonetic implementation is the size of the processing window used by speakers.". It is assumed that insights about the size of the processing window can be gained by the investigation of choices of pitch register (initial F0) in relation to sentence/phrase length. This is commonly known under the heading of preplanning (Gussenhoven, 2004) or of anticipatory raising (Rialland, 2001). Roughly speaking, speakers tend to start with higher F0 if the sentence is long than if it is short. The interaction of pitch register choice and sentence length does not convey attitudinal and/or informational meaning in the sense of Gussenhoven (2002, 2004), see section 1.6, but may reflect speaker internal spacing[17] considerations; either production or perception based (see Lindblom, 1990 for a general account). Studies, testing the influence of the factor sentence length on the scaling of initial F0 values, have shown mixed results in intonation (Liberman & Pierrehumbert, 1984; Prieto, Shih & Nibert, 1996; Prieto, D'Imperio, Elordieta, Frota & Vigário, 2006; van Heuven, 2004; Thorson, 2007 among others) as well as in tone languages (Snider, 1998; Rialland & Somé, 2000; Rialland, 2001; Laniran & Clements, 2003; Connell, 2003, 2004; Scholz, 2012); see chapter 5 section 5.2 for further details. Languages seem to differ with regards to their preplanning strategy (see Rialland, 2001 for an overview and Laniran & Clements, 2003 for "foresight" and "hindsight" strategies in Yoruba) and with regards to the size (first pP or IP) of the prosodic constituent they take into account (e.g.

[17] The term is borrowed from Ladd (1996:74), "'Tonal space' is intended to echo 'vowel space' and has a similar-informal and essentially metaphorical-role in phonetic description: it reifies the limits on a set of observed values in phonetic data…. Clements's 'tonal level frame' (1979), Pierrehumbert and Beckman's 'transform space' (1988:182), and 'register' (e.g. Poser 1984; Connell and Ladd 1990) are all instantiations of the tonal space idea.".

Krivokapić, 2007; Petrone et al., 2011; Scholz, 2012; Fuchs, Petrone, Krivokapić & Hoole, 2013). Furthermore, the flexibility of the pitch register choice seems to be constrained by the number and nature of tonal contrasts in a language (Connell, 2003, 2004). Latest research (e.g. Petrone, Fuchs & Krivokapić, 2011) suggests that the speaker-specific difference observed in many studies on preplanning can be attributed to the working memory capacity of a speaker.

The issue of anticipatory raising has already been raised for Akan by Christaller (1875:183) and has received further attention by Stewart (1965) who claimed that Akan speakers would anticipate the number/existence of non-automatic downstep in a sentence by raising the initial H tone. In his reply to Stewart, Schachter (1965:32) rejects his claim. Laniran & Clements (2003:206) note that the dispute has not been solved instrumentally so far. In chapter 5, the following main question will be experimentally approached:

- Do Akan speakers employ anticipatory raising?

1.8 Tonal variations in connected speech

After the calculation of articulatory and/or acoustic goals, articulation takes place. In this section, the background on presumably co-articulatory processes that apply to adjacent tones in connected speech will be provided. Three effects are discussed by Chen (2012) in her survey article on the topic: L raising, H raising and H lowering. H lowering (carry-over lowering) will not be further discussed here; see section 1.3 for arguments that the lowering effect on H tones in Akan is not co-articulatory. L raising refers to a local carry-over raising from an H onto the following L (H↗L). It has been attested in the tone languages Mandarin Chinese (Xu, 1997), Thai (Gandour et al, 1994), Yoruba (Laniran & Clements, 2003; Yu, 2009), and Bole (Yu, 2009). Gandour et al. (1994) and Laniran & Clements (2003) observe that a sentence-initial L tone is realized lower than a following L tone in a sentence with alternating LH/HL tones and argue that L raising is not taking place because no H tone precedes the L tone i.e. trigger for L raising is absent. L raising can be experimentally investigated by the paradigm in (39). If L raising is taking place, L2 should be realized systematically higher (↗) than L1.

```
(39)    L1 H ↗L2 H
         |  |  |  |
        CVCVCVCV
```

Another possibility to investigate L raising is presented by Yu (2009). The F0 height of L tones in sentences containing only L tones can be compared to the height of L tones in sentences with alternating tones (HL/LH) as in (40). L raising is taking place if L1 in (40)b. is realized higher (↗) than the L2 in (40)a.

```
(40)   a.    L   L2  L         b.   H↗L1 H
              |   |   |              |   |   |
             CVCVCV                 CVCVCV
```

L raising has not been experimentally investigated in Akan. However, Dolphyne (1994) presents data of sentences with alternating tones in which the first L tone is realized slightly lower than the second. In chapter 3 section 3.1, the following question will be experimentally approached for Akan:

- Are L tones subject to local carry-over raising from a preceding H tone?

An H tone, which is followed by an L tone, has been observed to be realized higher than an H tone which is followed by another H tone (↗HL). H raising, also sometimes called upstep (Gussenhoven, 2004) has been observed in tone languages e.g. Yoruba (Connell & Ladd, 1990; Laniran, 1992; Laniran & Clements, 2003; Yu, 2009), Thai (Gandour et al., 1994), Igbo (Laniran & Gerfen, 1997), Mandarin Chinese (Xu, 1997; Wang & Xu, 2011), Bimoba (Snider, 1998) and Bole (Yu, 2009). Similar effects have also been reported in intonation languages; for English, thus only in initial position referred to as initial raising (Cooper & Sorensen, 1981; Liberman & Pierrehumbert, 1984) and for German only in final position, more precisely sentence-final bi-tonal accents (Féry & Kügler, 2008). H raising has been analyzed as dissimilatory process (Laniran & Clements, 2003). It is usually investigated in sentence initial position by comparing the height of initial H tone in (H1) a sentence with only H tone, as in (41)a., to the height of H1 in a sentence with alternating tones (e.g. HL), as in (41)b. (e.g. Laniran & Clements, 2003; Wang & Xu, 2011). If H raising is taking place, H1 in (41)b. is realized higher (↗) than H1 in (41)a. H raising on later H tones (H3) may be masked by pitch lowering effects (downstep) as in Yoruba (Laniran & Clements, 2003).

```
(41)   a.    H1 H  H3H        b. ↗H1  L ↗H3L
              |  |  |  |            |   |   |  |
             CVCVCVCV             CVCVCVCV
```

To my knowledge, H raising has not been reported for Akan. In chapter 3 section 3.2 the following question will be addressed:

- Are H tones subject to local raising when they are followed by an L tone?

1.9 Summary of the main research questions

The overall goal of the thesis is to identify which factors contribute to the F0 contour of a sentence, at which stage of the speech production process the specific factors come into play, which functions they serve, how they interact and how they can be modeled. The table 2 summarizes the main points of interest and provides a preliminary classification of the factors according to their place in the grammar of Akan and their sphere of action.

<table>
<tr><th></th><th>Local</th><th>Global</th></tr>
<tr><td>Phonetics</td><td>H raising
L raising</td><td>Anticipatory raising</td></tr>
<tr><td rowspan="2">Language-specific phonetics</td><td>Downstep</td><td>Declination</td></tr>
<tr><td colspan="2">Implementation of lexical & post-lexical tones</td></tr>
<tr><td>Phonology</td><td>Post-lexical tones/elements</td><td></td></tr>
</table>

table 2: Summary of the main research entities; ordered by the presumed domain of the entity and their presumed place in the grammar.

The first pillar of the thesis constitutes the investigation of purely phonetic factors, such as H raising, L raising and anticipatory raising. The former two are regarded to be local tonal interactional processes. The latter is assumed to be a global process depending on the length of higher level prosodic constituents. I want to explore whether the factors occur and if yes, to which amount they contribute to the scaling of F0 targets. The results are presented in chapter 3.

The second pillar represents the investigation of the F0 lowering phenomena downstep and declination. It has been argued in section 1.3 that downstep can be regarded as lowering process arising due to local tonal interaction and cannot be regarded as purely co-articulatory (phonetic) in Akan. Declination, on the other hand, is assumed to be a global F0 lowering process whose occurrence has been argued to be language-specific, hence also not universal (phonetic); see section 1.6.1. I want to investigate, with the help of controlled experimental data, whether declination and downstep are distinct processes in Akan. Furthermore, I want to explore whether local and/or global F0 lowering should be represented in the phonological surface representation, if the answer is positive I want to propose such a represenation. The results are presented in chapter 4.

The third pillar of the thesis constitutes the investigation of the form and function of intonational morphemes, which may take on the form of post-lexical tones but need not necessarily so (Hyman & Monaka, 2008). Post-lexical tones/elements are regarded to be represented locally on the tonal string, see figure 15, but may have global effects on the F0 realization. Concretely, I want to show how sentence type and focus are phonetically realized,

which aspects of the phonetic signal should be regarded as phonological and how these can be represented in the phonological surface representation. The chapters 6 and 7 are concerned with the presentation of the expression and the analysis of sentence type and focus in Akan.
Finally, I want to contribute to the understanding of the implementation of lexical and post-lexical tones, not only in a descriptive way but also concretely by adjusting the Liberman & Pierrehumbert (1984) algorithm for Akan. The results are presented in chapter 4.
The materials and methods used to shed light on the research questions are described in the following chapter 2.

2. Chapter
Material and methods

This chapter provides general information relevant for the understanding of the data analysis presented in chapters 3-7. Section 2.1 supplies details on the speakers that participated in the production experiments. The corpus designed to address the research questions, raised in chapter 1, is introduced in section 2.2. The following section 2.3 provides information on the procedure and equipment used to record the participants. Finally, section 2.4 presents details on methods applied to the data to extract F0 values and on statistical analysis used in the thesis.

2.1 The participants

Altogether seventeen speakers participated in the experiments. All seventeen participants were native speakers of Asante Twi and declared English as their second language. The material presented in section 2.2 was recorded in two sessions with different speakers. The first group of participants, which I will refer to as the 'Minot' group, was recorded in Minot, North Dakota in 2010. The 'Minot' group consists of six speakers (one female and five male), all of them were students and their average age was 28. The second group, which I will call the 'Ghana' group, was recorded in Accra and Kumasi in 2009. The 'Ghana' group consists of eleven speakers (six female and five male). Eight were students, one was doing his civil service, one was working in the pharmacy and one was working as university lecturer. The average age was 26. Each speaker was paid a small fee for participation.

2.2 The corpus

In order to address the research questions formulated in chapter 1, I designed a corpus containing of five blocks (A - E). In the following each block will be explained in detail.

Block A consists of sentences with only H and only L tones differing in length. Length is indicated as number of syllables. The materials presented in block A serve as a baseline to investigate L raising and H raising, presented in chapter 3. It is further used to explore the issues of declination (chapter 4 section 4.3, 4.3.1) and anticipatory raising (chapter 5 section 5.2). Furthermore, the length of the bi-moraic final vowel of the test sentence with only H tones (42)c. is analyzed in the discussion of chapter 6. All test sentences presented in block A were recorded with the 'Minot' group. The data set consists of 144 sentences (2 tone x 4 length x 6 speakers 'Minot' group x 3 repetitions). The short sentence containing only H tones is made up of five syllables, see (42)a., the medial one of seven syllables, see (42)b., the long sentence of ten syllables, see (42)c., and the longest one of twelve syllables; see (42)d.

(42) a. Kúkúó-bá nó.
pot-DIM DEF
'The small pot.'

b. Kúkúó-bá pápá nó.
pot-DIM good DEF
'The good small pot.'

c. Kúkúó-bá pápá nó bɔ́ dáá.
pot-DIM good DEF break.HAB daily
'The good small pot breaks everyday.'

d. Kúkúó-bá pápá páá nó bɔ́ dáá.
pot-DIM good very DEF break.HAB daily
'The very good small pot breaks everyday.'

The short sentence containing only L tones is made up of seven syllables, see (43)a., the medial ones of eight syllables, see (43)b., the long sentence of ten syllables, see (43)c., and the longest one of twelve syllables; see (43)d.

(43) a. Yàẁ fì Àkyèm̀fò.
proper name originate.HAB proper name
'Yaw comes from Akyemfo.'

b. Àsàrè fì Àkyèm̀fò.
proper name originate.HAB proper name
'Asare comes from Akyemfo.'

c. Wɔ̀fà Àsàrè fì Àkyèm̀fò.
uncle proper name originate.HAB proper name
‘Uncle Asare comes from Akyemfo.’

d. Wɔ̀fà Àdò Àsàrè fì Àkyèm̀fò.
uncle proper name proper name originate.HAB proper name
‘Uncle Ado Asare comes from Akyemfo.’

Block B comprises sentences with alternating LH and HL tones differing in length. The materials presented in block B are analyzed in chapter 3 and chapter 4 section 4.2 which is concerned with the issue of downstep in initial position. It is furthermore used to investigate the relationship of downstep and declination, see section 4.3, 4.3.3, and to explore anticipatory raising. Parts of the material with alternating LH tones serve as a baseline for the investigation of the prosodic marking of Yes – No questions presented in chapter 6, see block D for further details. All test sentences presented in block B were recorded with the ‘Minot’ group. The data set consist of 126 sentences (2 tone x 3/4 length x 6 speakers ‘Minot’ group x 3 repetitions). The short sentence consisting of alternating LH tones is made up of six syllables, see (44)a., the long one of ten syllables, see (44)b., and the longest one of twelve syllables; see (44)c.

(44) a. Pàpá Kòfí kàsá.
father proper name talk.HAB
‘Father Kofi talks.’

b. Pàpá Kòfí kàsá kyèrɛ́ nè bá.
father proper name talk.HAB point_to.HAB PRO child
‘Father Kofi talks to his child.’

c. Pàpá Kòfí kàsá kyèrɛ́ nè bá bíò.[18]
father proper name talk.HAB point_to.HAB PRO child again
‘Father Kofi talks to his child again.’

The short sentence containing alternating HL tones is made up of five syllables, see (45)a., the medial one of seven syllables, see (45)b., the long sentence of nine syllables, see (45)c., and the longest of eleven syllables; see (45)d.

[18] The alternating LH pattern is not maintained throughout; the final word does not exhibit LH but HL tones.

(45) a. Ánàné bìsá.
proper name ask.HAB
'Anane asks.'

b. Ánàné bìsá sìká.
proper name ask.HAB money
'Anane asks for money.'

c. Ánàné bìsá sìká ɛ̀nɛ́.
proper name ask.HAB money today
'Anane asks for money today.'

d. Ánàné bìsá sìká bɔ̀né ɛ̀nɛ́.
proper name ask.HAB money bad today
'Anane asks for bad money today.'

Block C contains complex sentences consisting of a matrix clause followed by a complementizer clause with only H and only L tones differing in length. The material serves to investigate cues to signal higher level prosodic phrasing and anticipatory raising presented in chapter 5 section 5.2. All test sentences presented in block C were recorded with the 'Minot' group. The data set consists of 108 sentences (2 tone x 3 length x 6 speakers 'Minot' group x 3 repetitions). The material consists of three complementizer clauses containing only H tones and three complementizer clauses containing only L tones. All six were preceded by the same matrix clause. Its segmental and tonal structure is illustrated in (46).

(46) Nàná kà-à ɛ̀-nórà sɛ́ ….
proper name say-PST NOM-yesterday COMP
'Nana said yesterday that ….'

The length of the complementizer clause, measured in number of syllables varied from short over medium to long. The short complementizer clause containing only H tones is made up of five syllables, see (47)a., the medial one consists of seven syllables, see (47)b., and the long one of ten syllables; see (47)c.

(47) a. …kúkúó-bá bɔ́.
pot-DIM break.HAB
'…the small pot breaks.'

b. …kúkúó-bá pápá bɔ́.
pot-DIM good break.HAB
'…the small good pot breaks.'

c. …kúkúó-bá pápá nó bɔ́ dáá.
pot-DIM good DET break.HAB everyday
'…the small good pot breaks everyday.'

The short complementizer clause containing only L tones is made up of seven syllables, see (48)a., the medial one consists of eight syllables, see (48)b., and the long one of ten syllables; see (48)c.

(48) a. …Yàw fì Àkyèmfò.
proper name originate.HAB proper name
'…Yaw comes from Akyemfo.'

b. …Àsàrè fì Àkyèmfò.
proper name originate.HAB proper name
'…Asare comes from Akyemfo.'

c. …wɔ̀fà Àsàrè fì Àkyèmfò.
uncle proper name originate.HAB proper name
'…uncle Asare comes from Akyemfo.'

Block D contains string-identical Yes – No question statement pairs, analyzed in chapter 6, which bears on the issue of the prosodic marking of sentence type. Parts of the materials were recorded with the 'Ghana' group, specifically items 1 (49)a., 2 (49)b., 3 (49)c. and 6 (49)f. Supplementary Yes – No questions to the statements with alternating LH tones presented in block B were recorded with the 'Minot' group. The data set consists of 196 sentences, 88 (4 items x 2 sentence type x 11 speakers 'Ghana' group) from the 'Ghana' group and 108 (3 items x 2 sentence type x 6 speakers 'Minot' group' x 3 repetitions) from the 'Minot' group.

Sentence type may be signaled by a modification of the final element(s). Therefore, emphasis will be put on it during the data description. Besides tonal information, I will additionally provide information about the vowel quality of the final vowel, as a preparation for the measurement of formant structure.

The first Yes – No question/statement pair is displayed in (49)a. I will refer to it as item 1 in the content of chapter 6. The final vowel has a rounded, mid, back quality and carries an H tone. The final vowel of item 2, see (49)b., has a rounded, mid, back quality and is associated with an H tone. Item 3, illustrated in (49)c., exhibits a final H toned rounded, mid, back vowel. The corresponding Yes – No question to the statement with alternating LH tones, see (44)a., is displayed in (49)d. I will refer to it as item 4. The final vowel has a central, unrounded, low quality and carries an H tone. Item 5, see (49)e., is the corresponding Yes – No question to the statement with alternating LH tones, see (44)b., the final vowel has a central, unrounded, low quality and is associated with an H tone. (49)f. represents item 6. The final vowel of the determiner carries an H tone and has a rounded, mid, back quality. Item 7, see (49)g., constitutes the corresponding Yes – No question to the statement with alternating LH tones, see (44)c. The final vowel carries an L tone and has a rounded, mid, back quality.

(49) a. ɔ̀-kɔ́./?
PRO-go.HAB
'He is gone./?'

b. Wó-nìm̀ hɔ́./?
PRO-know.HAB there
'You know the place./?'

c. Kɛ̀sɛ́ɛ̀ nó./?
big DEF
'The big one./?'

d. Pàpá Kòfí kàsá?
father proper name talk.HAB
'Father Kofi talks?'

e. Pàpá Kòfí kàsá kyèrɛ́ nè bá?
father proper name talk.HAB point_out.HAB PRO child
'Father Kofi talks to his child?'

f. Mààmé nó dì-ì à-dùá nó./?
woman DEF eat-PST NOM-beans DEF
'The woman ate the beans./?'

g. Pàpá Kòfí kàsá kyèrɛ́ nè bá bíò?
father proper name talk.HAB point_out.HAB PRO child again
'Father Kofi talks to his child again?'

Block E consists of simple SVOAdv and complex OnaSVAdv structures containing either an L toned object (target word) or an object carrying an H tone. The material was uttered in different contexts to gain insights into the prosodic marking of focus in Akan presented in chapter 7 section 7.2. Note that it was already published as Kügler & Genzel (2012); only parts of the experiments reported in the paper are presented here. All test sentences presented in block E were recorded with the 'Ghana' group. The data set consists of 132 sentences (3 focus conditions x 2 tone on the target word x 2 syntactic construction x 11 speakers 'Ghana' group).

The neutral (wide informational focus) renditions serve as baseline for comparison and were elicited without context. Two sentences with simple SVOAdv structure were used as baseline for the investigation of prosodic effects of focus on the in-situ object, one with an object carrying L tones, see (50)a., and one exhibiting an H tone displayed in (50)b. Two other structures served as baseline for the exploration of prosodic effects of focused objects in non-canonical position (ex-situ). The baseline for the sentence containing the L toned target word (*Ado*) is illustrated in (50)c. and the baseline for the sentence containing the target word *amango* is displayed in (50)d.

(50) a. Àgyèmàǹ bóà-à Àdò à-nɔ̀pá yí.
proper name help-PST proper name NOM-morning this
'Agyeman helped Ado this morning.'

b. Ànúm̀ tɔ̀-ɔ̀ à-máǹgó á-nɔ̀pá yí.
proper name buy-PST NOM-mango NOM-morning this
'Anum bought a mango this morning.'

c. Àdò bóà-à Àgyèmàǹ ànɔ̀pá yí.
proper name help-PST proper name NOM-morning this
'Ado helped Agyeman this morning.'

d. À-máǹgó á-tè fìrì dùá nó só
NOM-mango PRF-exist leave.PST tree DET down
á-nɔ̀pá yí.
NOM- morning this
'A mango has fallen down the tree this morning.'

Narrow informational focus was evoked by wh-questions illustrated in (51)a. and b. asking for the object of the sentences in (50)a. and b. respectively. Narrow corrective focus on the object of the sentences in (50)a. and b. was elicited by Yes – No questions, see (51)c. and d., which included an object that the speakers should correct.

(51) a. Hwàń nà Àgyèmàǹ bóà-à à-nɔ̀pá yí?
whom PART proper name help-PST NOM-morning this
'Whom did Agyeman help this morning?'

b. Déɛ̀bɛ́ǹ nà Ànúm̀ tɔ̀-ɔ̀ à-nɔ̀pá yí?
what PART proper name buy-PST NOM-morning this
'What did Anum buy this morning?'

c. Àgyèmàǹ bóà-à Ànúm̀ à-nɔ̀pá yí?
proper name help-PST proper name NOM-morning this
'Agyeman helped Anum this morning?'

d. Ànúm̀ tɔ̀-ɔ̀ kùbé á-nɔ̀pá yí?
proper name buy-PST coconut NOM-morning this
'Anum bought coconut this morning?'

The answers to the questions presented in (51)a-d. either exhibit the focused object in-situ, see (52)a. and b. or ex-situ, see (52)c. and d. Focus is marked by square brackets a superscripted F.

(52) a. Àgyèmàǹ bóà-à [Àdò]$_F$ à-nɔ̀pá yí.
proper name help-PST proper name NOM-morning this
'Agyeman helped Ado this morning.'

b. Ànúm̀ tɔ̀-ɔ̀ [à-máǹgó]$_F$ á-nɔ̀pá yí.
proper name buy-PST NOM-mango NOM-morning this
'Anum bought a mango this morning.'

c. [Àdò]$_F$ nà Àgyèmàǹ bóà-à à-nɔ̀pá yí.
proper name PART proper name help-PST NOM-morning this
It was Ado who helped Agyeman this morning.

d. [À-máǹgó]$_F$ nà Ànúm̀ tɔ̀-ɔ̀ à-nɔ̀pá yí.
NOM-mango PART proper name buy-PST NOM-morning this
'It is a mango that Anum bought this morning.'

The lab speech data of block E is also available as a semi-spontaneous version obtained from a situation description task, which was conducted with the 'Ghana' group. The data is published as Genzel & Kügler (2010). However, no acoustic analysis was provided in the paper. Parts of the semi-spontaneous data are listed in chapter 7 section 7.1 and are further analyzed in the discussion of section 7.2.

Additional material is used in chapter 3 section 3.2, to explore the locality of H raising. Furthermore, the thesis contains already published data (Genzel & Kügler, 2011) which is presented in chapter 4 section 4.1 to examine the realization of automatic and non-automatic downstep. Both sets were recorded with the 'Minot' group three times each. The material is explained in the particular chapters.

2.3 Recording procedure

All speakers were recorded at a sampling frequency of 44.1 kHz and 32 bit resolution, directly on a laptop (Levono R61) using Audacity (Version 1.2.6) and a headset (Logitech Internet Chat Headset). The headphones were binaural with a frequency spectrum from 20 to 20000 Hz and an acoustic impedance of 32 Ohm with an integrated volume control, so that every participant could adjust the volume. The experiments were carried out in a quiet room and conducted using presentation software. The participants were familiarized with the task through written and oral instructions, followed by four practice trials. The material was presented in a pseudo-randomized order. Items from other unrelated experiments were interspersed as fillers. All test sentences were prepared in Akan orthography with English translation below the target sentence, since the orthography lacks marking for tone, which is the crucial variable here. The participants got the instruction to read the sentence on the slide silently and consult the English translation in case of tonal ambiguities. After this step, they were asked to produce the sentence aloud. Each sentence was presented on a separate slide. The presentation flow was self-paced.

For the recording of the material represented in section 2.2 (block D), sentence type was signaled visually by either question mark (?) or full stop (.). The questions ((51)a. - d.), eliciting the desired information structure on the test sentences presented in section 2.2 (block E), were spoken by a young female Akan speaker and pre-recorded in a quiet room in Berlin using the same equipment. The participants were instructed to first listen to the pre-recorded question and to subsequently read the answer quietly. Then, the question was presented again and the participants had to produce the answer as a response to the question. All instances of narrow corrective focus started with a negation particle *daabi* 'no'. The participants were asked to put a pause after it.

2.4 Data pre-processing and statistical analysis

The syllables of all sentences were labeled by hand in Praat (Boersma & Weenink, 2012) based on visual evaluation of the spectrogram and listening to the sound file. Standard cues for segmental labeling were used (Turk, Nakai, & Sugahara, 2006). The F0 analysis was based on a Hanning window of 0.4 seconds length with a default 10 ms analysis frame. Every pitch object was visually checked for octave jumps and algorithm faults. They were manually corrected.

Tones were labeled for each syllable. The tonal label was set manually in the middle of each of the TBUs, assuming that the mid point is sufficient to reflect the phonetic interpretation of the abstract phonological tone and to avoid microprosodic influences[19] as far as possible. The corresponding F0 values were extracted in Hertz (Hz) using a Praat script. In two cases of tone delay, the actual minimum and maximum were measured.[20] Additional measures and conversions applied to the data and extracted Hz values will be presented in the particular content chapters.

All statistical analyses presented in the thesis were done in R (R Development Core Team, 2011). Mostly linear mixed-effect models were performed. The actual specification of the model is explained in the particular content chapters. A *t*-value greater than 2 is assumed to indicate significance. Additionally, *p*-values (*p*MCMC) were calculated from a MONTE CARLO sampling by Markov chain (Baayen, 2008)[21], a *p*MCMC value smaller than 0.05 is taken as cut-off point for significance. The goodness of fit of the pitch implementation model applied to predict the data, presented in chapter 4 section 4.3, was determined using the square (R^2) of Pearson's correlation coefficient (Shih, 2000). A R^2 (coefficient of determination) near 1 is assumed to indicate a good fit, whereas an R^2 closer to 0 is taken to indicate that the predicted values do not fit the obtained values well. A correlation was calculated to determine the linear relation between the variable F0 and intensity presented in the discussion of chapter 6. A Pearson's *r* of 1 is taken to reflect that the relationship between two variables is perfectly linear, whereas a correlation coefficient close to 0 is taken to reflect that there is no linear relationship between them. A few *t*-test's were conducted mainly supplementary on the data on automatic and non-automatic downstep in Akan from Dolphyne

[19]Generally, high vowels may have intrinsically higher F0 than low vowels; see Connell (2002b) for a survey on African tone languages and references therein. Voiceless fricatives and aspirated stops may have an uplifting effect on F0 and sonorants do rather have the opposite effect; see Yip (2002:31f.) for an overview.

[20] This method was applied to the initial H tone and following L tone on the subject of sentences (45)a.-d. presented in section 2.2 (block B) and to the first H tone on the subject of the material used to investigate anticipatory raising in relation to the presence of non-automatic downstep presented in chapter 5 section 5.1. The material is introduced in chapter 4 section 4.1.

[21] My deepest gratitude goes to Thomas Westkott for his guidance.

(1994) presented in the introduction to section 4.1 chapter 4. One further *t*-test was undertaken in the discussion of chapter 6 to compare the length of a mono-moraic to the length of a bi-moraic vowel, again, a *p*-value smaller than 0.05 is taken as cut-off point for significance. Finally, the statistical analysis (repeated measures ANOVAs with speakers as random factor & post-hoc *t*-tests) conducted on the already published data (Kügler & Genzel, 2012) is reproduced in chapter 7 section 7.2, the interested reader is referred to the paper for further details.

The following chapter 3 bears on the issue of local tonal raising of L and H tones due to neighboring tones of the opposite identity in connected speech. Each section starts with an introduction, which contains a definition of the issue under discussion. After that empirical evidence on the issue from other (tone) languages and Akan, if available, is presented. Out of this information a research hypothesis is generated, which in turn is evaluated on the basis of the material introduced in section 2.2 of this chapter. The analysis of the data is explicated under the heading results and is followed by a discussion of the outcome. I will use expressions such as “L1 is realized higher/lower than L2”. It is important to note that I refer to the surface realization (F0) of these tones. Pitch level, tonal height and F0 will be used synonymously. This procedure is applied to all following chapters.

3. Chapter

The influence of adjacent tones in connected speech

Like segments, which are influenced by neighboring segments (e.g. Farnetani & Recasens, 2010) in connected speech, tones are subject to co-articulation (e.g. Chen, 2012). Rightward carry-over effects and leftward anticipatory co-articulation have been reported. Most studies focus on contour tone languages, such as Mandarin Chinese (e.g. Xu, 1997), Vietnamese (e.g. Brunelle, 2009) and Thai (e.g. Abramson, 1979). However, some works on co-articulation in African level tone languages exist; see Laniran (1992) and Laniran & Clements (2003) for data on Yoruba and Yu (2009) for data on Yoruba and Bole. In this chapter, I will explore local tonal effects in Akan. Section 3.1 is concerned with L raising and section 3.2 with H raising.

It will be shown that the nature of an adjacent tone affects the scaling of a tone in Akan, both H and L tones are affected

3.1 L raising

Introduction: L raising refers to a local carry-over raising from an H tone onto the following L tone (Laniran & Clements, 2003:241). The figure 17 shows that the initial L tone in Yoruba is realized lower than the second and third L tone in sentences with alternating LH tones (black circles) and in sentences with alternating HL tones (white circles). The second L tone is raised by about 5 Hz. Further data presented in their paper revealed a maximal L raising of about 10 Hz.

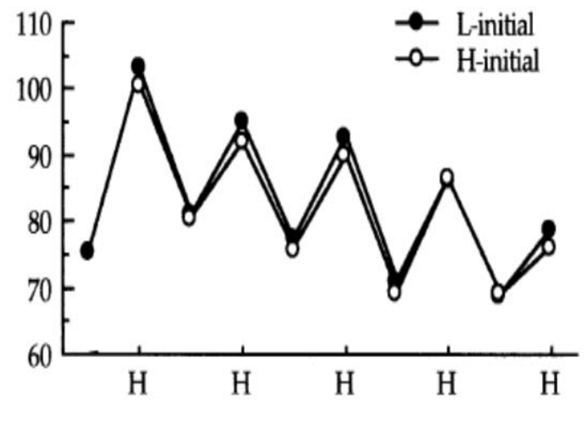


figure 17: L raising and downstep in Yoruba in sentences consisting of alternating HL and alternating LH tones; reproduced from Laniran & Clements (2003:213).

Empirical evidence for the existence of L raising in Akan can be found in Dolphyne (1994). She compared L tones in a sentence with alternating LH tones, as illustrated in (53).

(53) L H1 LH2 L H3 H L H4
| | | | | | | | |
Papa Kofi kɔtɔ n-toma
father proper name go_buy.HAB PL-cloth
'Father Kofi has gone to buy cloth.'
(Dolphyne, 1994:6; association lines and glosses added to the original)

The F0 values for the L tones with differences, referred to pitch drop (Δ), of the sentence in (53), aggregated over five speakers, are presented in table 3. The relevant information is that the initial L tone (L1) is realized slightly lower than the second (L2). The amount of the first pitch drop (L1-L2) is positive and lower than in Yoruba (Laniran & Clements, 2003).

Tone	value	Δ
L1	127 Hz	+3 Hz
L2	130 Hz	15 Hz
L3	115 Hz	13 Hz
L4	102 Hz	

table 3: F0 values for L tones in Hz and differences in pitch drop (Δ), L1-L2, L2-L3, L3-L4, aggregated over speakers (5), for sentences exhibiting downstep; reproduced from Dolphyne (1994:8).

Clements, Michaud & Patin (2011:6f.) point to the importance of distinguishing "…(phonological) assimilation, which is category-changing and phonetic assimilation or co-articulation, which is gradient. A rule by which an L tone acquires a higher contextual variant before H in a language with just two contrastive tone levels, L and H, is not phonological.". In line with Clements et al. (2011), Laniran & Clements (2003) suggest a co-articulatory analysis, following (Gandour et al., 1994), i.e. L raising is absent in initial position because there is no H tone preceding the L tone. Xu (1997:80) also offers an articulatory explanation for L raising: an L tone which is adjacent to an H tone might appear to be raised at the surface because there is evidence to believe that it is more difficult to implement a L tone than a H tone due to phonatory constraints. The production of an L involves the activities of the strap muscles (mainly thyrohyoid, sternohyoid and sternothyroid) (Erickson, 1976), "…these muscles only contribute actively to lowering f_0 when it is to drop below a threshold level, usually near the midrange. This indicates that to reach the lower f_0 range, extra effort by the speaker may be needed. Thus, the lower an f_0 target, the more difficult it is to reach."; see also Erickson, Honda, Hirai & Beckman (1995) for English.

Given the fact that L raising is widely assumed to result out of tonal co-articulation (Gandour et al, 1994; Xu, 1997; Laniran & Clements, 2003) and following Dolphyne's (1994) empirical observation, the following hypothesis will be tested for Akan.

Hypothesis: An L tone flanked by H tones is realized higher on the surface. **Material:** As proposed in chapter 1 section 1.8, the hypothesis can be tested by comparing the height of L tones, at a specific location, in a sentence containing only L tones to the height of L tones, at the same location, in sentences with alternating LH/HL tones. The table 4 provides an overview of the material used to test L raising in Akan. The numbers refer to the detailed material displayed in chapter 2 section 2.2, block A. An example of the short sentence containing only L tones and the short sentence containing alternating HL tones are presented in (54) a. and b., respectively.

Tone/Length	**L**	**LH**	**HL**
7 σ's	(43)a.		(45)a.
10 σ's	(43)c.	(44)b.	
12 σ's	(43)d.	(44)c.	

table 4: Material used to test L raising; method 1.

(54) a. Yàẁ fì Àkyèm̀fò. (43)a.
proper name originate.HAB proper name
'Yaw comes from Akyemfo.'

b. Ánàné bìsá. (45)a.
proper name ask.HAB
'Anane asks.'

Planned comparison: The height of L2 in a sequence of only L tones will be compared to the height of L1, which is flanked by two H tones, as illustrated in (55). Akan is a terraced level tone language, which means that the scaling of a tone depends on the position within the utterance. Hence, it is crucial to compare the height of tones in different sentences only at the same position.

```
(55)   L1 L2 L3  vs.  H  L1 H
        |  |  |       |  |  |
       CVCVCV        CVCVCV
```

Further, the height of L3 in a sequence of only L tones will be compared to the height of L2 which is followed by an H tone, as illustrated in (56).

(56) L1 L2 L3 L4 vs. L1 H L2 H
| | | | | | | |
CVCVCVCV CVCVCVCV

Additionally, L raising can be tested within the same sentence, by comparing the height of an initial L tone in sentences with alternating LH tones to a later (e.g. second) L tone. The table 5 provides an overview of the material for this comparison. An example of the short sentence containing alternating LH tones in presented in (57).

Tone/Length	**LH**
6 σ's	(44)a.
10 σ's	(44)b.
12 σ's	(44)c.

table 5: Material used to test L raising; method 2.

(57) Pàpá Kòfí kàsá. (44)a.
father proper name talk.HAB
'Father Kofi talks.'

The height of L1 in a sentence with alternating LH tones will be compared to the height of L2 in the same sentences, as illustrated in (58).

(58) L1 H L2 H
| | | |
CVCVCVCV

Speaker 5 was excluded from parts of the analysis, since F0 values in the short renditions of LH sentences could not be reliably obtained due to creakiness.

Results: The figure 18 shows the F0 of a sentence with only L tones (black solid line) and the F0 of a sentence with alternating LH tones (black dashed line), produced by a female speaker, aggregated over repetitions. L2 and L4, in sentence with alternating LH tones, are realized higher (raised) in comparison to the L tones at the same position (L3, L7) in sentences with only L tones.

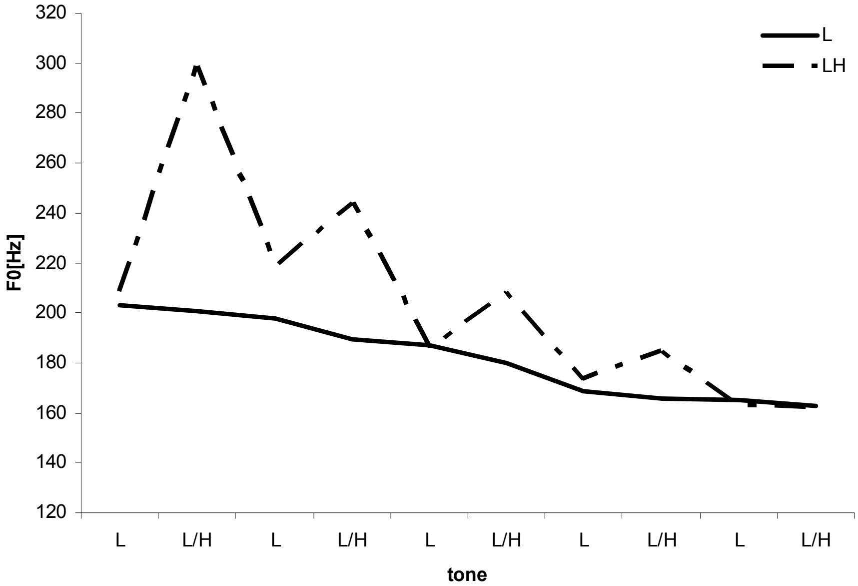


figure 18: Mean F0, aggregated over repetitions, of sentences consisting of alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba.* 'Papa Kofi talks to his child.'; Mean F0, aggregated over repetitions, of sentences containing only L tones (solid line), *Wɔfa Asare fi Akyemfo.* 'Asare come from Akyemfo.'; female speaker.

The figure 19 presents the F0 of a sentence, produced by a female speaker, with alternating LH tones (solid line) and the F0 of a sentence with alternating HL tones (dashed line), aggregated over repetitions. The initial L tone is realized lower than the second L tone in sentences with alternating LH tones. L2 is raised by about 10 Hz in this figure. The raising effect is smaller, but still present, on the first L tone in sentences with alternating HL tones.

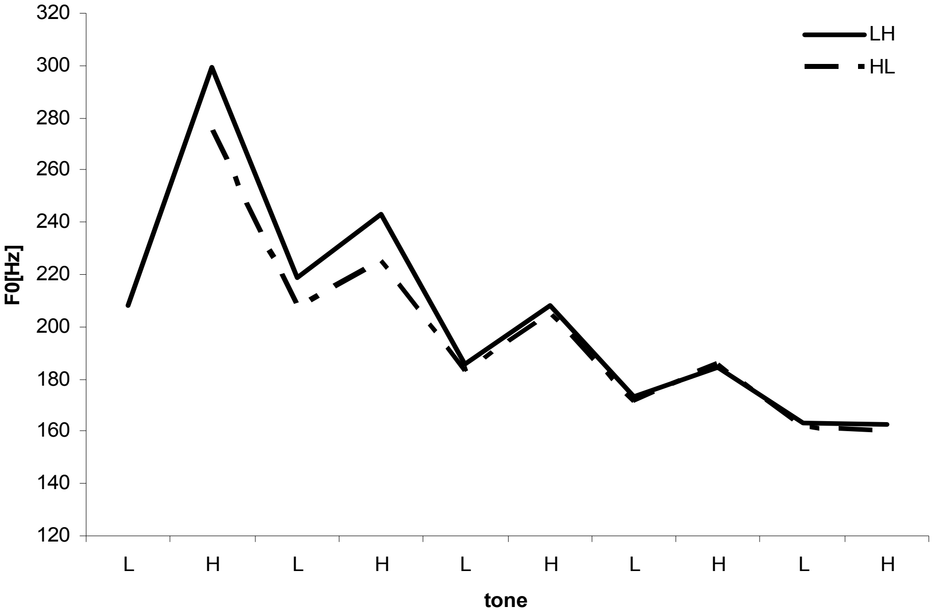


figure 19: Mean F0, aggregated over repetitions, in sentences consisting of alternating LH (solid line), *Papa Kofi kasa kyerɛ ne ba.* 'Papa Kofi talks to his child.'; Mean F0, aggregated over repetitions, in sentences consisting of alternating HL tones (dashed line), *Anane bisa sika ɛnɛ.* 'Anane asks for money.', female speaker.

In what follows, I will present the results for all speakers. In a first step, mean Hz values are presented with standard deviation in parentheses. In a second step, boxplots providing information of the smallest observation, lower and upper quartile, median and largest observation for each speaker is visualized. A statistical analysis concludes the result section. This procedure applies to all following sections and chapters.

Turning to the results obtained with the help of method 1, see (55) and (56), the mean F0 value obtained for L2, in sentences with only L tones, amounts to 133.63 Hz (29) and to 145.01 Hz (30) for L1, in sentences with alternating HL tones. The figure 20, left hand side, presents the data for L2 (only L) and L1 (alternating HL) for each speaker separately, split by tonal configuration and aggregated over repetitions. All speakers, except speaker 2, show a higher value for the L tone (L1), which is flanked by H tones (HL environment).

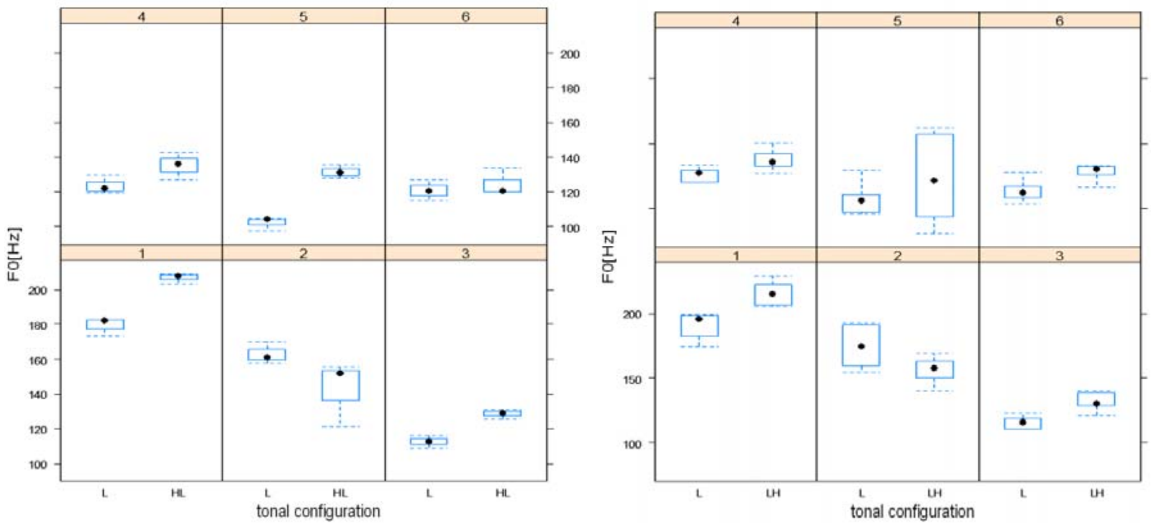


figure 20: Left boxplot: F0 of L2 in sentences with only L tones and L1 in sentences with alternating HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. Right boxplot: F0 of L3 in sentences with only L tones and L2 in sentences with alternating LH tones, aggregated over repetitions and lengths, split by speakers and tonal configuration.

The mean F0 value obtained for L3, in sentences with only L tones, aggregated over speakers, repetitions and lengths, amounts to 138.02 Hz (34) and to 148.52 Hz (35) for L2 in sentences with alternating LH tones. The figure 20, right hand side, provides information for each speaker. All speakers, except speaker 2, show a higher value for L2, which is wedged between two H tones, than for L3 in an only L environment.

A linear mixed effects model was calculated on the height of the L tone. Tonal configuration (only L/alternating LH) was treated as fixed factor. Speakers, repetitions and position of the tone (L1, L2, L3) were considered as random factors. The factor tonal configuration significantly affected the height of the L tone (t = -4.147, pMCMC < 0.001, 108 observations); L tones are realized higher in alternating environments than in sentences containing only L tones.

Turning to the results obtained with the help of method 2, see (58), the F0 means obtained for the first L tone (L1), aggregated over speakers (without speaker 5), repetitions and lengths, amount to 138.36 Hz (33) and to 149.27 Hz (33) for the second L tone (L2) in sentences with alternating LH tones. The figure 21 provides information of the L1 and L2 in sentences with LH alternating tones for all speakers, split by position and aggregated over repetitions. L1 is realized lower than L2, for all speakers.

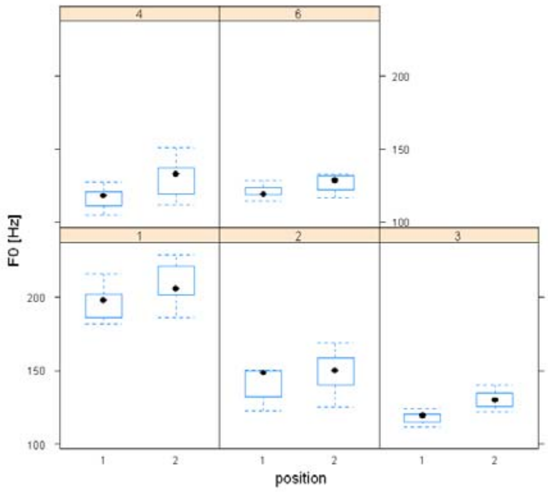


figure 21: Boxplot F0 of L1 and L2, in sentences with alternating LH tones, aggregated over repetitions and lengths, split by speakers and positions.

A linear mixed model was run on the height of the first and second L tone. Position (first vs. second) was included as fixed factor. Repetitions, speakers and sentence lengths were treated as random factors. The factor position showed a significant effect (t = -4.830; pMCMC < 0.001, 87 observations); L2 is realized higher than L1 in sentences with alternating LH tones.

Summary: Like in Thai (Gandour et al., 1994), Yoruba (Laniran & Clements, 2003; Yu, 2009) and Bole (Yu, 2009), an L tone in Akan which is squeezed between two H tones is realized higher on the surface than an L tone which is not preceded by an H tone; see e.g. figure 20. Initial L tones in sentences with alternating LH tones are not subject to L raising; see figure 21. Both methods, the comparisons of L tones in HLH/LHLH environment with L tones in only L environment and the sentence internal comparison of L1 with L2 in a sentence with alternating LH tones, applied here, showed a mean raising of about 10 Hz which is comparable to the amount of raising found in Yoruba (Laniran & Clements, 2003).

3.2 H raising

Introduction: An H tone which is followed by an L tone is realized higher than if it is followed by another H tone (e.g. Connell & Ladd, 1990; Gandour et al. 1994; Xu, 1997; Laniran & Clements, 2003; see Féry & Kügler, 2008 for optional H raising in German). The figure 22 shows that the first H tone in sentences with alternating LH tones (black circles) is realized higher than the first H tone in sentence with only H tones (white circles), in Yoruba. The raising of the initial H tone amounts to 15 Hz on average (Laniran & Clements, 2003:237). Other raising values for other African tone languages available in the literature, 12 Hz for Bole (Yu, 2009) and 20 Hz in Bimoba (Snider, 1998), are comparable to the data from Yoruba.

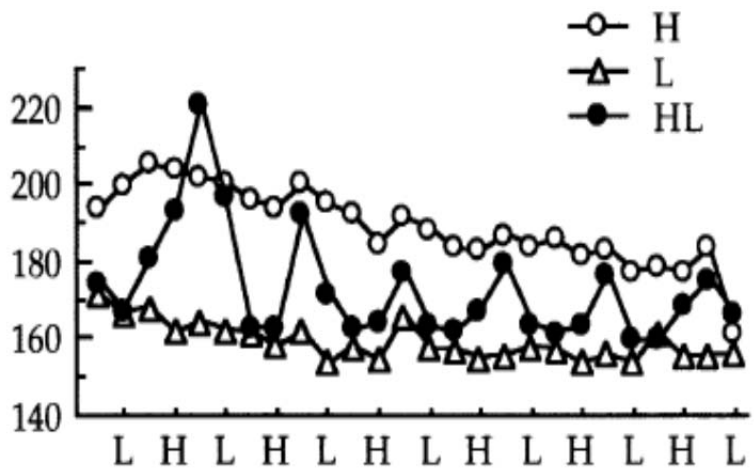


figure 22: Yoruba sentence containing only H tones (white circles), Yoruba sentence containing only L tones (white triangles), Yoruba sentences with alternating L and H tones (black circles), female speaker; reproduced from Laniran & Clements (2003:217).

H raising in Yoruba always applies to H tones which are followed by L tones. The figure 23 shows that the second H tone (b), which is followed by an L tone, is realized higher than the second H tone which is followed by an M tone (a); see Yu (2009) for data showing that H tones in Yoruba are also raised before M.

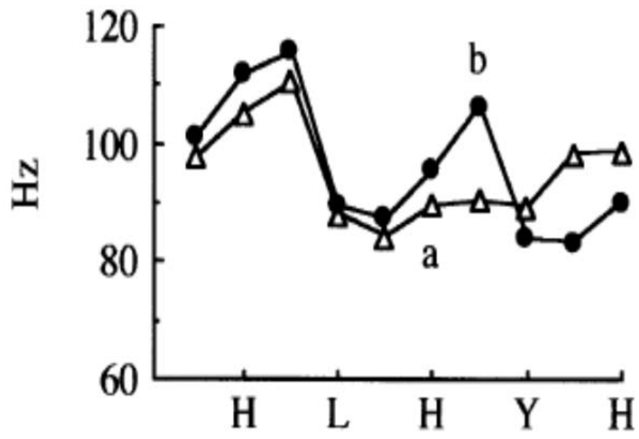


figure 23: F0 of sentences with HLHM (white triangles) and HLHL (black circles) tones, male speaker; reproduced from Laniran & Clements (2003:232).

Laniran & Clements (2003:232) argue that "…although Yoruba speakers implement downstep and H raising by quantitatively different means, their realization strategies ''conspire'' to insure that downstepping H tone will not penetrate the frequency band reserved for M tones.". Thus in Yoruba, which has a three way tonal contrast L, M, H and downstep, H raising serves to ensure perceptual differentiation between a lowered H tone and M. However they (2003:244) did not address "…the question of whether H raising and downstep are purely phonetic principles in Yoruba, or whether in spite of their largely nondistinctive nature they have become phonologized, creating phonologically raised and downstepped tones, as has happened in the evolution of many other African languages (see Snider (1998) for pertinent discussion).". However, Liberman et al. (1992) show that downstep and H raising do not necessarily co-occur. In Igbo, H raising is absent albeit the presence of downstep.

H raising which is anticipatory/dissimilatory in nature has been claimed to reflect controlled articulatory planning, related to the tonal-acoustic/perceptual space (Chen, 2012). Gandour et al. (1993) and Potisuk et al. (1997) propose that H raising takes place to maximize the perceptual distance between H and L. Peng (1997) shows in a perception study that H raising indeed facilitates the perception. It cannot reflect pure co-articulation, since it also appears before floating L tone (Gussenhoven, 2004:108 and references therein).

In sum, the grammatical status of H raising may be language-specific. At the outset it seems to be a reflex of the phonetic system to enhance the tonal contrast between H and L to make room in the tonal/perceptual space for upcoming tonal events, which want to be maximally distinct (Gandour et al., 1993; Potisuk et al., 1997; Chen, 2012). H raising can become grammaticalized (part of the phonology), as mentioned for Yoruba, or can even lead to the emergence of "…new surface-contrastive tone levels…" Laniran & Clements

(2003:205), as in Moba (Rialland, 1983); see Hyman (1993) for a general discussion. Unlike Yoruba, Akan does not exhibit M tones in its grammar. There is thus no reason to expect that H raising is grammaticalized. Although to my knowledge, H raising has not been mentioned or described in the literature on Akan, I act on the unmarked assumption that H raising is present because it is a dissimilatory phonetic process which applies to enhance the perception of tonal contrasts (Gandour et al., 1993; Potisuk et al., 1997; Chen, 2012). Hence, the following hypothesis will be tested.

Hypothesis: An H tone, which is followed by an L tone, is raised.

Material: As proposed in chapter 1 section 1.8, the hypothesis can be tested by comparing the height of an (first) H tone in a sentence with only H tones to the height of an (first) H tone in a sentence with alternating LH/HL tones. The table 6 provides an overview of the material used to test H raising in Akan. An example of the long sentence containing only H tones and of a long sentence containing alternating LH tones is represented in (59)a. and b., respectively. The complete list of materials is introduced in chapter 2, section 2.2, block A.

Tone/Length	**H**	**LH**	**HL**
5 σ's	(42)a.		(45)a.
7 σ's	(42)b.		(45)b.
10 σ's	(42)c.	(44)b.	
12 σ's	(42)d.	(44)c.	

table 6: Material used to test H raising.

(59) a. Kúkúó-bá pápá nó bɔ́ dáá. (42)c.
pot-DIM good DEF break.HAB daily
'The good small pot breaks everyday.'

b. Pàpá Kòfí kàsá kyèrɛ́ nè bá. (44)b.
father proper name talk.HAB point_to.HAB PRO child
'Father Kofi talks to his child.'

Planned comparison: The height of H1 in a sequence of only H tones will be compared to the height of H1 which is followed by an L tone. This is illustrated in (60)a. and (60)b.

```
(60)  a.   H1 H2 H3  vs.  H1 L  H2
            |  |  |        |  |  |
           CVCVCV         CVCVCV
```

```
b.      H1 H2 H3  vs.  L  H1 L
         |   |   |          |   |   |
        CVCVCV        CVCVCV
```

To see whether H raising is still detectable on later H tones, the height of H3 in sentences with only H tones will be compared to the height of H2 in sentences with alternating HL tones, see (61)a., and the height of H4 in sentences with only H tones will be compared to the height of H2 in sentences with alternating LH tones; see (61)b.

```
(61)  a.      H1 H2 H3  vs.  H1  L  H2
               |   |   |          |   |   |
              CVCVCV        CVCVCV

      b.      H1 H2 H3 H4  vs.   L  H1 L   H2
               |   |   |   |          |   |   |   |
              CVCVCVCV         CVCVCVCV
```

To gain insights into the domain of H raising, I recorded additional material. It is illustrated in (62). To answer the question: “How many tones before the appearance of an L tone does H raising apply?” sentences with three, see (62)a., and four H tones, see (62)b., preceding the first L tone were recorded[22].

(62)	a.	Kúkúó	wɔ̀	Yàẁ	Àsàrè	hó.
		pot	be_at_a_place.HAB	proper name	proper name	there
		‘The pot is with Yaw Asare.’				
	b.	Kúkúó-bá	wɔ̀	wɔ̀fà	Àdò	hó.
		pot-DIM	be_at_a_place.HAB	uncle	proper name	there
		‘The small pot is with uncle Ado.’				

Planned comparison: The height of H1 in a sequence of only H tones will be compared to the height of H1 in LHL environment, to the height of H1 in HHHL environment and to the height of H1 in HHHHL environment; see (62). If H raising is strictly local, H1 should differ significantly between H1 in only H and H1 in LHL environment only. Additionally, the height of H3 in a sentence with only H tones will be compared to the height of H3 in HHHL environment and to the height of H3 in HHHHL environment. If H raising is strictly local, H3 should differ significantly between H3 in sentences with only H tones, H3 in which the L tone is one tone away (HHHHL) and H3, which is immediately followed by an L tone. Finally, the

[22] The material was recorded three times each with speakers from the ‘Minot’ group, which results in a data set of 36 sentences.

height of H4 in a sentence with only H tones will be compared to the height of H4, which is immediately followed by an L tone (HHHHL). If H raising is strictly local, H4 is expected to differ significantly between H4 in a sentence with only H tones and H4 which is immediately followed by an L tone.

(63) H1H2 H3 H4 H5 vs. L H1 L H2 L vs. H1 H2 H3 L L vs. H1 H2 H3 H4 L
| | | | | | | | | | | | | | | | | | | |
CVCVCVCVCV CVCVCVCVCV CVCVCVCVCV CVCVCVCVCV

Note that speaker 2 was excluded from the analysis since he shows falsetto voice.

Results: The figure 24 shows the F0 of a sentence with only H tones (solid line) and the F0 of a sentence with alternating HL tones (dashed line), produced by a female speaker; aggregated over repetitions, see (60). H1 in sentences with alternating HL tones is considerably raised in comparison to H1 in sentences with only H tones. The raising of H1 is about 30 Hz in this example. The effect is not detectable on the later H tone(s).

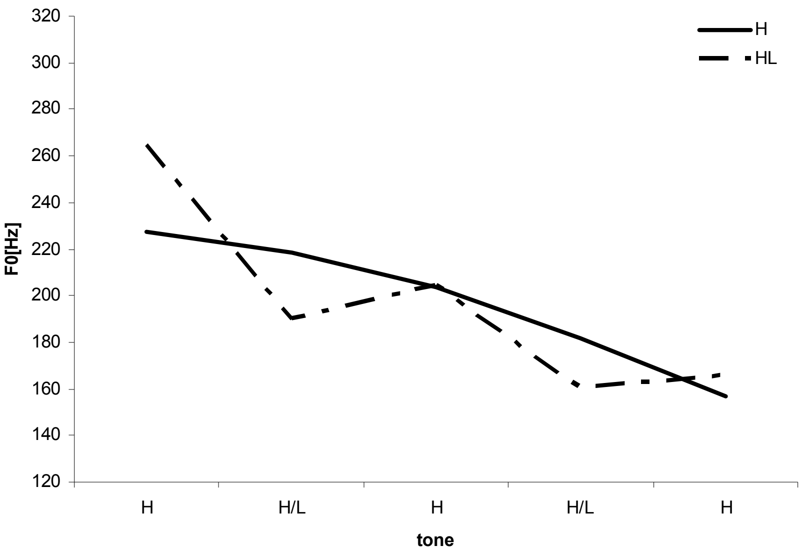


figure 24: H raising and downstep in Akan. Mean F0, aggregated over repetitions, of sentences with alternating HL tones (dashed line), *Anane bisa.* ‘Anane asks.’; Mean F0, aggregated over repetitions, of sentences containing only H tones (solid line), *Kukuoba no.* ‘The small pot.’; female speaker.

The mean F0 value, aggregated over speakers, repetitions and lengths, obtained for H1 in sentences with only H tones amounts to 174.13 Hz (41) and to 185.85 Hz (56) for H1 in sentences with alternating LH/HL tones. The figure 25 provides information of H1 for all speakers, split by tonal configuration (all H vs. alternating LH/HL) and aggregated over repetitions. All speakers, except speaker 5, show a raising effect on initial H tones, which are followed by an L tone.

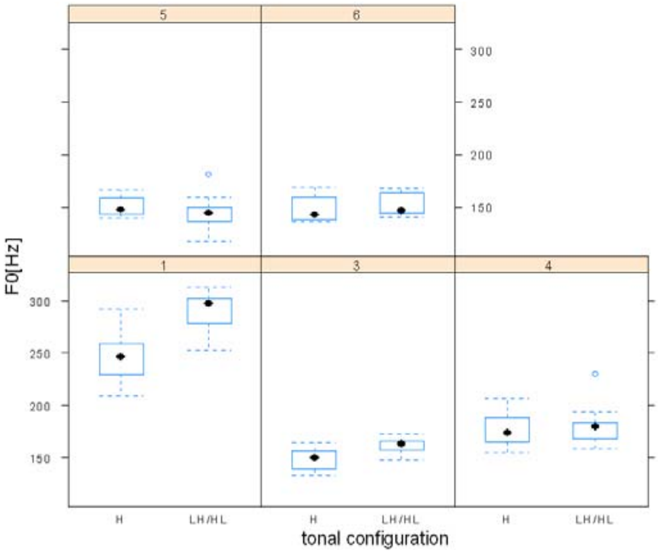


figure 25: Boxplot F0 of H1, in sentences with only H tones and in sentences with alternating LH/HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration.

A linear mixed model was run on the height of the first H tone. Tonal configuration (only H/alternating LH, HL) was included as fixed factor. Repetitions, speakers and sentence lengths were treated as random factors. The factor tonal configuration showed a significant effect (t = -3.981; pMCMC < 0.005; 120 observations); H1 is realized higher in alternating LH/HL environment than in sentences with only H tones.

Turning to the results of later H tones, see (61), the mean F0 value, aggregated over speakers, repetitions and lengths, obtained for H3 and H4 in sentences with only H tones amount to 156.70 Hz (37) and to 156.03 Hz (41) for H2 in sentences with alternating tones (LH/HL). The figure 26 provides information of H3 and H4 in sentences with only H tones and H2 in sentences with alternating LH/HL tones for all speakers, split by tonal configuration and aggregated over repetitions. Speakers 1, 3 and 6 show a slightly higher value for later H tones in alternating sequences and speakers 4 and 5 exhibit a tendency into the opposite direction.

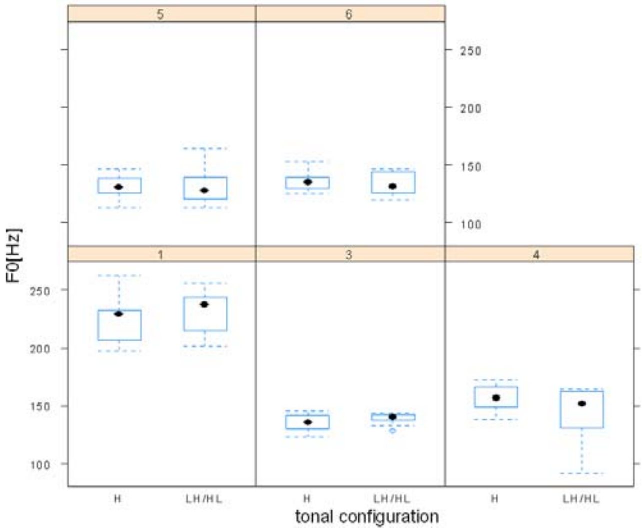


figure 26: Boxplot F0 of H3 & H4, in sentences with only H tones and of H2 in sentences with alternating LH/HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration.

A linear mixed model was run on the height of the later H tones. Tonal configuration (only H/alternating LH, HL) was included as fixed factor. Repetitions, speakers and sentence lengths were treated as random factors. The factor tonal configuration did not show a significant effect.

Turning to the results of the domain of H raising, see (63), the mean F0 values, aggregated over speakers and repetitions, obtained for H1 in sentences with only H tones amount to 184.54 Hz (49), to 191.73 Hz (60) in sentences with alternating LH tones, to 180.57 Hz (46) in sentences with initial HHHL configuration and to 175.59 Hz (42) in sentences with initial HHHHL tones. The figure 27 provides information of H1 for each speaker in sentences with only H tones, in sentences with alternating LH tones, in sentences with initial HHHL tones and in sentences with initial HHHHL tones, split by tonal configuration and aggregated over repetitions. Speakers 1, 3 and 4 show a higher H1 in sentences with alternating LH tones compared to H1 in all other configurations. All remaining speakers do not show a systematic difference for H1 in the different tonal environments.

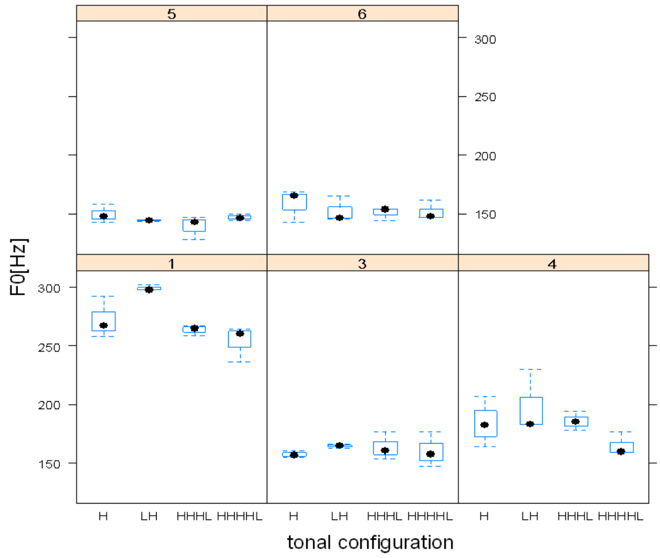


figure 27: Boxplot F0 of H1, in sentences with only H tones, alternating LH tones, initial HHHL tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration.

A linear mixed effects model was calculated on the height of the first H tone. Tonal configuration (only H/alternating LH/HHHL/HHHHL) was treated as fixed factor. Speakers and repetitions were treated as random factors. All comparisons were calculated with H1 in sentences with only H tones as baseline. Only the contrast between the baseline and H1 in LH environment showed a significant effect (t = 2.96, pMCMC < 0.05, 60 observations); H1 which is immediately followed by L (LH environment) is realized higher than H1 in sentences with only H tones and H1 in HHHL and HHHHL contexts.

The mean values of the third H tone, aggregated over speakers and repetitions, obtained for H3 in sentences with only H tones amount to 174.42 Hz (45), to 187.46 (54) in sentences with initial HHHL configuration and to 178.13 Hz (45) in sentences with initial HHHHL tones. The figure 28 provides information of H3 for each speaker in sentences with only H tones, in sentences with initial HHHL tones and in sentences with initial HHHHL tones, split by tonal configuration and aggregated over repetitions. Speakers 1, 3, 4 and 6 show a higher value for H3 which is immediately followed by an L tone in comparison to H3 in only H sentences or H3 in HHHHL environment. The opposite effect can be seen in the data of speaker 5.

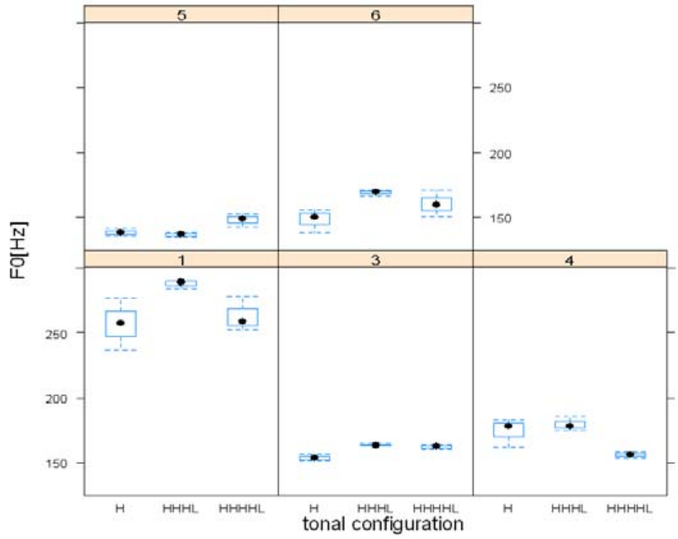


figure 28: Boxplot F0 of H3, in sentences with only H tones, initial HHHL tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration.

A linear mixed effects model was calculated on the height of the third H tone. Tonal configuration (only H/HHHL/HHHHL) was treated as fixed factor. Speakers and repetitions were treated as random factors. All comparisons were calculated with H3 in sentences with only H tones as baseline. Only the contrast between the baseline and H3 which is immediately followed by an L tone (HHHL) approached significance (t = 3.39, pMCMC < 0.067, 45 observations); H3 which is immediately followed by L (HHHL context) is realized higher than H3 in sentences with only H tones and H3 in HHHHL contexts.

The mean F0 values of the fourth H tone, aggregated over speakers and repetitions, obtained for H4 in sentences with only H tones amount to 167.88 Hz (44) and to 178.48 Hz (48) in sentences with initial HHHHL tones. The figure 29 provides information of H4 for each speaker in sentences with only H tones and in sentences with initial HHHHL tones, split by tonal configuration and aggregated over repetitions. Speakers 1, 3, 5 and 6 show a higher value for H4 when it is immediately followed by an L tone in comparison to H4 in all H sentences. The opposite effect can be seen in the data of speaker 4.

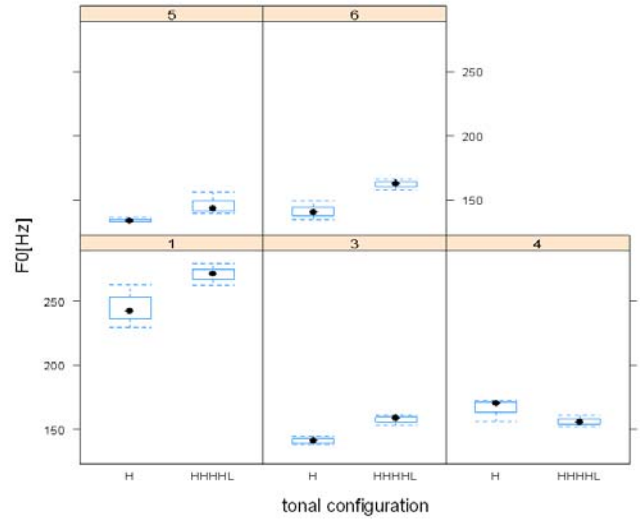


figure 29: Boxplot F0 of H4, in sentences with only H tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration.

A linear mixed effects model was calculated on the height of the fourth H tone. Tonal configuration (only H/HHHHL) was treated as fixed factor. Speakers and repetitions were treated as random factors. The factor tonal configuration approached significance (t = -3.601, pMCMC < 0.0982, 30 observations); H4 which is immediately followed by L (HHHHL context) is realized higher than H4 in sentences with only H tones.

Summary H: Like in many other tone languages, e.g. Bimoba (Snider, 1998), Yoruba (Laniran & Clements, 2003; Yu, 2009) and Bole (Yu, 2009), an H tone in Akan which is followed by an L tone is realized higher on the surface than an H tone which is not (immediately) followed by an L tone; see figure 24. H raising was reliably detected by comparing the first H tone in a sentence with only H tones to the first H tone in sentences with alternating LH/HL tones. The raising amounts to 10 Hz on average which is comparable to the amount of raising found in Yoruba and Bole. The effect vanished on later H tones in alternating tones like in Yoruba (Laniran & Clements, 2003), see figure 26, although it may still be present and masked by other pitch lowering principles to be introduced in chapter 4. The data has further shown that H raising applies to the first H tone which is immediately followed by an L tone also in later positions; see figure 29.

Discussion: In this chapter evidence has been provided that both L and H tones are influenced by an adjacent tone of the opposite identity in Akan. L tones are considerably raised when they occur between two H tones. L tone raising is clearly a co-articulatory effect. Two observations speak in favor of this view. First, the comparison of the height of the L tone (145.01 Hz) on the proper name *Ànánè* to the height of the L tone (148.52 Hz) in the sequence *pàpá kò(fi)*, see figure 20 for illustration, shows that L raising is slightly stronger if phonation is continuous. Second, the L tone in the former case reaches its minimum very late as illustrated in figure 30.

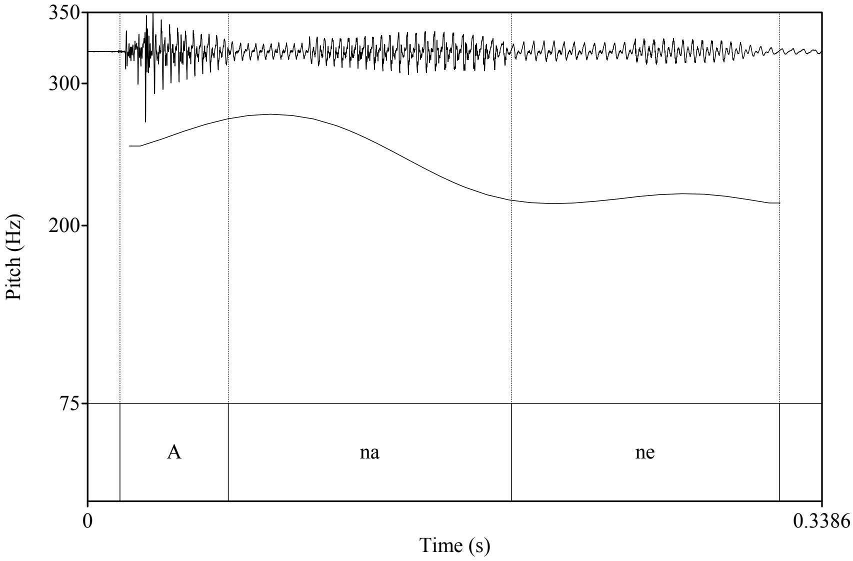


figure 30: Pitch track of the word *Anane*, uttered by the female speaker.

Besides L raising, the data has shown that H tones are raised when they are followed by an L tone and that the effect is strictly local in Akan. However, unlike L raising H raising is at least arguably not due to co-articulation. Most researchers agree that H raising is functional. Hyman (2011:218) notes that it "could thus be useful as a counterforce to processes which lower tones." e.g. downstep. H raising has been shown to be perceptually relevant (Peng, 1997). Furthermore, it has been observed that H raising is also triggered by floating L̥ tones, which are per definition not phonetically realized (Gussenhoven, 2004), whether this is also true for Akan has to be left as subject to future research.

However, as already pointed out in the introduction to this section, H raising might occur without downstep as in Igbo (Liberman et al., 1992). In the following chapter 4 section 4.3, I will show that Akan is another example of a language in which H raising appears independent of downstep since downstep is phonetically similar to declination. It thus seems that H raising in Akan is best analyzed as a local controlled articulatory planning effect (Chen, 2012), which is employed to optimize the tonal space and to facilitate the perception of the contrast between H and L; see chapter 5 for global anticipatory planning effects.

In connected speech, adjacent L and H tones influence each other. H tones have an uplifting effect on L tones and vice versa. H raising and L raising contribute to the scaling of tones in Akan sentences. The amount of both effects is comparable to those found in other African tone languages.

4. Chapter Downtrends

One of the central issues in the study of African tone languages is tone terracing (e.g. Clements, 1979). Akan is well known for its terracing property in the linguistic community since Stewart's (1965) seminal work on automatic and non-automatic downstep. It has been a matter of debate whether to treat the two types of downstep as similar or as distinct processes (e.g. Snider, 1998, 2007 vs. Liberman et al., 1992). For Akan, Dolphyne (1994) has claimed that they are phonetically distinct, whereas Stewart (1965) originally suggested that automatic and non-automatic downstep are similar in their phonetic realization. In section 4.1 of this chapter, the relationship between the two types of downstep will be examined on the basis of controlled material. Furthermore, it has been observed by several researchers that tone terracing does not take place phrase-initially (e.g. Armstrong, 1968; Huang, 1985; Laniran & Clements, 2003). Section 4.2 will be concerned with the issue of downstep in initial position. Moreover, declination is a downtrend which largely lacks controlled experimental investigation in African tone languages, although it has been found to show an interesting interaction with the number of tonal contrasts in a language (Hyman, 1975). Section 4.3.1 draws on the issue of declination in Akan. The relationship between declination and downstep has been an interesting topic. Some researchers equate the two (e.g. Hombert, 1974), some consider declination is a prerequisite of the emergence of downstep in a language (e.g. Yip, 2002; Gussenhoven, 2004) and others argue that downstep is independently phonological (e.g. Snider, 1999). Section 4.3.3 provides insights into the relationship of declination and downstep in Akan.

It will be shown that downstep, which will be reanalysed as declination in this chapter, is the most dominant factor in the determination of the surface F0 in Akan and that is has to be represented phonologically.

4.1 Types of downstep

Introduction: Downstep refers to the lowering of an H tone following an L tone (e.g. Stewart, 1965, Hyman, 1975, Yip, 2002). Classically, a distinction between automatic downstep and non-automatic downstep [23] exists in the literature, which goes back to Stewart (1965). According to Stewart (1965), non-automatic downstep, also referred to as downstep, is defined as H lowering which is triggered by a phonetically not realized L tone; also know as floating L tone (L̥). Connell (2002a:6), referring to Stewart (1965), observes: "One of Stewart's important contributions was the recognition that the lowering of the second H in a HH sequence paralleled the lowering of the second H in a HLH sequence. In the latter case, the lowering was attributed to the influence of the intervening L; in the former it was argued that the lowering was due either to an underlying (floating) L, or one that had been lost historically.". Two observations are of importance here. Firstly, the two types of downstep should be phonetically similar, since they share the same trigger (an L tone); see also Huang (1985). Secondly, all types of non-automatic downstep including lexical downstep[24] should be analyzable as underlying HLH sequence (at least in Akan); see chapter 1 section 1.1 for details. The first observation has been experimentally tested for some African tone languages, though with different results even within the same language. Liberman et al. (1992) report for Igbo that downstep causes a greater degree of lowering than non-automatic downstep, whereas, Laniran (1992) concludes that both types exhibit the same amount of lowering. Snider (1998) examined the phonetic realization of downstep in Bimoba. In Bimoba, associated and floating L tones and M tones cause downstep on H tones. Crucially, no phonetic difference between the two types of downstep was observed. Snider (2007) also concludes for Chumburung that the degree of lowering is the same.

Dolphyne (1994) examined the phonetic realization of automatic downstep and non-automatic downstep in Akan. The analysis of the recordings of five male speakers has shown that non-automatic downstep causes a greater degree of lowering than automatic downstep. The outcome is rather surprising having in mind Stewart's original claim and the

[23] Note that there are also languages in which downstep has to be called non-automatic because there is no intervening L tone on the surface but the presence of downstep seems to be constrained to syntactic/prosodic boundaries e.g. Namwanga (Bickmore, 2000), Tsonga (Lee, 2009) and Tswana (Zerbian & Kügler, 2012). Following Bickmore (2000), the H tone in Namwanga spreads from left to right. When the H tone from the left encounters another underlying H tone, spreading is blocked and downstep appears on the second underlying H tone; but see Paster & Kim (2011) for an alternative proposal.

[24] Lexical downstep, also known as phonologically distinctive downstep, refers to a lowering of the second H tone in a HH context where no trace of an L tone is detectable, neither historically nor due to phonological processes such as hiatus resolution; see e.g. Dolphyne (1988); Obeng (1989) for Akan. However, it has been convincingly argued in Abakah (2000:263f.) that all types of lexical downstep in Akan can be reanalyzed as classic cases of non-automatic downstep involving floating L. They have been analyzed as compounds whose derivation involves the deletion of an L-toned nominal prefix; see chapter 1 section 1.1 for illustration.

experimental results from other African tone languages; see above. Her study will be presented in detail since it is the only experimental contribution on the topic in Akan. Dolphyne's material for testing automatic downstep, displayed again in (64), is made up of alternating LH tones. The first H tone (H1) is preceded by an L tone; however the initial tone is not subject to lowering; see section 4.2. H2, H3 and H4 are subject to "downstep" since they are preceded by an L tone. The H tone following H3 is not separated by an L tone and hence not subject to downstep.

(64)	L H1	LH2	L H3	H L H4
	| |	| |	| |	| | |
	Papa	Kofi	kɔtɔ	n-toma
	father	proper name	go_buy.HAB	PL-cloth
	'Father Kofi has gone to buy cloth.'			

(Dolphyne, 1994:6; association lines and glosses added to the original)

The test sentence involving non-automatic downstep (!) is displayed in (65). Again, H1 should not be affected by any lowering process. H2, H3 and H4 are non-automatically downstepped. According to Dolphyne (1994), downstep on H2 and H3 results from tone spreading and H4 is an instance of lexical downstep; but see Abakah (2000:263f.) for a reanalysis into non-automatic downstep. The verbs in (65), *bɛ̀ká* – 'to come' and *kyɛ̀ré* – 'to tell', are underlyingly associated with an L tone on the first syllable of the verb stem and with an H tone on the second syllable of the verb stem. Unfortunately, Dolphyne (1994) does not provide a detailed analysis of the processes causing non-automatic downstep in her material. A solution for explaining Dolphyne's data can be found in Paster (2010); see also chapter 1 section 1.2 (11). Under Paster's (2010:104) analysis, the perfective comes along with a toneless prefix /a/[25] and a floating grammatical H tone. The prefix receives an H tone via tone spread from the preceding (H toned) syllable, as illustrated in (65). The grammatical H tone associates with the first syllable of the verb root (*bɛ́ká*). Afterwards, "perfect polarity" applies, changing the H tone on the first syllable into an L tone (*bɛ̀ká*). The L tone is dislodged and set afloat due to a "tonal plateauing" process (H tone spread in Dolphyne's terminology) which forces the H tone on the second syllable to spread onto the first syllable ((L̥) *bɛ́ká*). The floating L tone causes downstep on the following H tone; H2 and H3 in (65).

[25] Dolphyne (1988:93) assumes that the perfective prefix is underlyingly L toned.

(65) LH1 !H2H H !H3H H !H4

Kofi a-bɛka a-kyɛre Ama

proper name PRF-come PRF-tell proper name

'Kofi has come to tell Ama (about it).'

(Dolphyne, 1994:5; association line and glosses added to the original)

Dolphyne's test sentences, reproduced in (64) and (65), differ slightly in length (number of syllables). The test sentence involving "downstep" is made up of 9 syllables and the test sentence involving non-automatic downstep of 10 syllables. Additionally, it remains unclear where pitch was exactly measured. Turning to the results of her investigation, table 7 presents the F0 values for the H tones with differences, referred to pitch drop (Δ), for the test sentence exhibiting downstep, see (64), and non-automatic downstep, (65). Note that tables, figures and statistics are self-created supplementary.

Tone	**Automatic downstep**	**Drop**	**Δ**	**Non-automatic downstep**	**Drop**	**Δ**
H1	181 Hz	1	31 Hz	180 Hz	1	25
H2	150 Hz	2	10 Hz	155 Hz	2	20
H3	140 Hz	3	20 Hz	135 Hz	3	20
H4	120 Hz			115 Hz		

table 7: F0 values for H tones in Hz and differences in pitch drop (Δ), H1-H2, H2-H3, H3-H4, aggregated over speakers (5), for sentences exhibiting automatic downstep and non-automatic downstep; based on Dolphyne (1994:5f.).

On the basis of the amount of pitch drop, Dolphyne (1994:6) observes a "…fairly uniform pitch drop between successive downstepped high tones…" and non-uniform drops between H tones in alternating LH tonal sequences (downstep environment). She concludes (1994:10f.) that there is a clear phonetic difference between downstep and non-automatic downstep and that this outcome makes it necessary to distinguish between the three types of downstep in Akan: lexical downstep, derived downstep (non-automatic downstep) and automatic downstep. However, supplementary statistical analysis (paired sample *t*-tests) for the absolute pitch levels reported, see table 7, did not show a significant difference between downstep and non-automatic downstep for any of the three H tones[26]. However, the second pitch drop between H2 and H3 is significantly smaller in the case of automatic downstep ($t(4) = -7.8$, $p = 0.001$). Hence, the difference between automatic and non-automatic downstep, reported by Dolphyne (1994), refers only to the second drop in pitch.

[26] tH1(4) = 0.6, $p = 0.56$; tH2(4) = 1.8, $p = 0.15$; tH3 (4) = 2.3, p = 0.083.

The figure 31 compares the course of F0 for the sentences with alternating LH tones (downstep), see (64), black line, with the course F0 for the sentence involving non-automatic downstep, see (65), grey line. The points of measurement are marked with numbers matching in color, as indicated in the written version.

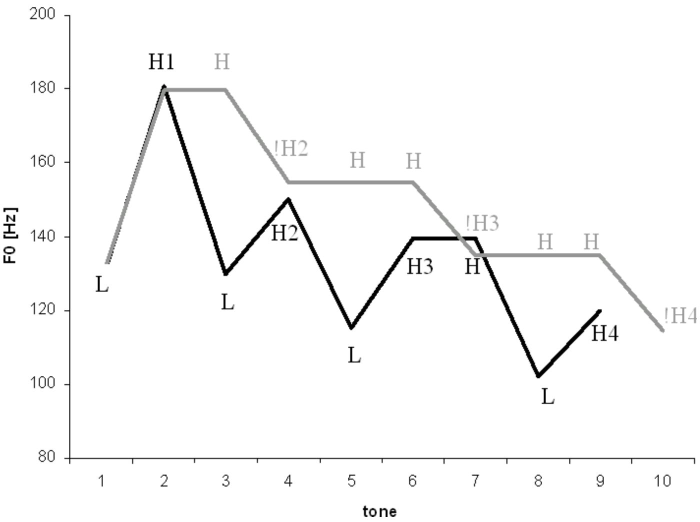


figure 31: Automatic downstep (black solid line) and non-automatic downstep (grey solid line), aggregated over speakers (5); based on Dolphyne (1994:23).

The observed difference (H2 – H3 vs. !H2 - !H3) is unexpected from the perspective that downstep as well as non-automatic downstep are triggered by an intervening L tone and calls for a replication of the experiment with more controlled material, matched for length and number of intervening tones. The following hypothesis, based on Stewart's (1965) original assumption on the similarity between automatic and non-automatic downstep, will be tested.

Hypothesis: The degree of lowering is the same for sentences exhibiting downstep and non-automatic downstep, in absolute terms (pitch level) as well as in relative terms (pitch drop), since the lowering is triggered by an intervening L tone (HLH), which is either present or non-present on the surface.

Material: The hypothesis will be tested with the help of material exhibiting downstep, see (66)a.[27], and non-automatic downstep, see (66)b. The two associative constructions, *Kòfí pàpá* – 'Kofi's father' and *Kòfí (ɔ̀)-dáń* – 'Kofi's house', were embedded into an identical carrier sentence; matched for sentence length and tonal make-up. In (66)b. the nominal prefix *ɔ̀* is deleted for hiatus resolution reasons and the L tone dissociates; see also (12). The L tone is set afloat and causes a lowering (!) on the following H tone.

[27]The material was recorded three times each with speakers from the 'Minot' group, which results in a data set of 36 sentences. The data has been already published as Genzel & Kügler (2011).

(66) a. LH1L L L L H2 L H3 L H H
Afua hunu Kofi papa a-nɔpa yi
proper name see.HAB proper name father NOM-morning this
‘Afua sees Kofi’s father this morning.’

b. LH1L L L L H2 L̥ !H3H4 L H H
Afua hunu Kofi (ɔ)-dan a-nɔpa yi
proper name see.HAB proper name NOM-house NOM-morning this
‘Afua sees Kofi’s house this morning.’

Planned comparison: In a first step, the pitch level will be investigated by comparing the height of H3 in (66)a. to the height of !H3 in (66)b. To exclude any positional influences, H3 in the downstep environment will, additionally, be compared to H4 in the non-automatic downstep environment. In a second step, the amount of pitch drop (H3-H2) in (66)a. will be compared to the amount of pitch drop in (66)b. (H2-!H3/H4).

Results: The figure 32 provides an overview of the results. The course of F0, aggregated over speakers and repetitions, for the sentences exhibiting downstep, represented by the solid line, and non-automatic downstep, represented by the dashed line, is displayed. The curves are nearly lying upon each other. The pitch level of the tones in the critical area (H3 vs. !H3 and/or H3 vs. H4) does not show a systematic difference.

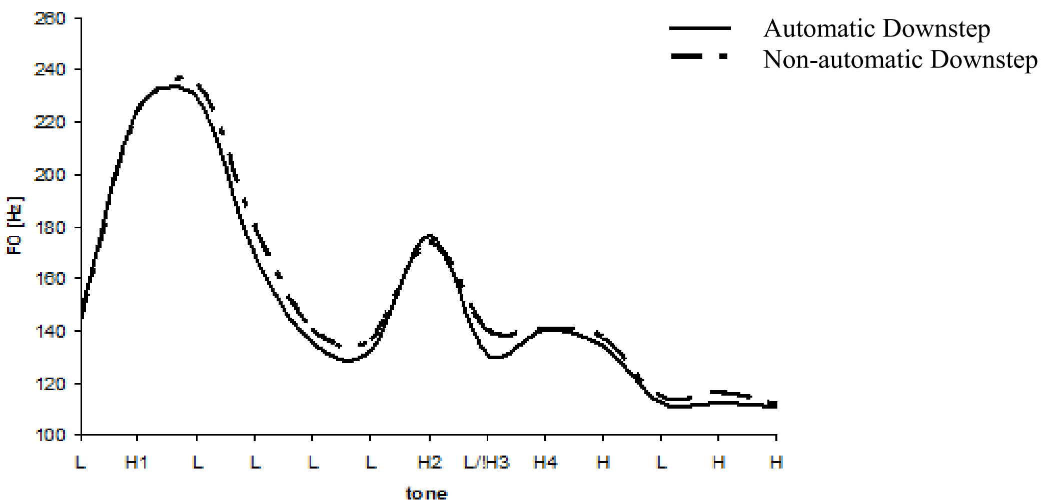


figure 32: Mean F0 of sentences, aggregated over speakers and repetitions, exhibiting automatic downstep (solid line), *Afua hunu Kofi papa anɔpa yi.* ‘Afua saw Kofi’s father this morning.’; Mean F0 of sentences, aggregated over speakers and repetitions, exhibiting non-automatic downstep (dashed line), *Afua hunu Kofi dan anɔpa yi.* ‘Afua saw Kofi’s house this morning.’; n=6.

The preceding H tone (H2) is realized at the same pitch level in both cases. Thus, the amount of pitch drop is not expected to differ as a function of downstep type.

Turning to the pitch level, the mean F0 values obtained for H3, aggregated over speakers and repetitions, amount to 140.28 Hz (35) in sentences containing downstep, and to 140.28 Hz (30) in sentence containing non-automatic downstep. The means for H4, aggregated over speakers and repetitions, are at 140.28 Hz (35) in sentences containing downstep and at 140.67 Hz (26) in sentences involving non-automatic downstep. The following boxplots, figure 33, provide information of H3 and !H3 (left plot) and of H3 in downstep environment and H4 in non-automatic downstep environment (right plot) for all speakers, split by downstep type and aggregated over repetitions. Speakers 2, 3, 5 and 6 do not show a systematic difference as a function of downstep type. Speaker 1 exhibits slightly lower values in the case of downstep, whereas for speaker 4 slightly higher values can be observed.

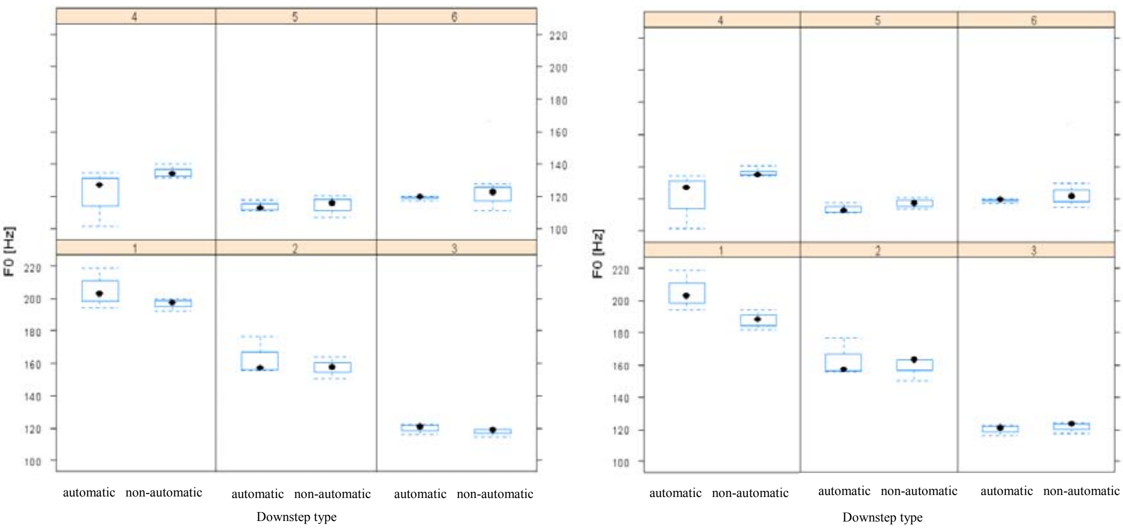


figure 33: Left boxplot: F0 of H3 in sentences containing downstep vs. F0 of !H3 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. Right boxplot: F0 of H3 in sentences containing downstep vs. F0 of H4 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers.

A linear mixed model was run on the height of the critical H tone for each comparison separately. Downstep type was included as fixed factor. Repetitions and speakers were treated as random factors. None of the comparisons yielded a significant result.

Turning to the pitch drop, the means for the pitch drop for H2-(!)H3, aggregated over speakers and repetitions, amount to 30.9 Hz (14) in sentences containing downstep and to 33.71 Hz (13) in sentences containing non-automatic downstep. The means for the pitch drop H2-H3/H4, aggregated over speakers and repetitions, are at 30.9 Hz (14) in sentences containing downstep and at 33.33 Hz (16) in sentences containing non-automatic downstep. The figure 34 provides information of H2-(!)H3 (left plot) and H2-H3/H4 (right plot), aggregated over repetitions and split by downstep type and speakers. Speakers 1, 3, 4 and 5 do not show a systematic difference. A bigger drop in the case of non-automatic downstep can be observed for speakers 2 and 6.

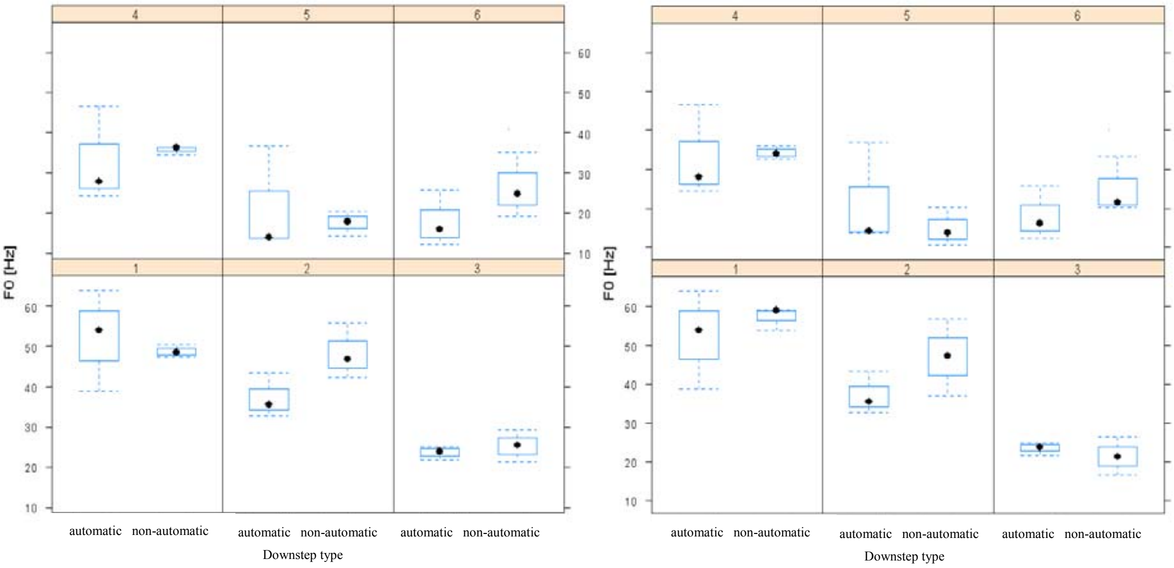


figure 34: Left boxplot: F0 drop between H2-H3 in sentences containing downstep vs. F0 drop between H2-!H3 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. Right boxplot: F0 drop between H2-H3 in sentences containing downstep vs. F0 drop between H2-H4 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers.

A linear mixed model was run on the pitch drop for each comparison separately. Downstep type was included as fixed factor. Repetitions and speakers were treated as random factors. None of the comparisons yielded a significant result.

Summary: As in Igbo (Laniran, 1992), Bimoba (Snider, 1998) and Chumburung (2007), the amount of lowering in Akan triggered by automatic and non-automatic downstep is phonetically similar; see figure 33 and figure 35. No difference between the two was observed by comparing the absolute pitch level of an H tone, which is preceded by an overtly realized L tone, to an H tone, which is preceded by a floating L tone. Further, the pitch drop did not differ systematically between the two types of downstep. This observation is in line with the hypothesis formulated on the basis Stewart's (1965) original claim and contradicts Dolphyne's (1994) observation that non-automatic downstep causes a greater degree of lowering than downstep. Hence, downstep as well as non-automatic downstep is triggered by an intervening L tone between two H tones. The floating L tone remains in the tonal string. Both types of downstep are automatic; see also Abakah (2000, 2002). Automatic and non-automatic downstep cause the same degree of F0 lowering. It thus can be regarded as one factor. The amount of lowering is with a mean of around 30 Hz, three times higher than the amount of raising caused by the phonetic effects L and H raising. Downstep is a major factor in determining the F0 contour of an Akan sentence that contains alterning L and H tones.

4.2 Downstep in initial position

Introduction: It has been observed that initial tones do not undergo lowering from a preceding L tone (Huang, 1980, 1985; Connell & Ladd, 1990; Laniran & Clements, 2003). The figure 17, displayed in chapter 3 section 3.1, illustrates the effect for Yoruba. The initial H tone, which is preceded by an L tone (black circles), is not lower than an initial H tone, which is not preceded by an L tones (white circles).

Armstrong (1968:51) defines downstep as "…the tendency of *non-initial low tones* to pull succeeding high and mid tones downward in pitch.". Huang (1985:213) remarks that the non-application of downstep in initial position is a property which should "…generally held to be true." and accounts for downstep in Akan theoretically by assuming tonal feet; see also Clements (1979, 1983, 1990). A tonal foot (Φ) is a formal object which organizes tones "…in much the same way that a sequence of segments is organized into a sequence of formal objects called syllables" (Huang, 1985:215). Clements (1983) as well as Huang (1985) assume that (67)a. is a possible tonal foot structure, whereas (67)b. is not. According to Clements (1983), a new tonal foot is started between each /LH/ sequence. The idea behind this approach is that H tones are more prominent than L tones (de Lacy, 1999), i.e. H tones serve as heads. Further, tonal feet want to be left-headed (Yip, 2002:153) "…a tonal upturn signals an increase in prominence, and thus a new foot boundary.".

(67) a. Φ b. *Φ

H L LH

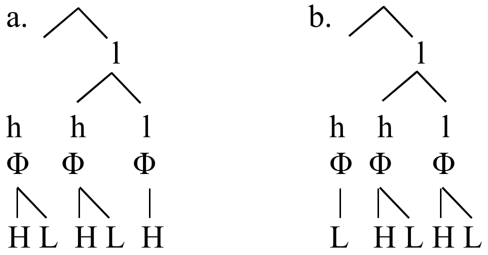
Following Clements (1983) and Huang (1985), all remaining tones are gathered into a tonal foot. Hence, an L tone preceding an H tone is parsed into a single tonal foot, as illustrated in (68).

(68) Φ Φ

L HL

The tonal feet are gathered into trees. According to Clements (1983), feet are grouped into right branching trees labeled with register tones, h and l. The tonal sequences with alternating LH and HL tones, which are of interest here, exhibit the foot structure illustrated in (69), reproduced from Huang (1985:221). Register tones are represented as capital letters in Huang's work. I replaced them with small letters for sake of clarity, following Clements (1983).

(69)



Tonal feet serve as domain for pitch register value assignment and are assumed to be the basic pitch bearing units (Huang, 1985:215). Pitch register is defined as a frequency band internal to the speaker's range, "which determines the highest and lowest frequency within which tones can be realized at any given point in the utterance." (Clements, 1990:59). It "…sets the range in within which contrasting tones are defined." (Huang, 1985:214). Tonal feet, which are dominated by an l register tone, like the second and third foot in (69), are realized one step lower in the register than the first tonal foot, which is immediately dominated by an h register tone. The third tonal foot is realized one step lower than the second. It is immediately dominated by an l register tone. This iterative application leads to the generation of the terracing surface pattern.

Although, the first H tone in (69)b. is dominated by a register l tone, it is not lowered. Huang (1985:222) remarks that "…the amount of tonal contrast is usually greater than the amount of contrast resulting from terracing, the H tone that occurs in the second foot is still higher than the L occurring in the first foot." i.e. the influence of terracing is outranked by the desire of the H tones to be maximally distinct from the preceding L tone. The pitch values of the tone(s) in initial position belonging to the first foot are seen as reference values (Huang, 1980; 1985). The first tonal foot in (69)a. contains both tonal entities. Hence, the following tones in the second tonal foot are scaled in relation to them. As a consequence, the second H tone is realized lower in relation to the first. The first H tone in (69)b. is the first in the phrase, and thus, an H tone default value is assigned to it. The non-application of downstep in initial position does not directly follow from the foot structure but crucially relies on the idea of initial tones as frame-setters or reference values[28].

The concept of tonal feet is problematic for Akan since there are, to my knowledge, no phonological processes which have been shown to have the tonal foot as a domain. Generally,

---

[28] Note that Huang (1980:266) assumed earlier that "Our theory allows the occurrence of an underlying initial downstep, but neutralizes it with a non-downstep with respect to the surface output. Such a situation is perfectly reasonable. Since a downstep established a ceiling, so just like a non-downstep H it is also the highest in pitch in an entire phrase.".

the grouping in (69) resembles the construction of metrical feet into a strong-weak opposition (Liberman & Prince, 1977) in languages with stress like e.g. English (Ladd, 1996). The notion of the metrical foot is intertwined with the issue of rhythm. Halliday (1985) describes English as (metrical) foot-timed language based on the observation that "There is a tendency to equalize the duration of each foot…." (Clark & Yallop, 1990:287), i.e. unstressed/weak syllables in a metrical foot tend to be reduced (e.g. shorter, more co-articulation), whereas stressed/strong syllables tend to be strengthened (longer, less co-articulation). Anderson (2009, 2011), who studies rhythm in Akan, points out that prominence relations (strong/weak; stressed/unstressed) are problematic: "… register tone languages do not usually exhibit stress, nor is there much evidence of phonological rules applying to the metrical structure…." (Anderson 2011:1). The question is whether there is empirical (rhythmic) evidence for tonal feet which resembles the pattern found for metrical feet in e.g. English. Anderson (2011) investigates rhythm in sentences with alternating LH/HL tones and in sentences with only H and only L tones. He suggests the rhythmic pattern in (70), based on his previous study (Anderson, 2009). Prominent syllables are marked by a beat (x), glosses and feet structure are added to the original version (Anderson, 2011:16). The structure for the HL sentences resembles the structure proposed by Clements and Huang. However, the LH structure differs in two respects, first L tones are parsed together with the following H tones and the initial L tone does not constitute its own foot.

(70) LH
x x x x
(Kòfí)$_{\Phi}$ (Dòkú)$_{\Phi}$ (bìsá)$_{\Phi}$ (mè)$_{\Phi}$.
proper name proper name ask.HAB PRO
'Kofi Doku asks me.'

HL
x x x x
(Wá-bà)$_{\Phi}$ (wá-bɔ̀)$_{\Phi}$ (Kwábì)$_{\Phi}$ (bí)$_{\Phi}$.
PRO-come.HAB PRO-hit.HAB proper name SPEC[29]
'You come and hit a certain Kwabi.'

Anderson (2011) found that syllables in sentences with alternating tones, uttered under time pressure, tend to shorten by a similar magnitude. Hence, no evidence for a greater amount of shortening in syllables carrying an L tone, which would correspond to a weak node, could be observed. Phrases with HL or LH tones "…were always alternating prominent …." (Anderson, 2011:22), as illustrated in (70). This tells us that the evidence from rhythm research does not

[29] Following Arkoh (2011:10), I glossed *bi* is a marker of specificity.

support the assumption of an additional prosodic entity like the tonal foot. An alternative way of analysing the non-occurance of initial downstep is proposed by Liberman & Pierrehumbert (1984). Furthermore, the idea that initial tones serve as reference values, mentioned above, is incorporated in the pitch implementation algorithm proposed for modeling downstep in English by Liberman & Pierrehumbert (1984:186), reproduced in (71). Their model will be explained in detail in section 4.3. For now, it is important to note that the F0 of a certain tone in a sentence e.g. H2 in a sentence with alternating LH/HL tones can be predicted from the value of the preceding tone (X_i), e.g. initial H tone (H1). The value of H1 is given and depends on the pitch register choice of the speaker. Thus, an initial H tone can not be subject to downstep since the lowering is relational. This model works without the postulation of additional prosodic constituents, e.g. tonal feet, and avoids the postulation of post-lexical rules, such as the grouping of tonal feet into right-branching trees.

(71) $X_{i+1}\text{-}r = s^*(X_i\text{-}r)$

Interestingly, their approach to downstep has been successfully used in tone languages such as Chichewa (Myers, 1996) and Yoruba (Laniran & Clements, 2003), suggesting that this approach is applicable in typologically different languages.

The introduction has shown that the non-application of downstep in sentence initial position is an inspiring topic. It has stimulated the development of the concept of tonal feet which I introduced and discussed, with the result that the empirical basis supporting it is at least questionable. The most unmarked position outlined, was that initial H tones are not lowered because lowering is relational and first occurrences of tones are realized with a default value, which depends on the pitch register choice of the speaker. Although the non-application of downstep in sentence's initial position can be regarded as one of the key arguments for not analysing downstep in Akan as a co-articulatory phonetic effect, empirical evidence for it is absent. To fill this gap the following hypothesis will be tested.

Hypothesis: Initial H tones in alternating tone sequences do not show a downstep effect.

Material: As proposed in chapter 1 section 1.3 (15), the hypothesis can be tested by comparing the height of an initial H tone (H1) in a sentence in which H1 is immediately preceded by an L tone to the height of H1 in a sentence in which it is the first tone in the sentence. The table 8 provides an overview of the material used to test the absence of downstep in initial position in Akan. An example of the short sentence with alternating LH and of a short sentence with alternating HL tones is presented in (72)a. and b., respectively. The complete list of materials is introduced in chapter 2, section 2.2, block B.

Length	LH	Length	HL
6 σ’s	(44)a.	5 σ’s	(45)a.
10 σ’s	(44)b.	9 σ’s	(45)c.
12 σ’s	(44)c.	11 σ’s	(45)d.

table 8: Material used to test initial downstep.

(72) a. Pàpá Kòfí kàsá. (44)a.
father proper name talk.HAB
‘Father Kofi talks.’

b. Ánàné bìsá. (45)a.
proper name ask.HAB
‘Anane asks.’

Planned comparison: The height of H1 in a sentence with alternating LH tones will be compared to the height of H1 in a sentence with alternating HL tones as in (73).

(73)
```
 L  H1  L     vs.    H1  L  H
 |  |   |            |   |  |
CVCVCV              CVCVCV
```

Results: The figure 35 shows the F0 of a sentence with alternating LH tones (solid line) and the F0 of a sentence with alternating HL tones (dashed line) produced by a female speaker, aggregated over repetitions. H1 in sentences with alternating LH tones is not realized lower than H1 in a sentence with alternating HL tones.

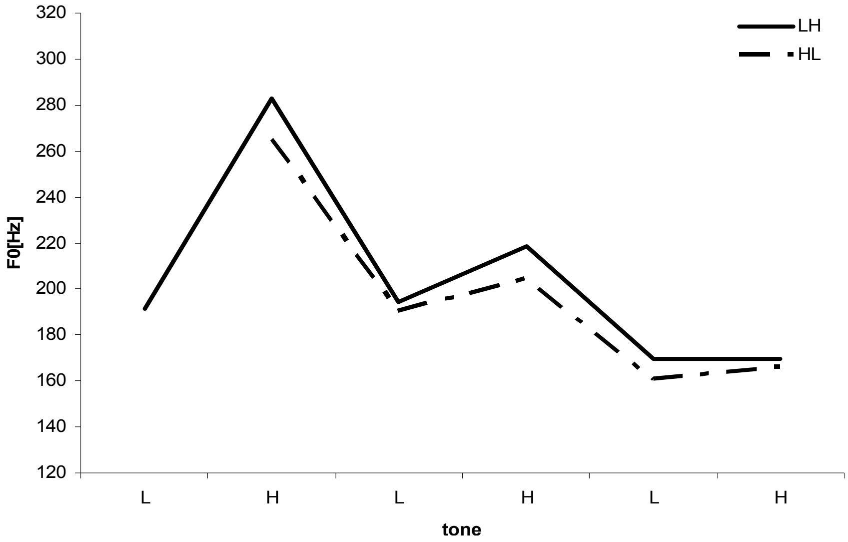


figure 35: Mean F0, aggregated over repetitions, of sentences consisting of alternating LH (solid line), *Papa Kofi kasa.* 'Father Kofi talks.'; Mean F0, aggregated over repetitions, of sentences consisting of alternating HL tones (dashed line), *Anane bisa.* 'Anane asks.'; female speaker.

The mean F0 value obtained for the initial H tone (H1) in the LH configuration, aggregated over repetitions, speakers and lengths, amounts to 193.84 Hz (57), and to 191.82 Hz (56) for H1 in the HL configuration. The figure 36 provides information on the height of H1, aggregated over repetitions and lengths, in sentences with either LH or HL tonal configuration, for all speakers. H1 is not realized slightly lower in an HL environment than in an LH environemt for speakers 1 and 5. Speakers 3 and 6, however, show an effect into the opposite direction. All remaining speakers do not show a difference of the tonal height. Thus, H1 seems not to be systematically affected (lowered) by the presence of a preceding L tone, in LH environment.

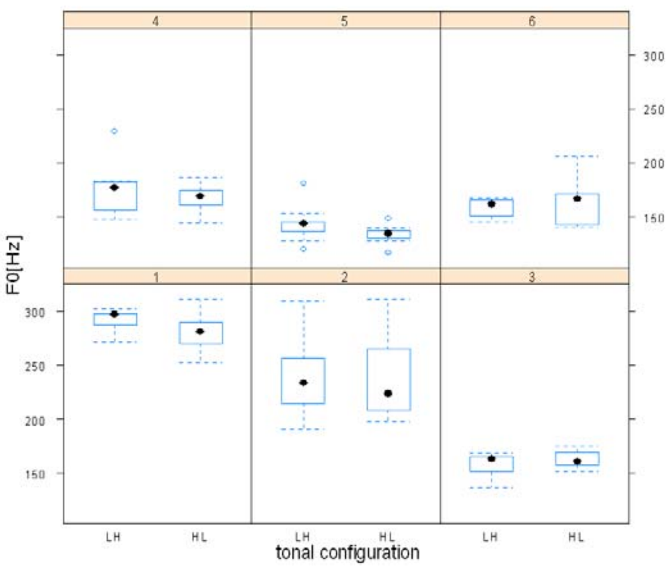


figure 36: Boxplot F0 of H1, in sentences with alternating LH tones and HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration.

A linear mixed effects model was run on the height of the first H tone. Tonal configuration (LH vs. HL) was included as fixed factor. Repetitions, speakers and sentence lengths were treated as random factors. The factor tonal configuration did not show a significant effect.

Summary: Like in Yoruba (Laniran & Clements, 2003), initial H tones in Akan which are preceded by an L tone are not lowered, see figure 35 for illustration. Speakers rather show consistent initial H tone values in sentences with alternating LH and HL tones. This finding is in line with Huang's (1985) claim that non-application of downstep in initial position is a general property because the lowering of an H tone is relational to a preceding H tone. The finding further provides evidence for the view that initial tones serve as reference values (e.g. Liberman & Pierrehumbert, 1984). In the next section, the relational nature of downstep and the relevance of initial tones as reference values will be shown to be essential for the modeling of F0 lowering in Akan.

4.3 Declination and downstep

In this subchapter, the pitch lowering phenomena declination and downstep will be explored. In section 4.3.1, declination in Akan will be investigated with the help of controlled experimental data. The amount of declination will be of interest, especially in relation to the tonal identity and sentence length. Finally, the mental representation of declination will be discussed and modeled. Section 4.3.3 is concerned with downstep. It bears on the issue whether declination and downstep have to be regarded as separate processes or not.

4.3.1 Declination

Introduction: The phenomenon has been already mentioned in Pike (1945). The terminus declination was introduced by Cohen & ´t Hart (1967) to describe a downward trend of the F0 during the course of a declarative utterance[30]; see also Collier (1975, 1985) and Becker (1979). Thus, the domain of declination is the whole utterance (IP). The figure 37 illustrates declination in a Hausa sentence consisting of only H tones.

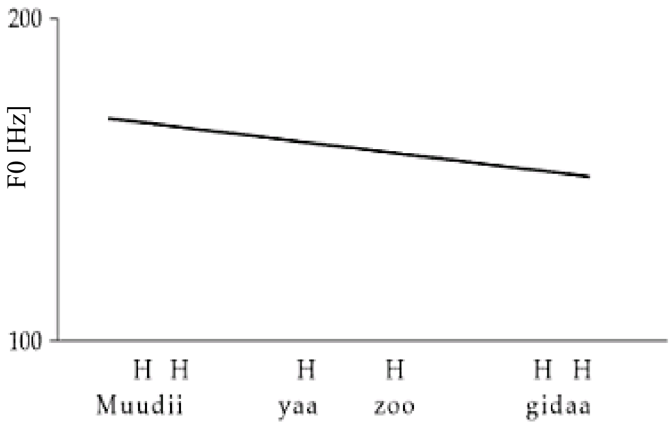


figure 37: F0 of a Hausa sentence containing only H tones, *Muudii yaa zoo gidaa.* 'Muudi came home.' (originally from Lindau, 1986:761); reproduced from Connell (2011:838).

The study of declination is particularly interesting in tone languages because they provide the opportunity to investigate declination in its pure form, in sentences with only H or only L tones. Further, it has been recognized that declination and tonal contrast interact. L tones are likely to show declination in general, whereas H tones do not decline if they contrast with either M or !H in a language (Hyman, 1975:228), i.e. the phonetic implementation can be constrained by the phonology of a language. Thus, the presence or absence of declination can serve as an indicator to determine which surface occurrences of tones are phonologically contrastive i.e. have to be regarded as part of the phonology of a language. This is relevant for

[30] The size of the domain of declination is not generally agreed upon; see Breckenridge (1977) "intonation group"; Pierrehumbert & Beckman (1988) "utterance" among others. I will use the term utterance here.

Akan, since Dolphyne (1988, 1994) has argued that !H is part of the phonology whereas Abakah (2000, 2002) holds the view that all surface occurrences of lowered H tones are due to L̥ in the underlying representation. Experimental investigations on the issue in other African tone languages largely confirm Hyman's observation (e.g. Hombert, 1974 for Shona and Dschang; Lindau, 1986 for Hausa; Connell & Ladd, 1990 and Laniran & Clement, 2003 for Yoruba; Liberman et al., 1992; Ikekeonwu, 1993 for Igbo; Snider, 1998 for Chumburung; Urua, 2002 for Ibibio; Connell, 2003 for Mambila). The table 9 provides an overview of the available experimental evidence from African tone languages.

Language	**Tone**	**Declination**
Chumburung (Snider, 1998)	L H	yes > H yes
Hausa (Lindau, 1986)	L H	not available yes
Shona (Hombert, 1974)	L H	not available yes
Yoruba (Connell & Ladd , 1990; Laniran & Clements, 2003)	L M H	yes ? yes ? yes ?
Mambila (Connell, 2003)	L upper M lower M H	yes > M yes yes no
Dschang (Hombert, 1974)	L !L H !H	no not available not available not available
Ibibio (Urua, 2002)	L H !H HL HL	yes > H yes not available not available not available

table 9: Interaction of number/type of tonal contrasts and declination.

The three languages with a two-way tonal contrast, Chumburung, Hausa and Shona show declination for sentences with only H tones. Sentences containing only L tones decline to a greater degree than H tones in Chumburung. The remaining four languages with M, !H and/or !L contrast do either show no declination, as in Mambila, or less declination in sentences with only H tones than in sentences with only L tones, as in Ibibio. The empirical facts for Yoruba are not clear. Hombert (1974), Connell & Ladd (1990) and Laniran (1992) found that sentences with only H tones exhibit no declination and that sentences with only L tones show a small pitch decline. Laniran & Clements (2003), on the other hand, present data from one speaker who shows a gradual decrease in the degree of declination from H over M to L. They also show that speaker-specific differences concerning the presence and the degree of declination are present in Yoruba. Laniran & Clements (2003:244) conclude that there is "…no strong evidence for any global, utterance-level declination in Yoruba, beyond the effects of downstep and what we have termed ''background downdrift'' across individual tone levels.".

Generally, utterances tend to correlate with breath groups (Lieberman, 1967). Since subglottal air pressure (P_{sg}) is higher at the beginning of a breath group/utterance than at the end, pitch is higher at the beginning than at the end (e.g. Gussenhoven, 2004). Consequently, pitch lowers gradually (declination)[31]. Gussenhoven (2002, 2004) relates this feature of speech production to universal form-meaning relations in intonation, also known as Production Code. He (2004:89) states that "there is no obvious meaning to be attached to this fact, other than that the utterance is progressing." i.e. a gradually declining F0 signals coherence (Hansson, 2003). Other meanings, associated to the Production code, are related to the discourse structure. Möhler & Mayer (2001) propose a phonetic discourse model of pitch range control based on German data. I will only present the basic ideas for single discourse elements, since the thesis is not concerned with discourse analysis and each utterance used here constitutes its own discourse element. The pitch range of a speaker is divided into two categorical register levels, high (h) and low (l), which associate with IPs. Register is seen as a phonological entity, which is phonetically interpreted as pitch range. The phonetic interpretation of the register tones manifests itself as width and position of the pitch register of an intonation phrase with respect to the speaker's overall range. An h register tone is realized in the higher portions of the speaker's range and l is realized in the lower half; see figure 38.

[31] I am aware that there is a huge amount of literature concerned with the relationship of P_{sg} and declination. Earlier works suggested that declination can be attributed solely to a decrease of P_{sg} (e.g. Lieberman, 1967; Collier & Gelfer, 1983; Strik & Boves, 1995). However, a growing amount of experimental evidence suggests that declination is under speaker control and not only related to the muscular control of the respiratory system. Declination seems to require muscular activity at the laryngeal level (e.g. Maeda, 1976; Ohala, 1978; Collier, 1985; Xu & Wang, 1997; Ohala et al., 2004).

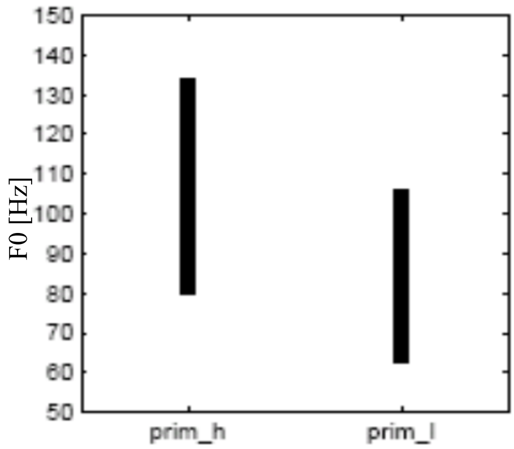


figure 38: Pitch range of the primary registers h and l for the male speaker; reproduced from Möhler & Mayer, (2001).

Assuming that the register tones provide a frame for the phrase internal tones, the question arises how actual F0 values can be assigned. The formula in (71), which is repeated in (74), has been used to predict declination in Mandarin Chinese (Shih, 2000). Let us suppose that an IP associated with left edge h and right egde l constitutes the unmarked/default case and that initial tones, either H or L, receive reference values according to their phonological identity plus register specification; see section 4.2. It has been shown that the lowering of F0 can be described as an exponential decay towards a non-zero asymptote (e.g. Anderson, 1978; Pierrehumbert, 1980; Liberman & Pierrehumbert, 1984; Bird, 1994[32]; Myers, 1996; Shih, 2000; Laniran & Clements, 2003).

(74) $X_{i+1}\text{-}r = s*(X_i\text{-}r)$

The non-zero asymptote is a speaker-specific baseline value (*r*), which is a F0 value representing the bottom of the speaker's pitch range. I will follow Liberman & Pierrehumbert (1984) and Shih (2000), who assume a constant value i.e. independent on the tonal specification; see Bird, 1994 for a tone-specific account. The value of any tone (X_{i+1}), L or H, in a sequence of only L/only H can be calculated from the preceding tone (X_i). X_{i+1} is proportionally lowered from X_i by a lowering coefficient (*s*) in relation to *r*. The coefficient *s* is a numerical value between 0 and 1 characterizing the degree of F0 decline, which may be language and even speaker-specific. If *s* is less than 1 each successive tone is rendered lower than the previous one by *s*, in relation to *r*.

[32] Thanks to Bernadett Smolibocki for discussing the paper with me.

An example illustrating the function of the algorithm is presented in (75). The coefficients, the initial and the final value are hypothetical. The initial H tone value (H1) for the speaker, whose range is illustrated in figure 39, is at 130 Hz, the baseline value *r* is at 60 Hz and the lowering coefficient *s* it at 0.88. The second H tone (H2) in a sequence of only H tone can be calculated by fitting the values into the formula in (74).

(75) H2-r = s*(H1-r)
H2-60Hz = 0.88*(130Hz-60Hz)
H2-60Hz = 61.6Hz
H2 = 121.6 Hz

This procedure applies iteratively from left to right with one tone look-ahead. Thus, H3 will be calculated on the basis of H2 and so forth. As a consequence, lowering intervals between tones becomes smaller as the utterance is progressing. This has been supported by experimental findings, such as declination rate is faster in the beginning of an utterance and decreases as the sentence progresses (e.g. Maeda, 1976; Cooper & Sorensen, 1981; Shih, 2000) and declination rate decreases with increasing sentence length (e.g. Cooper & Sorensen, 1981 for English; Lindau, 1986 for Hausa; Swerts et al., 1996 for Swedish; Connell, 2003, 2004 for Mambila; Shih, 2000, Yuan & Liberman, 2010, for Mandarin Chinese).

The introduction has shown that the domain of declination is the entire utterance (IP) and that it functions to signal coherence (Hansson, 2003; Gussenhoven, 2004). It can be represented by register tones associated to the IP (Möhler & Mayer, 2001) and phonetically interpreted as gradual decay towards a non-zero asymptote (Shih, 2000). Declination rate interacts with utterance length (e.g. Cooper & Sorensen, 1981) and its presence in (tone) languages depends on the number and nature tonal contrasts on a language (Hyman, 1975). Languages with phonologized downstep (!H) and M tones only show declination in sentences with only H tones as illustrated for e.g. Dschang (Hombert, 1974) and Mambila (Connell, 2003). The study by Laniran and Clements (2003) has shown that declination is not only language-specific but even speaker-specific. Declination has not been experimentally investigated for Akan. However, it has been classified as terraced level tone system (Welmers, 1959; Clements, 1979; Abakah, 2000) which means that the realization of phonologically identical tones differs depending on their position in the utterance (Hyman, 2001:1369), the later the lower. Hence there are good reasons to assume that Akan sentence with only H and only L tones show declination. Further, following Abakah's (2000, 2002) proposal that !H is not part of the phonology of Akan, it is expected that H tones in sentences with only H tones show declination. The factor sentence length is expected to affect the declination rate, the longer the flatter (e.g. Ohala et al., 2004). The following hypotheses will be tested.

Hypotheses:

i. F0 in sentences containing only H tones and only L tones declines.
ii. Declination rate decreases with length.

Material & Measurements: The first hypothesis (i.) can be tested by comparing the declination rate in sentences with only H and only L tones. The second hypothesis (ii.) can be investigated by comparing the declination rate between short, medium, long and longest sentences. The table 10 provides an overview of the material used to test declination in sentences with only H and only L tones differing in length. An examples of the medium sentence containing only H tones and of the medium sentence containing only L tones is presented in (76)a. and b., respectively. The complete list of materials is introduced in chapter 2, section 2.2, block A.

Length	**H**	**Length**	**L**
short (5 σ's)	(42)a.	short (7 σ's)	(43)a.
medium (7 σ's)	(42)b.	medium (8 σ's)	(43)b.
long (10 σ's)	(42)c.	long (10 σ's)	(43)c.
longest (12 σ's)	(42)d.	longest (12 σ's)	(43)d.

table 10: Material used to test declination.

(76) a. Kúkúó-bá pápá nó. (42)b.
pot-DIM good DEF
'The good small pot.

b. Àsàrè fì Àkyèmfò. (43)b.
proper name originate.HAB proper name
'Asare comes from Akyemfo.'

To determine the declination rate, the total length of the utterance was measured in seconds (sec). The initial and final Hz values were converted into semitones (st) using the following equation: f(st) = 12log2 (f (Hz) / 100 Hz) (Rietveld & van Heuven, 1997:370). The reference value of 100 Hz was replaced by a F0 mean calculated over the whole utterance; see Lai, Evanini & Zechner, 2011 for usage of F0 median. This was done to make the declination rate values of the individual speakers more comparable. Finally, the difference between the two values (Δ), initial F0 and final F0, was divided by the total length (Adriaens, 1991). This is illustrated by the following equation: rate (st/sec) = f_{initial} (st) - f_{final} (st) / duration (sec). Speaker 2 was discarded from the analysis of sentences with only H tones because of falsetto voice.

Results: The figure 39 presents a high resolution F0 course[33], aggregated over repetitions, for the Akan sentences containing only H tones, uttered by a female speaker. All sentences decline in pitch. The solid line refers to the short renditions, the dotted line to the medium renditions, the dashed & dotted line to the long renditions and the dashed line to the longest renditions. The longest sentence shows a decrease in declination rate compared to the shorter realizations. The initial F0 increases gradually with sentence length, except for the longest rendition. This aspect will be further explored under the heading anticipatory raising in chapter 5 section 5.2.

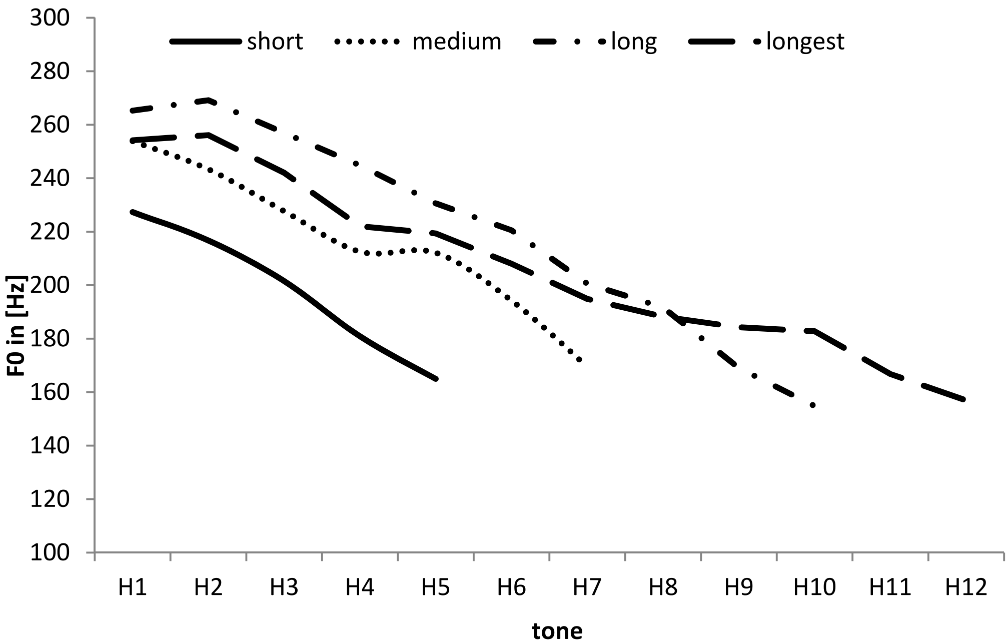


figure 39: Mean F0 of sentences containing only H tones differing in length, aggregated over repetitions; short sentences (solid line), *Kukuoba no.* 'The small pot.', medium sentences (dotted line), *Kukuoba papa no.* 'The good small pot.', long sentences (dashed & dotted line), *Kukuoba papa no bɔ daa.* 'The good small pot breaks everyday.', longest sentences (dashed line), *Kukuoba papa paa no bɔ daa.* 'The very good small pot breaks everyday.'; female speaker.

The figure 40 presents a high resolution F0 course, aggregated over repetitions, for the sentences containing only L tones uttered by a female speaker. Again, all sentences decline in pitch and longer sentences show a slower declination rate than shorter ones. As for the H tones, this speaker shows a gradual increasing of initial F0 with sentence length, except for the longest rendition.

[33] A Praat script was used, which divided TBU in 10 intervals. For each interval F0 was extracted in Hz. Thanks to Tobias Günther for assistance.

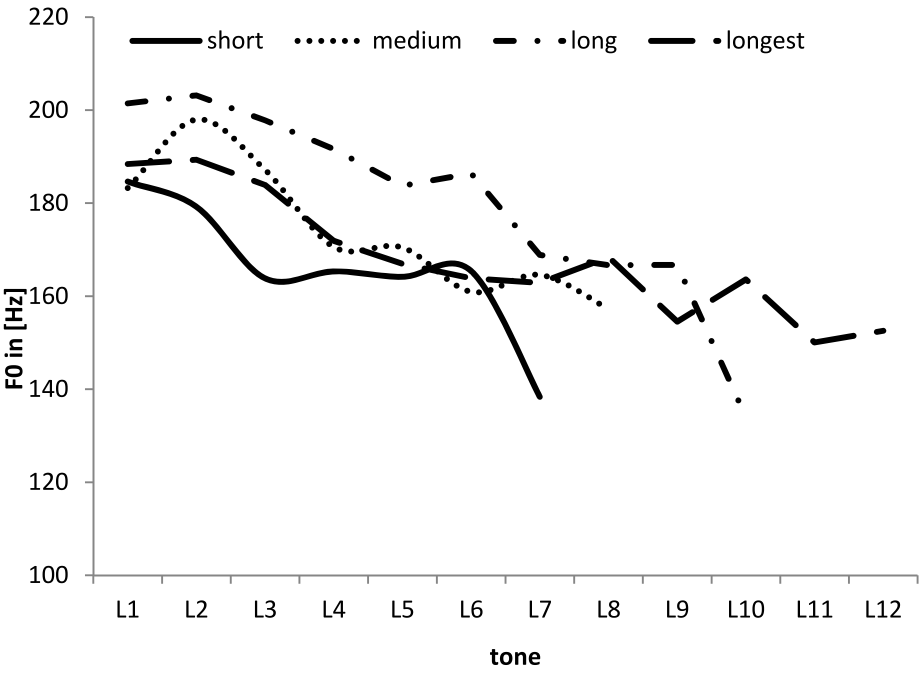


figure 40: Mean F0 of sentences containing only L tones differing in length, aggregated over repetitions; short sentences (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akyemfo.', medium sentences (dotted line), *Asare fi Akyemfo.* 'Asare comes from Akyemfo.', long sentences (dashed & dotted line) *Wɔfa Asare fi Akyemfo.* 'Uncle Asare comes from Akyemfo.', longest sentences (dashed line), *Wɔfa Ado Asare fi Akyemfo.* 'Uncle Ado Asare comes from Akyemfo.'= 12 syllables; female speaker.

The mean declination rate in sentences with only H tones, aggregated over speakers and repetitions, amounts to 8.37 st/sec (3.6) for the short sentences, to 7.12 st/sec (1.6) for the medium sentences, to 5.28 st/sec (1.6) for the long sentences and to 4.3 st/sec (1.3) for the longest sentences.

The mean declination rate in sentences with only L tones, aggregated over speakers and repetitions, amounts to 5.21 st/sec (1.2) for the short sentences and to 5.32 st/sec (1.5) for the medium sentences. The value for the long sentences is at 4.52 st/sec (1.7) and at 3.21 st/sec (0.8) for the longest sentences. The declination rate decreases as a function of sentence length for both tonal entities. The figure 41 provides speaker-specific information of the declination rate in sentences with only H tones (left hand side) and sentences with only L tones (right hand side), differing in length, aggregated over repetitions. For the sentences with only H tones, a gradual decrease with increasing length can be observed for speaker 1 and 4. Speaker 5 shows the same effect but not between long and longest renditions. The data of speakers 3 and 6 exhibits the effect for short vs. longest renditions only. For the sentences with only L tones, a gradual decrease with increasing length can be observed for speaker 1 and 4. Speakers 3 and 5 show a higher value for the medium length but a decrease in declination rate between short, long and longest. The data of speaker 6 exhibits no difference in declination rate for short, medium and long but a decrease for the longest renditions.

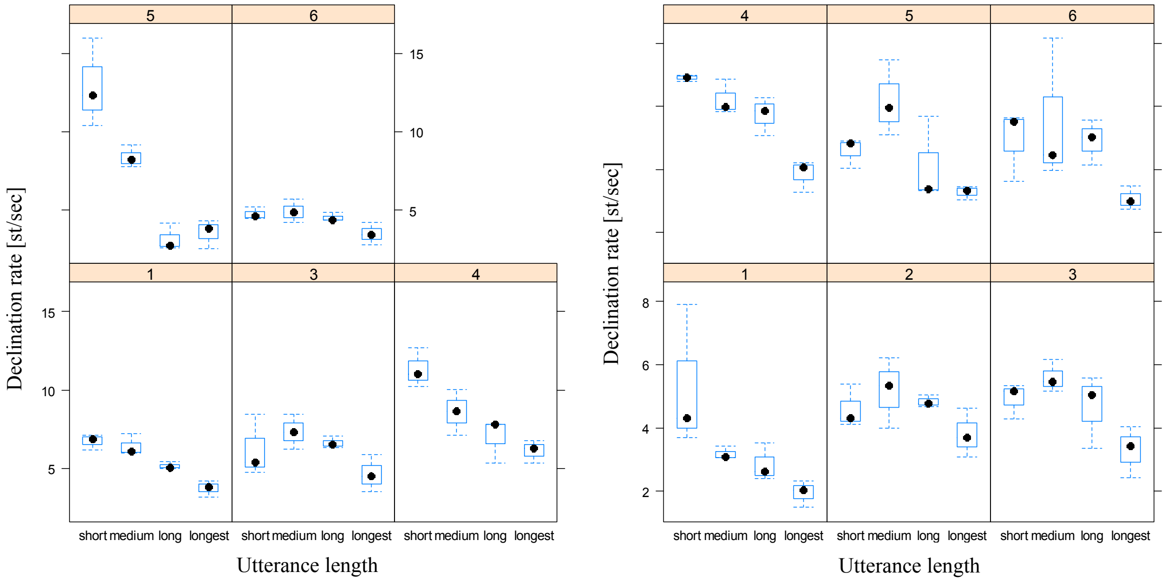


figure 41: Left boxplot: declination rate for sentences with only H tones, aggregated over repetitions, split by speakers and utterance lengths (short = 5σ's, medium = 7σ's, long = 10σ's, longest = 12σ's). Right boxplot: declination rate for sentences with only L tones, aggregated over repetitions and split by speakers and utterance lengths (short = 7σ's, medium = 8σ's, long = 10σ's, longest = 12σ's).

A linear mixed effects model was calculated on the declination rate. Length (short as reference category vs. medium/long/longest) and tone (L/H) were treated as fixed factors. Speakers (without speaker 2) and repetitions were included as random factors. The contrast between short and medium was significant (t = 6.104, pMCMC < 0.0001, 132 observations). The contrast between short and long was also significant (t = 3.575, pMCMC < 0.001) and the contrast between short and longest yielded a significant result (t = -2.289, pMCMC < 0.05); declination rate decreases with increasing length. The factor tone had a significant effect (t = 6.350, pMCMC < 0.0001); declination rate is slower in sentences with only L tones than in sentences with only H tones. Moreover, there was a significant interaction for tone and the length comparison short vs. medium (t = 3.234, pMCMC < 0.01) and for tone and the length comparison short vs. longest (t = 2.095, pMCMC < 0.05). The interactions are illustrated in figure 42. The first interaction is present because the decrease in declination rate that can be observed from short (dotted and dashed line) to medium (dotted line) length for the sentences containing only H tones is not existent for the sentences containing only L tones. The second interaction is significant because the lines representing the short and the longest (solid line) realizations may meet in the future for the sentences containing only L tones, but not for the short and the longest realizations containing only H tones.

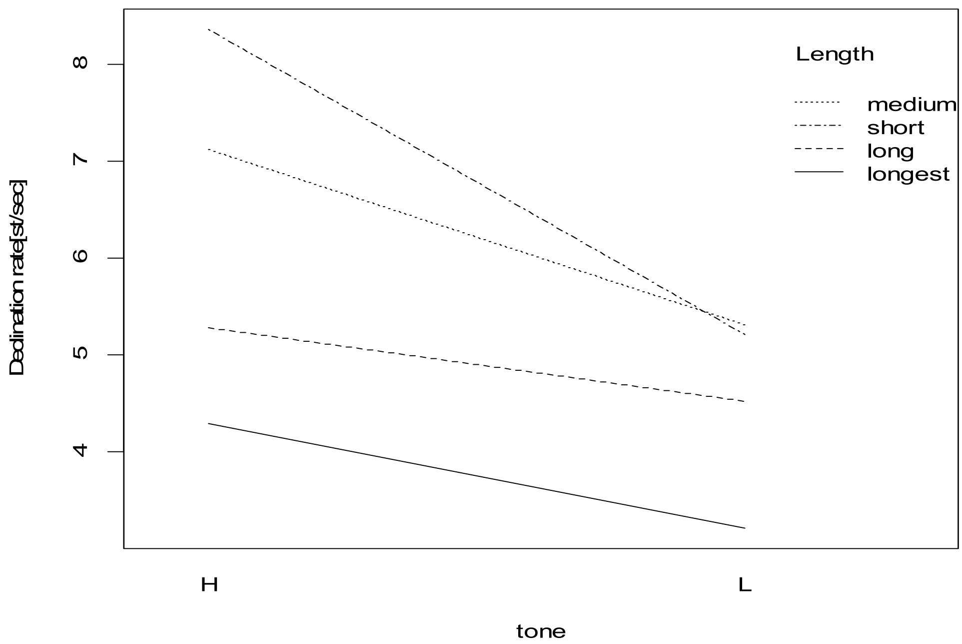


figure 42: Interaction of tone, length (short – dotted and dashed line, medium - dotted line, long – dashed line, longest – solid line) and declination rate.

Summary: Akan declarative sentences with only H and only L tones show declination; see figure 39 and figure 40. Unlike in Yoruba (Laniran & Clement, 2003), declination in Akan is systematic. Akan sentences with only H tones generally decline faster than sentences with only L tones, see figure 41, contrary to the situation in Chumburung (Snider, 1998). As in many other tone (e.g. Lindau, 1986) and intonation languages (e.g. Cooper & Sorensen, 1981), the rate of declination decreases with the length of the sentence.

Discussion: The data has provided evidence that !H is not lexicalized in Akan since sentences with only H tones exhibit a considerable amount of declination. It has been shown in other African tone languages, that declination in sentences with only H tones is suspended if !H is lexicalized (Hyman, 1975:228). The finding speaks against Dolphyne's (1988, 1994) claim that Akan exhibits "lexical" downstep and supports Abakah's (2002) proposal that occurrences of !H are instances of non-automatic downstep; see chapter 1 section 1.1.

Turning to the discussion of the grammatical status of declination in Akan, the data has provided good reasons to represent declination phonologically. I have shown that the declarative sentences used in the present study show declination. With a few exceptions e.g. short renditions of speakers 4 and 5 in the sentences containing only H tones, see figure 41, declination rate is relatively constant between and within the speakers, contrary to declination in Yoruba (e.g. Laniran & Clements, 2003). Unlike in e.g. Hausa and Chichewa, declination in Akan is not affected by sentence type. Anticipating the prosodic analysis of Yes – No questions in chapter 6, figure 43 shows that declination occurs irrespective of the register raising effect in Yes – No questions; see chapter 1 section 1.6.1 (38)b. for glosses and chapter 6 for further details on the intonational marking of Yes – No questions.

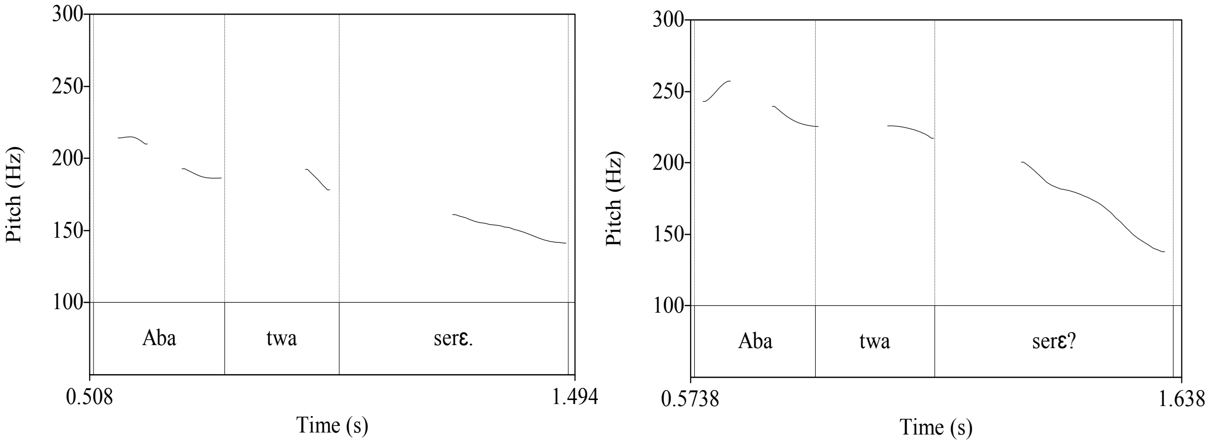


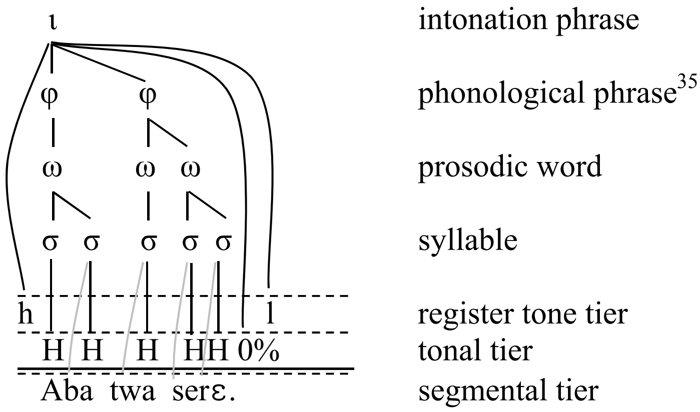
figure 43: Left pitch track: statement consisting of only H tones, *Aba twa serɛ.* 'Aba cuts grass.'; Right pitch track: Yes – No question consisting of only H tones, *Aba twa serɛ?* 'Aba cuts grass?'; female speaker.

I thus propose that declination in Akan is phonological. To represent it, register tones, as recommended by Möhler & Mayer (2001), will be used. The register tones are not strictly "discoursal" as in the original sense but serve to generate the gradual lowering effect. Its linguistic function is to signal coherence (Hansson, 2003). The mental construct of an Akan declarative sentence with only H tones is illustrated in (77). The phonological surface representation is inspired by Pierrehumbert & Beckman (1988) and Möhler & Mayer (2001). The phonemes are represented on the segmental tier. Lexical (H) tones are displayed on the tonal tier, above the segmental tier. Lexical tones are connected with their TBUs[34] by association lines. The final element on the tonal tier is a tonally unmarked boundary tone, referred to as 0% following Grabe (1998). Grabe (1998:59) notes "…'0%' is not assumed to reflect a phonological category but is a place holder indicating the end of an intonation phrase which does not appear to be associated with a tone….". and continues by saying that 0% reflects "…absence of pitch movement at IP boundaries…". I have not provided any independent analysis of the absence of any distinctive pitch movement at the IP boundary here. In chapter 6, it will be shown that declarative sentences exhibit final lowering. It is a matter of debate whether final lowering should be seen as the phonetic manifestation of L% (Pierrehumbert & Beckman, 1988) or as a phonetic effect, which arises due to a decrease of articulatory effort towards the end an utterance (Herman et al. 1996, Herman, 2000; Arvaniti,

[34] I choose the syllable as TBU here, but as pointed out in chapter 1 section 1.2 the identity of the TBU in Akan is still an unresolved issue.

2007). I will assume a combination of both views for Akan and show in chapter 6 that if an intonational morpheme is tonally specified, articulatory effort increases. Thus, unlike Grabe (1998), 0% is taken to reflect a phonological category which does not mark the end of an IP but is an intonational morpheme, in the sense of Ladd (1996), indicating sentence type. Since 0% is by definition not tonally specified, its phonetic manifestation is a decrease in articulatory effort which results in the surface pattern of final lowering. The tonal tier is dominated by the register tone tier. All other projections (syllable, prosodic word…) are represented for the sake of completeness to motivate the IP node. The h register tone associates to the left edge of the IP and the l register tone to the right edge. The phonetic interpretation of the tones on the register tier will manifest itself, following Möhler & Mayer (2001), in the realization of the initial H tone and the final H tone, in this particular case. The initial H tone on the tonal tier is dominated by h on the register tier and is thus realized in the higher portions of the speaker's range. The final H tone on the tonal tier is dominated by l on the register tier and is thus realized in the lower half of the speaker's range.

(77)



The phonological representation in (77) makes several predictions. First, the domain of declination is the IP. Hence, no uses of declination reset at smaller boundaries, in particular phonological phrase boundaries, are expected; see chapter 5, section 5.2 for evidence. Second, the declination pattern should not be influenced by intonational morphemes since it would undermine its form and thereby also its function. This is illustrated in figure 43 and more specifically outlined in chapter 6. Declination is present in Yes – No questions, which exhibit a global register raising effect. Third, the lack of intra- and interspeaker variation in declination in Akan is challenging under the standard view that phonetics is the domain of gradient entities (see Cohn, 2007 and references therein). Further, the amount of F0 decline is stable irrespective of the tonal specification. More evidence will be provided in section 4.3.3

[35] Only maximal phonological phrases are shown. A maximal phonological phrases is a phrase, which is not dominated by any other phonological phrase (Itô & Mester, 2007; Selkirk, 2009).

and can be seen in the fact that declination in sentences with only H and only L tones can be modeled with the same lowering quotient (*s*), as we shall see in the following section.

4.3.2 The implementation of declination

Turning to the phonetic implementation of declination, as pointed out in the introduction, the aim is to model declination as an exponential decay towards a non-zero asymptote, following Shih (2000). The model predicts that declination slope decreases with increasing length. The data presented for Akan sentences with only H and only L tones has shown this effect. Further, the model predicts a higher declination rate for sentences with only H tones than for sentences with only L tones since the input value (T1) is per definition lower in the latter case. Hence, the distance between the starting F0 and the baseline value is smaller in sentences with only L tones. Lowering of the F0 by the lowering quotient thus results in a smaller rate of decline for the sentences with only L tones than for the sentences with only H tones. This effect was also present in the Akan data.

Recall the formula introduced in (71). To fit the model, the coefficients *r* and *s* have to be established. The variable *r* is specified for each speaker individually. All values at the last reliably trackable pitch point were collected to estimate *r*. The lowest value was chosen as a representative of the speaker-specific baseline reference value. Uncertainty surrounds the question of whether the lowering factor *s* has to be regarded as speaker-specific or language-specific. Liberman & Pierrehumbert (1984), Myers (1996) and Shih (2000) use speaker-specific values. Since *s* is seen as the phonetic manifestation of the phonological register tone specification, it seems to be more attractive to establish a language-specific lowering quotient. At the same time, speaker-specific differences (e.g. size of the larynx), affecting the pitch range, exist. For now, I decided for a speaker-specific lowering quotient, see table 10, but all values are located between 0.7 and 0.8. Thus, the phonetic implementation algorithm would also provide satisfactory results with a language-specific *s* value at about 0.75. The value of the lowering quotient, *s*, was determined from the goodness of fit of predicting the height of the second tone from the mean initial tone in the short renditions[36]. An example of the calculation procedure for the short renditions containing only L tones for the female speaker is presented in (78) with the obtained values and differences between obtained and predicted value as orientation; see figure 44 for visualization.

[36] Except for speaker 3; see discussion section for reasoning. For this speaker the medium rendition was used.

(78) L2-r = s*(L1-r)

L2-140Hz = 0.76*(188.65Hz-140Hz)

L2-140Hz = 36.97Hz

$L2_{predicted}$ = 176.97Hz $L2_{obtained}$ = 178.98Hz Δ = 2.01 Hz

L3-140Hz = 0.76*(176.97Hz-140Hz)

$L3_{predicted}$ = 168.1Hz $L3_{obtained}$ = 174.38Hz Δ = 6.28 Hz

L4-140Hz = 0.76*(168.1Hz-140Hz)

$L4_{predicted}$ = 161.36Hz $L4_{obtained}$ = 165.72Hz Δ = 4.36 Hz

L5-140Hz = 0.76*(161.36Hz-140Hz)

$L5_{predicted}$ = 156.23Hz $L5_{obtained}$ = 163.11Hz Δ = 6.88 Hz

L6-140Hz = 0.76*(156.23Hz-140Hz)

$L6_{predicted}$ = 152.33Hz $L6_{obtained}$ = 162.63Hz Δ = 10.3 Hz

L7-140Hz = 0.76*(152.34Hz-140Hz)

$L7_{predicted}$ = 149.38Hz $L7_{obtained}$ = 147.14Hz Δ =-2.24 Hz

The table 11 provides information of the lowering quotient *s*, the baseline quotient *r*, the mean initial tone values (T1) for the sentences containing only L tones, aggregated over repetitions, and R^2, the coefficient of determination for obtained vs. predicted F0 values for each speaker. A R^2 near 1 is assumed to indicate a good fit, whereas an R^2 closer to 0 is taken to indicate that the predicted values do not fit the obtained values well.

First, the formula is applied to predict declination in sentences containing only L tones since this seems to be the unmarked case (Hyman, 1975). After that, declination in sentences containing only H tones will be predicted using the same quotients (*r*, *s*).

Speaker	Length	R	s	T1	R^2
1	1	140	0.76	196.52	0.91
	2			197.32	0.88
	3			207.73	0.87
	4			197.34	0.88
2	1	90	0.77	165.08	0.81
	2			159.51	0.59
	3			182.97	0.62
	4			170.04	0.73
3	1	79	0.75	120.40	0.33
	2			129.66	0.61
	3			127.24	0.48
	4			119.41	0.52
4	1	75	0.8	129.92	0.78
	2			132.29	0.84
	3			131.82	0.70
	4			127.92	0.74
5	1	80	0.72	104.18	0.78
	2			126.07	0.98
	3			115.84	0.57
	4			114.56	0.91
6	1	78	0.7	126.83	0.77
	2			128.35	0.59
	3			122.29	0.50
	4			113.52	0.53

table 11: Reference values *r* and *s*; mean initial F0 (T1), aggregated over repetitions, for short, medium, long and longest sentences containing only L tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6.

The model successfully predicts the declining pattern in sentences containing only L tones in Akan. The R^2 values range from 0.33 to 0.98. The figure 44 illustrates the goodness of the model by comparing the obtained and predicted values of the short renditions of the female speaker. The obtained values, aggregated over repetitions, are represented by the solid line and the predicted values by the dashed line. It can be seen that the curves nearly lie upon each other.

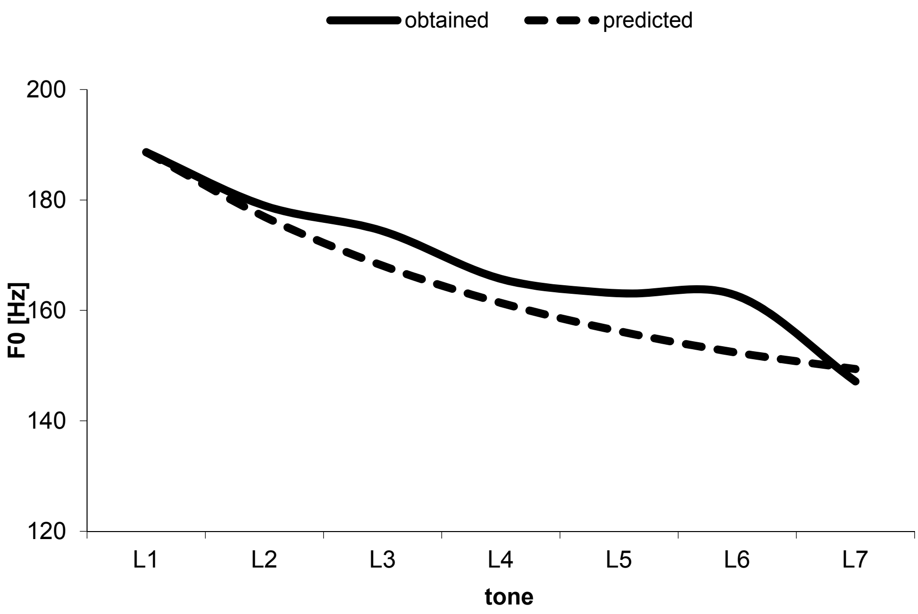


figure 44: Obtained mean F0, aggregated over repetitions, of short sentences containing only L tones (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' and predicted F0 values (dashed line); female speaker.

The table 11 reveals a low coefficient of determination for the short renditions of speaker 3. In his data, the second (113.44 Hz) and third mean L tone value (113.06 Hz) renditions is only very slightly lowered. Additionally, speaker 3 shows surprisingly high values on the antepenultimate and penultimate syllable (*è.m̀*) of the proper name *Akyemfo*. This is illustrated in figure 45. Again, the obtained values, aggregated over repetitions, are represented by the solid line and the predicted values by the dashed line. None of the other speakers shows this effect. It thus seems that we are dealing with a slip of the tongue here.

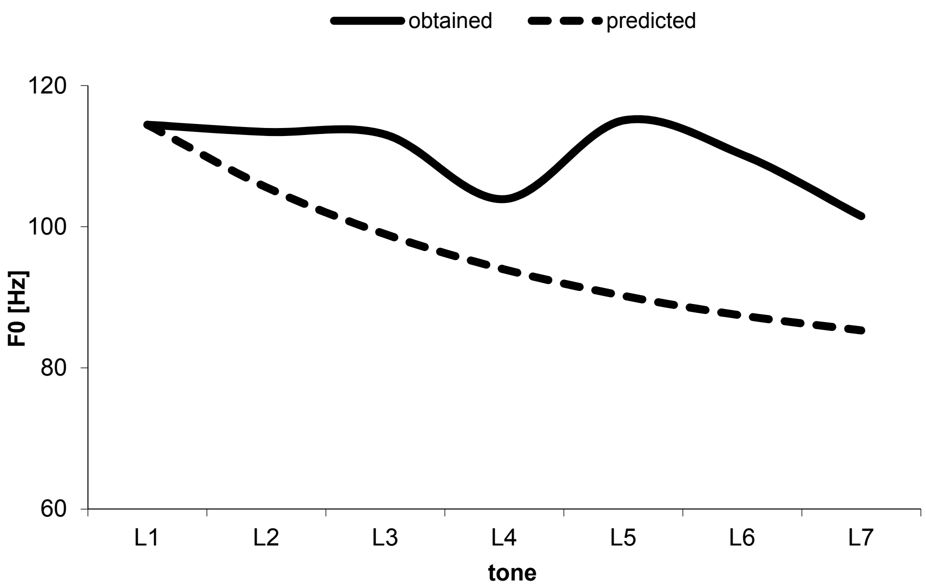


figure 45: Obtained mean F0, aggregated over repetitions, of short sentences containing only L tones (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' and predicted F0 values (dashed line); male speaker 3.

Turning to the prediction of declination in sentences with only H tones, table 12 provides information of T1 for the sentences containing only H tones, aggregated over repetitions, and the R^2 values for each speaker. The same lowering quotient *s* and baseline value *r* as for the prediction of the L tones in the only L tone sentences were used.

Speaker	Length	r	s	T1	R^2
1	1	140	0.76	227.31	0.89
	2			253.83	0.82
	3			272.29	0.86
	4			231.87	0.77
3	1	79	0.75	137.85	0.81
	2			150.80	0.80
	3			157.73	0.57
	4			150.70	0.66
4	1	75	0.8	158.66	0.87
	2			177.72	0.86
	3			184.33	0.64
	4			183.42	0.54
5	1	80	0.72	144.97	0.98
	2			143.95	0.78
	3			149.56	0.71
	4			158.35	0.78
6	1	78	0.7	136.40	0.74
	2			148.66	0.80
	3			158.81	0.77
	4			145.64	0.71

table 12: Reference values *r* and *s*; mean initial F0 (T1), aggregated over repetitions, for short, medium, long and longest sentences containing only H tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6.

The model in its current form predicts the declination in sentences containing only H tones in a satisfactory manner. The R^2 values range from 0.57 to 0.98. The figure 46 illustrates the goodness of the model by comparing the obtained and predicted values of the short renditions of the male speaker 5. The obtained values, aggregated over repetitions, are represented by the solid line and the predicted values by the dashed line. It can be seen that the curves nearly lie upon each other.

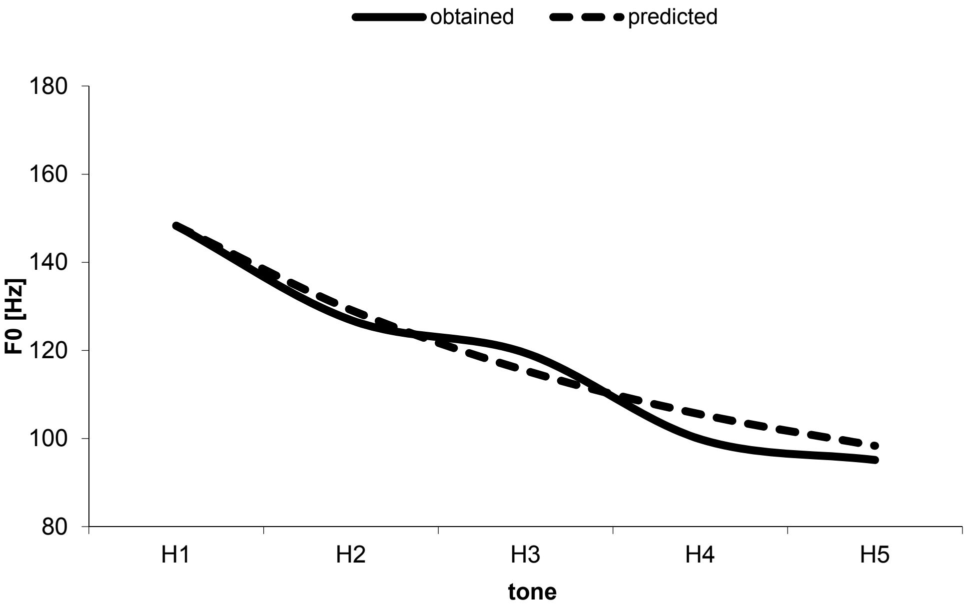


figure 46: Obtained mean F0, aggregated over repetitions, of short sentences containing only H tones (solid line), *Kukuoba no.* 'The small pot.' and predicted F0 values (dashed line); male speaker 5.

To sum up, it has been shown that Akan has a clear declination effect. The effect has been successfully modeled as an exponential decay (Shih, 2000). The pitch implementation model predicts the F0 in sentences with only H and only L tones using three values: the initial value, whose relationship with sentence length will be further explored in chapter 5 section 5.2, the baseline value *r,* which was estimated for each speaker individually, and the lowering quotient *s,* which was assigned individually but centers around a language-specific value of 0.75. It was assumed that *s* is activated by the phonological register tones, see (77).

The next section will be concerned with downstep in sentences with alternating tones (LH/HL). The scaling of H and L tones will be explored as well as the relationship of downstep and declination.

4.3.3 Downstep

Introduction to Downstep: As pointed out in chapter 1 section 1.3 and section 4.2 of this chapter, downstep in African terraced level tone languages like Akan is best defined as a lowering of H in reference to a preceding H (Huang, 1985); see figure 35 for illustration. This definition fundamentally differs from that proposed by Wang & Xu (2011:597) for Mandarin Chinese. In Mandarin Chinese, downstep is a co-articulatory phenomenon[37]; see also chapter 1 section 1.3, figure 5 for illustration. Concerning the attribution of downstep to the phonetic or phonological component of the grammar, Laniran & Clements (2003:204) remark that "Downstep of H tones differs from assimilation in that it lowers the ceiling at which following H tones are realized; as a result, each successive H tone in longer downstepping sequences is lower than the preceding one, creating a cumulative "staircase" pattern.". Assimilation in this citation refers to carry-over lowering, which can be regarded as phonetic effect. The lowering of the ceiling (abstract reference line for H tones), also known under the heading key lowering (Stewart, 1981) or register lowering (e.g. Clark, 1993), has been considered as one of the major properties of downstep and has inspired numerous phonological proposals involving register tones (e.g. Hyman, 1985; Inkelas, Leben & Cobler, 1987; Snider, 1999). Phonological proposals have been introcuded in section 4.2 (69).

However, one of the main findings of the investigation of phonetic details of downstep in Yoruba revealed that there is "no evidence for the manipulation of the full set of tone levels as a whole, corresponding to what some linguists have called "register shifts or "key lowering" – i.e., a shift of all tone levels downward or upward as a block." and "…for some speakers, it is similar to downstep as it has been described in English, Japanese, Dutch, and other nontonal languages. This result supports the widely assumed, but previously unsubstantiated view that downstep in nontonal languages may be of the same fundamental nature as downstep in African languages." (Laniran & Clements, 2003:243). Furthermore, Huang (1985:211f.) remarks that "…the amount of drop involved in terracing in relation to the distance between phonemic tones need not be constant in a given language or even for a given speaker.". These observations together with the fact that downstep is not contrastive (e.g. Huang, 1985) make a phonological representation less appealing. If downstep would result out of the phonetic interpretation of a register tone specification as illustrated in (69), less variation/gradience is expected.

[37] Carry-over lowering has been explained by constraints of the physiology of the speech apparatus. Xu and Sun (2002:1411) show that speakers of Mandarin Chinese need 142 ms to complete a pitch rise of 6 st. Thus "downstepped" H tones in Mandarin Chinese may just reflect a tonal undershot due to time pressure.

Automatic downstep in English has been modeled as an exponential decay towards a non-zero asymptote as already outlined in section 4.2; see formula (71) from Liberman & Pierrehumbert (1984). The L tones, which are seen as part of a bi-tonal accent in English (e.g. H*+L), trigger the phonetic realization rule which lowers each H tone in relation to the preceding one by *s*. Non-adjacent H tones in an HLH environment are scaled relationally; see also Bird (1994). General advantages of this approach are that it gives an explicit phonetic output which can be evaluated empirically, that it has been successfully applied to a wide range of typologically different languages, e.g. English (Liberman & Pierrehumbert, 1984), Chichewa (Myers, 1996) and Yoruba (Laniran & Clements, 2003) and that it has been successfully implemented in text-to-speech systems (see e.g. Prieto et al., 1996). However, the scaling of L tones has not been explicitly approached in Liberman's & Pierrehumbert's proposal. Liberman & Pierrehumbert (1984:219) speculate that the scaling of L tones is symmetric to that of H tones[38]. In support of this view, Liberman et al. (1992:746) remark for Igbo that downstepped H and L tones seem to behave identically. L tones in Akan are underlyingly present, unlike in Chichewa, in which syllables with a surface L tone are assumed to be underlyingly toneless. Firstly, they play no role in the phonology (e.g. no spreading) and secondly, they appear as medial dip between H tones on the surface which is interpreted as a return to default laryngeal configuration (Myers, 1996). In Akan, L tones are active in the phonology (see e.g. Abakah, 2005:115f. for examples involving L tone spreading) and are not realized as medial dip on the surface. Thus L tones have to be modeled in Akan.

Introduction to the relationship of Downstep & Declination: Most researchers agree on the fact that both have to be regarded as distinct phenomena. Laniran & Clements (2003:204) state that "Downstep must be carefully distinguished from other principles that contribute to pitch lowering (see also Connell & Ladd, 1990, pp.2-3)." such as declination; see also Connell (2002a, 2011), but also mention the possibility that downstep may be equivalent to declination: "A declining pitch ramp across H tones could result from other principles such as general background declination of all tones…." (Laniran & Clement, 2003:216). Others suggest that declination is a necessary prerequisite for the emergence of downstep (e.g. Yip, 2002:10; Gussenhoven, 2004:98). Hyman (1975:227f.), referring to Schachter (1965), notes

[38] The phonetic implementation model has been revised to account for the Japanese data (Pierrehumbert & Beckman, 1988:184f.) at the expense of simplicity (Dilley, 1997:65). H and L tones are scaled in reference to prominence values (from 0 to < 1.0). The prominence values define the scaling of L and H tones with respect to a high reference line (h). As pointed out in section 4.2, prominence relations in Akan are problematic. Furthermore, unlike Japanese, Akan does not show scaling differences of tones due to focus (Kügler & Genzel, 2012). Therefore the revised model will not be further explored here.

that downstep might be an intonational property; see also Clements et al. (2011:5). Downstep as well as declination can be suspended for purposes of emphasis or in questions, as in e.g. Hausa (Inkelas & Leben, 1990) or Chichewa (Myers, 1996, 2004). Furthermore, Hyman (1975:227f.), referring to Hombert (1974), establishes an analogy between declination and downstep and states that "Virtually all tone languages exhibiting automatic downdrift have only two tones, H and L. Most dialects of Yoruba, Nupe, Ewe and Jukun, all of which have H, M, and L, do not have downdrift….if a three-tone language were to let the second H of a H-L-H sequence undergo downdrift, it would be likely to be confused with an underlying M…". This reasoning is reminiscent of that proposed for the absence of declination in tone systems with more than two levels. If downstep is an intonational property it should serve a linguistic function. Laniran & Clements (2003:204) mention that "…downstep is localized at specific junctures and is usually conditioned by the tonal, lexical, morphological, and/or syntactic structure of the utterance in which it applies, often serving distinctive or demarcative functions.". In the previous section, I have argued with the help of declination that downstepped H tones (!H) in Akan are not lexicalized and hence do not serve distinctive functions. The domain of downstep in terraced level tone languages like Akan is the IP and like declination, a downstep pattern throughout the whole IP may signal coherence (e.g. Bruce, Granström, Gustafson & House, 1991), i.e. an interruption of declination and/or downstep serves a demarcative function.

The introduction to downstep has provided arguments against a phonological representation involving register tones and has pointed out similarities between declination and downstep; see also chapter 1 section 1.3. There are not many controlled phonetic investigations on the relation of the two lowering phenomena. However, the classical empirical argument that declination and downstep are distinct is that downstep causes a greater degree of lowering than declination (e.g. Snider, 1998; Connell, 2002a, 2011; Laniran & Clements, 2003). In what follows, empirical works on the relation of downstep and declination in African tone languages will be introduced.

Connell (2002a, 2011) presents data for Hausa from Lindau (1986); reproduced in figure 47. The plot on the left hand side shows declination in a sentence with only H tones and the plot on the right hand side shows downstep in a sentence with alternating HL tones. Only the H tones are plotted. The lowering of H tones in the sentence with alternating tones is greater compared to the lowering in a sentence with only H tones because both lowering effects, declination and downstep, accumulate; downstep is analyzed as the local assimilation of Hs to

Ls (Connell, 2002a:7). What is however not clear is whether the local assimilation has to be understood as carry-over lowering in the sense of Xu & Sun (2002).

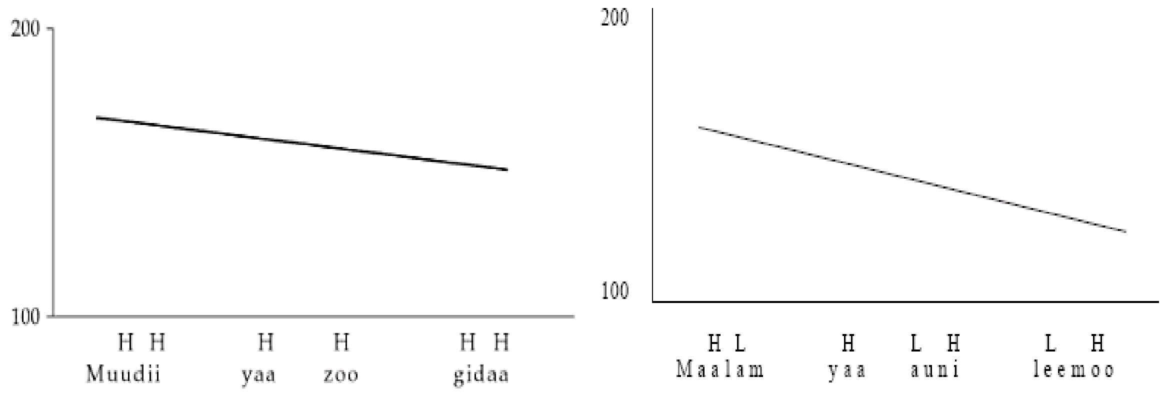


figure 47: Left hand side, F0 of a Hausa sentence containing only H tones, *Muudii yaa zoo gidaa.* 'Muudii came home.'; (originally from Lindau, 1986:761); reproduced from Connell (2011:838); Right hand side, F0 of a Hausa sentence containing alternating HL tones, *Maalam yaa auni leemoo.* 'The teacher weighed the oranges.'; (originally from Lindau, 1986:758); reproduced from Connell (2002a:7).

Ikekeonwu (1993) presents data of sentences with only H, only L and alternating LH/HL tones for Igbo but does not address the question of the relation between declination and downstep explicitly. However, from her plots it becomes evident that pitch in sentences with only L tones declines more than in sentences with only H tones. H tones in sentences with only H tones only show very slight if any declination; see also Liberman et al. (1992) and section 4.3.1. H tones in alternating HL tone sequences exhibit a greater rate of F0 decline than H tones in sentences with only H tones. Note that Liberman et al. (1992) assume that Igbo has three tone levels (L, M, H). For Bimoba which also has three contrastive tones (L, M, H), Snider (1998) shows that H tones in only H tone sequences and L tones in only L sequences decline very slightly. H tones in alternating LH sequences decline to a greater degree than H tones in only H tone sequences. Interestingly, H tones following the lowered H tone show a slight increase in pitch. This behavior is unexpected since in terracing languages like Akan a lowered H tone causes a lowering of the ceiling. It thus seems that the "downstep" effect in Bimoba is rather due to co-articulation (carry-over lowering). Laniran & Clements (2003) investigated for Yoruba, which is also a three tone language, whether H tones and L tones decline at a faster rate in sentences with alternating tones than in sentences with only H and only L tones. They show that H tones in alternating sequences decline more rapidly than in only H tone sequences, whereas L tones in alternating sequences behave similarly to L tones in only L tone sequences. Laniran & Clements (2003:244) remark that "Our discussion has…not addressed the question of whether H raising and downstep are purely phonetic principles in Yoruba….".

To sum up, the available empirical resources have revealed a lack of controlled empirical works on the relationship of declination and downstep in African tone languages with two level tones, like Akan. Besides Hausa, all languages reviewed are three tone languages, which behave differently from Akan with respect to declination. As pointed out in the introduction to section 4.3.1, languages with more than two tone levels are not likely to show declination in sentence with only H tones. To determine whether downstep causes a greater degree of lowering than declination alone, it is favorable to investigate a language with two level tones (L and H) that exibits declination for both tonal entities like Akan. The previous section has shown that H and L tones in Akan sentences with only H and only L tones exhibit a clear declination effect and that H tones in sentences with only H tones decline to a greater degree. Dolphyne's (1994) data has shown that the pitch of L tones in sentences with alternating tones also lowers considerably, see table 3, and that the pitch of H tones lowers, too i.e. H tones are downstepped; see table 7. The size of the pitch drop decreases as the utterance is progressing[39]. The results presented in 4.1 have shown that the lowering effect on H tones is also present when the L tone is not phonetically realized (L̥), thus carry-over lowering might play a minor role, if any. In what follows, I will investigate whether the lowering of L and H tones in alternating LH/HL sequences is greater than the lowering which can be attributed to declination, i.e. whether declination and downstep are different phenomena in Akan. Based on the observations by Snider (1998), Connell (2002a, 2011), Laniran & Clements (2003) the following hypothesis will be tested.

Hypothesis: Downstep causes a greater degree of lowering than declination.
Material & Measurements: As proposed in chapter 1 section 1.3 (17), the hypothesis can be tested by comparing the pitch level of a non-initial H tone in a sentence with only H tones to the pitch level of a non-initial H tone in a sentence with alternating LH/HL tones. Additionally, the amount of pitch drop between a non-initial H tone in a sentence with only H tones and a later H tone can be compared to the pitch drop calculated between a non-initial H tone in a sentence with alternating LH/HL tones and a later H tone. The same procedure will be applied to L tones. The material used to test the difference between declination and downstep is presented in table 13. An example of the long sentence containing only L tones and of a long sentence containing only H tones is illustrated in (79)a. and b., respectively. (79)c. presents an example of the long sentence containing alternating LH tones. The complete list of materials is introduced in chapter 2, section 2.2, blocks A and B.

[39] The greater value for the last pitch drop in table 7 is likely to be attributed to final lowering, which will be elaborated in chapter 6.

Tone/Length	H	L	LH	HL
7 σ’s	(42)b.	(43)a.		(45)b.
10 σ’s	(42)c.	(43)c.	(44)b.	
12 σ’s	(42)d.	(43)d.	(44)c.	

table 13: Material used to test the difference between declination & downstep.

(79) a. Wɔ̀fà Àsàrè fì Àkyèm̀fò. (43)c.
uncle proper name originate.HAB proper name
‘Uncle Asare comes from Akyemfo.’

b. Kúkúó-bá pápá nó bɔ́ dáá. (42)c.
pot-DIM good DEF break.HAB daily
‘The good small pot breaks everyday.’

c. Pàpá Kòfí kàsá kyèrɛ́ nè bá. (44)b.
father proper name talk.HAB point_to.HAB PRO child
‘Father Kofi talks to his child.’

Planned comparison: In a first step, the pitch level will be investigated. The height of H3 and H5, see (80), in sentences with only H tones will be compared to the height of H2 and H3 in sentences with alternating HL tones, respectively. In a second step, the pitch drop between H3 and H5 (H3-H5) in sentences containing only H tones will be compared to the pitch drop between H2 and H3 (H2-H3) in sentences with alternating HL tones. Initial values were not considered because they do not show downstep, see section 4.2, and final values were not included since they are subject to final lowering; see chapter 6 for details.

```
(80)   H1 H2  H3 H4  H5 …        vs.     H1  L1  H2 L2  H3 …
        |   |    |   |    |                  |   |    |   |    |
       CVCV  CVCV CV                     CVCV  CVCV CV
```

The same method will be applied to the L tones. The pitch level will be investigated by comparing the height of L2 and L4, see (81), in sentences with only L tones to the height of L1 and L2 in sentences with alternating HL tones, respectively. Further, the pitch drop between L2 and L4 (L2-L4) in sentences with only L tones will be compared to the pitch drop between L1 and L2 (L1-L2) in sentences with alternating HL tones.

```
(81)   L1  L2  L3 L4  L5… vs.     H1 L1  H2  L2  H3…
        |    |   |   |    |            |   |    |   |    |
       CVCV CVCV CV                CVCV CVCV  CV
```

In the long and longest sentences, the pitch level will be investigated by comparing the height of H4 and H6, see (82), in sentences with only H tones to the height of H2 and H3 in sentences with alternating LH tones, respectively. Further, the pitch drop between H4 and H6 (H4-H6) in sentences with only H tones will be compared to the pitch drop between H2 and H3 (H2-H3) in sentences with alternating LH tones.

```
(82)    H1 H2   H3 H4 H5 H6…      vs.      L1 H1   L2  H2  L3 H3…
         |   |    |   |    |   |                |   |    |   |    |   |
        CVCV  CVCV CVCV                   CVCV  CVCV CVCV
```

The same method will be applied to the L tones. The pitch level will be investigated by comparing the height of L3 and L5, see (83), in sentences with only L tones to the height of L2 and L3 in sentences with alternating LH tones, respectively. Further, the pitch drop between L3 and L5 (L3-L5) in sentences with only H tones will be compared to the pitch drop between L2 and L3 (L2-L3) in sentences with alternating LH tones.

.

```
(83)    L1  L2   L3 L4  L5 L6…      vs.      L1 H1  L2  H2  L3 H3…
         |   |    |   |    |   |                |   |    |   |    |   |
        CVCV  CVCV CVCV                   CVCV  CVCV CVCV
```

Again, Speaker 2 was discarded from the analysis of sentences with only H tones because of falsetto voice.

Results: The results will be presented for L and H tones separately, since sentence length has been shown to influence the declination rate in sentences with only H or only L tones. The figure 48 presents the mean pitch level of the L tones in the medium sentences consisting of only L tones (solid line) and the mean pitch level of L tones in the medium sentences with alternating HL tones (dashed line), aggregated over speakers and repetitions. The F0 decline of the L tones is slightly greater in sentences with alternating tones than in sentences that contain only L tones. The mean F0 at the beginning of the sentence is higher in the sentences with alternating tones than in the sentences containing only L tones.

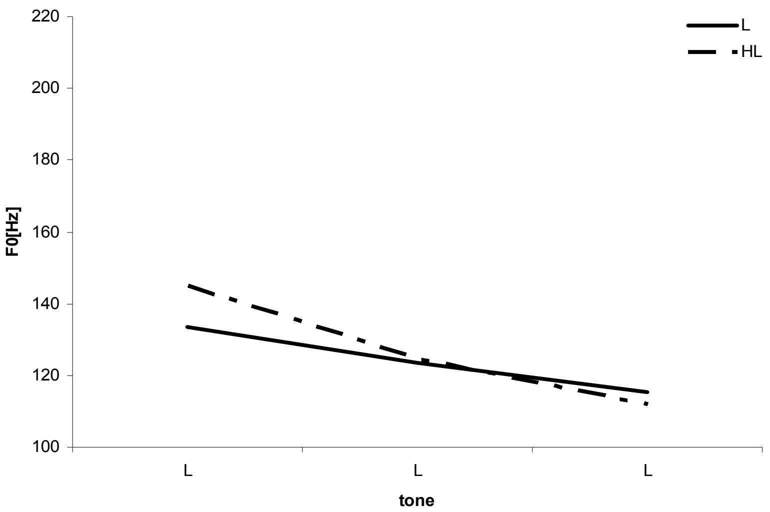


figure 48: Mean F0 of L2, L4 and L6, aggregated over speakers and repetitions, of the medium sentences containing only L tones (solid line), *Asare fi Akyemfo.* ‘Asare comes from Akyemfo.’; Mean F0 of L1, L2 and L3, aggregated over speakers and repetitions, of the medium sentences with alternating HL tones (dashed line), *Anane bisa sika*. ‘Anane asks for money.’; n=6.

The figure 49 displays the mean pitch level of the L tones in the longest sentences consisting only L tones (solid line) and the mean pitch level of L tones in the longest sentences with alternating LH tones (dashed line), aggregated over speakers and repetitions. For the longest sentences declination of the L tones does not differ as a function of the tonal environment. The values for the L tones in sentences containing alternating tones are overall higher (except the last tone) than the L tones in the only L sentence, presumably due to L raising, see chapter 3 section 3.1. This effect is to a lesser extent also present in the medium sentence.

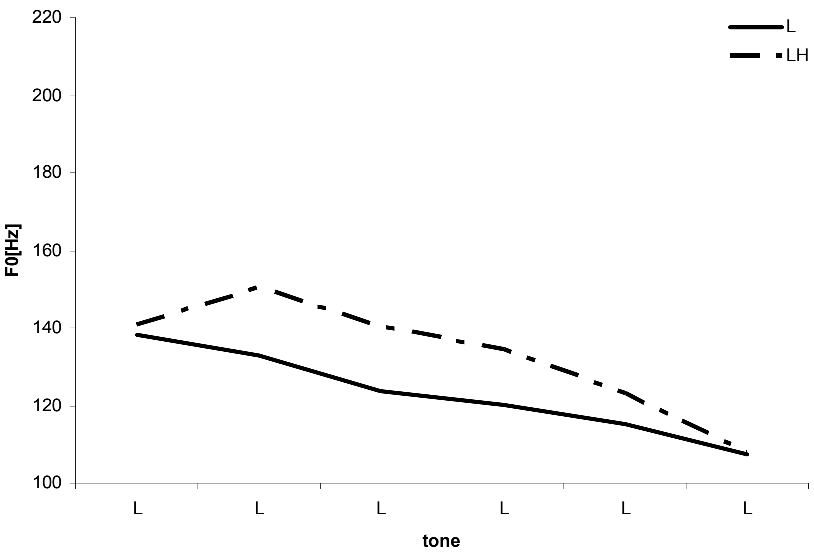


figure 49: Mean F0 of L1, L3, L5, L7, L9 and L12, aggregated over speakers and repetitions, of the longest sentences containing only L tones (solid line), *Wɔfa Ado Asare fi Akyemfo.* ‘Unlce Ado Asare comes from Akyemfo.’; Mean F0 of L1, L2, L3, L4, L5 and L6, aggregated over speakers and repetitions, of sentences with alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba bio.* ‘Father Kofi asks for his child again.’; n=6.

The figure 50 presents the mean pitch level of the H tones in the medium sentences with only H tones (solid line) and of the H tones in the medium sentences with alternating HL tones (dashed line), aggregated over speakers and repetitions. The declination of the H tones does not differ as a function of the tonal environment. The initial H tone value is higher in the alternating tone environment due to H raising; see chapter 3 section 3.2.

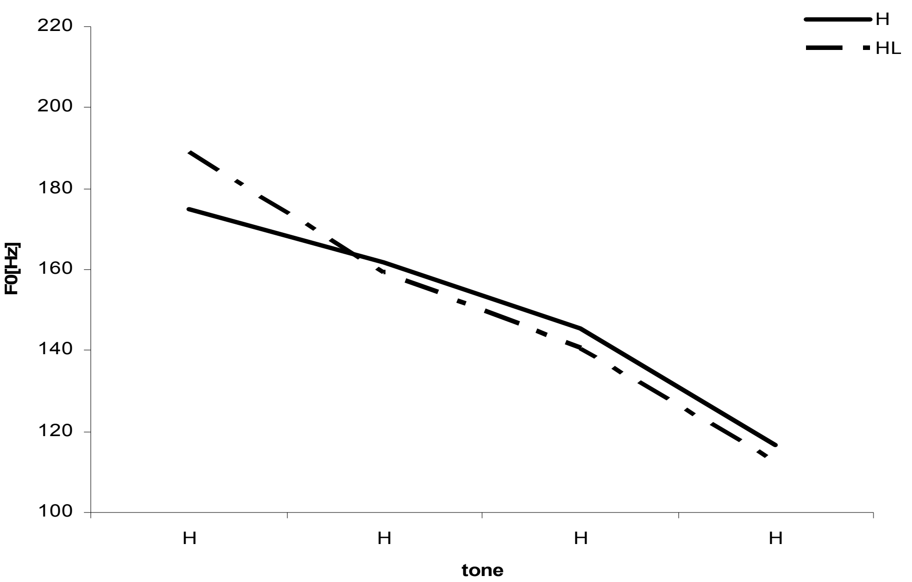


figure 50: Mean F0 of H1, H3, H5 and H7, aggregated over speakers and repetitions, of the medium sentences containing only H tones (solid line), *Kukuoba papa no.* 'The good small pot.'; Mean F0 of H1, H2, H3 and H4, aggregated over speakers and repetitions, of the medium sentences with alternating HL tones (dashed line), *Anane bisa sika.* 'Anane asks for money.'; n=5.

The figure 51 presents the mean pitch level of the H tones in the longest sentences with only H tones (solid line) and of the H tones in the longest sentences with alternating HL tones (dashed line), aggregated over speakers and repetitions. In the longest sentences the H raising effect is also present. It affects the first three H tones. Declination rate is clearly higher in sentences with alternating tones than in sentences with only H tones. F0 in sentences with only H tones in this condition has been shown to be subject to a length dependent declination rate decrease effect; see section 4.3.1.

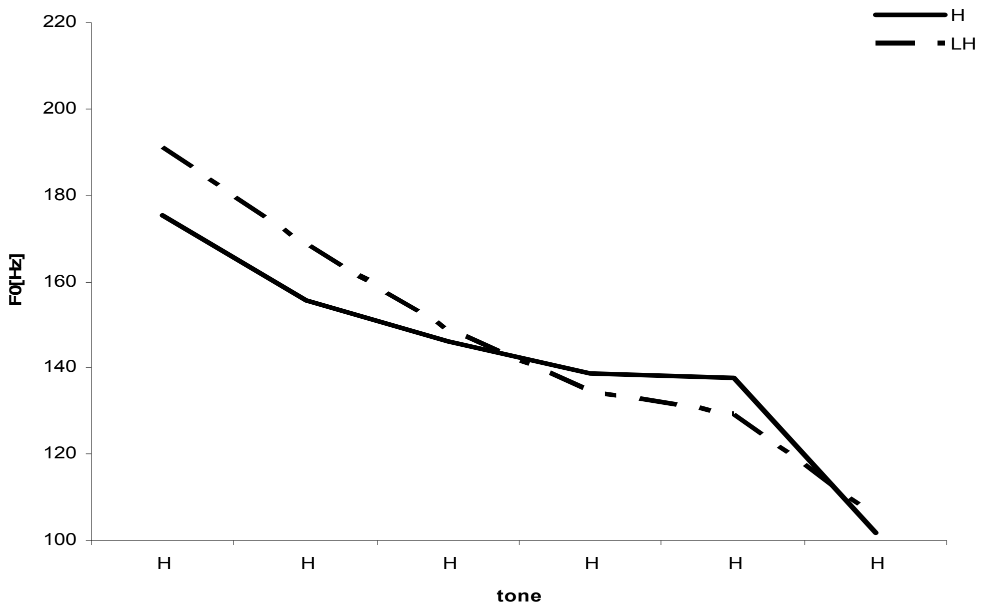


figure 51: Mean F0 of H2, H4, H6, H8, H10 and H12 aggregated over speakers and repetitions, of the longest sentences containing only H tones (solid line), *Kukuoba papa paa no bɔ daa.* 'The very good small pot breaks everyday.'; Mean F0 of H1, H2, H3, H4, H5 and H6 aggregated over speakers and repetitions, of the longest sentences with alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba bio.* 'Father Kofi talks to his child again.'; n=5.

The table 14 presents an overview of the mean F0 values of the selected L tones and the pitch drop between the selected L tones, aggregated over speakers and repetitions, in sentences with only L tones and sentences with alternating (LH/HL) tones; split by sentence lengths.

Tone/ Length	only L	LH	HL
7 σ's	L2 = 133.75 Hz (29) L4 = 120.79 Hz (27) drop = 12.96 Hz (5.4)		L1 = 144.33 Hz (30) L2 = 131.68 Hz (25) drop = 12.65 Hz (6.37)
10 σ's	L3 = 144.7 Hz (35) L5 = 134.64 Hz (32) drop = 10.06 Hz (4.5)	L2 = 144.38 Hz (31) L3 = 133.74 Hz (28) drop = 10.64 Hz (4.37)	
12 σ's	L3 = 133.01 Hz (31) L5 = 123.78 Hz (28) drop = 9.22 Hz (4.28)	L2 = 148.7 Hz (28) L3 = 139.6 Hz (25) drop = 9.1 Hz (3.62)	

table 14: Mean F0 values of the selected L tones plus standard deviation and mean pitch drop between the selected L tones, aggregated over speakers and repetitions for sentences with only L tones and sentences with alternating (LH/HL) tones; n=6.

The table 14 shows that the pitch level of earlier L tones is overall higher than that of later L tones in both tonal configurations. This is illustrated in figure 52. The figure 52 provides speaker-specific information of the earlier tones L2/L3 in sentences containing only L tones and L1/L2 in sentences containing alternating tones. Position 1 refers to the early tones.

Position 2 in figure 52 refers to the later tones L4/L5 in sentences containing only L tones and L2/L3 in sentences containing alternating tones. All values are aggregated over repetitions, lengths and tonal configurations. All speakers realize the later L tones lower than the early ones.

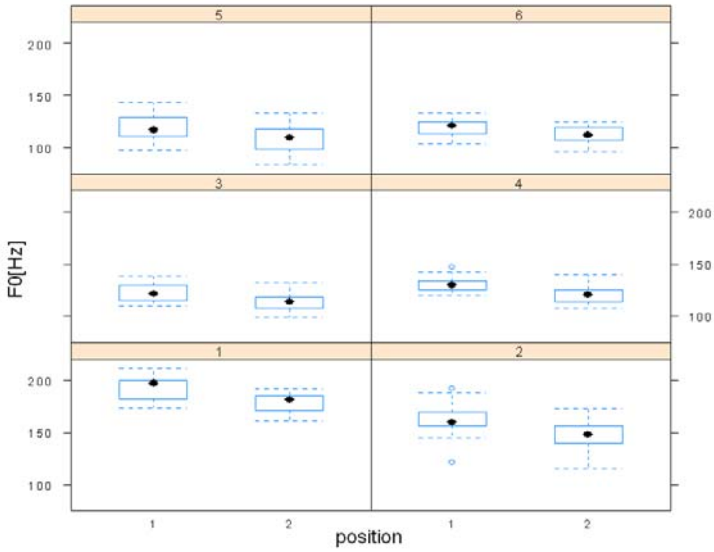


figure 52: Boxplot of pitch level of L tones, aggregated over repetitions, lengths and tonal configurations, split by positions (1 = early)/(2 = late), and speaker.

Furthermore, table 14 reveals that L tones are overall higher in alternating environments. The effect has been elaborated in chapter 3 section 3.1 under the heading L raising. However, the mean F0 values for the long rendition do not show L raising. The figure 53 provides information of the pitch level of the L tones in the long rendition for each tonal configuration, L refers to only L tones and LH refers to alternating tonal configuration, aggregated over repetitions and positions, for all speakers. The figure 53 shows that the lack of L raising in the mean value of the long sentences containing alternating tones in table 14 is due to speakers 2 and 5. All other speakers show the L raising effect in sentences containing alternating tones, as in the rest of the data set, but speakers 2 and 5 show an effect into the opposite direction.

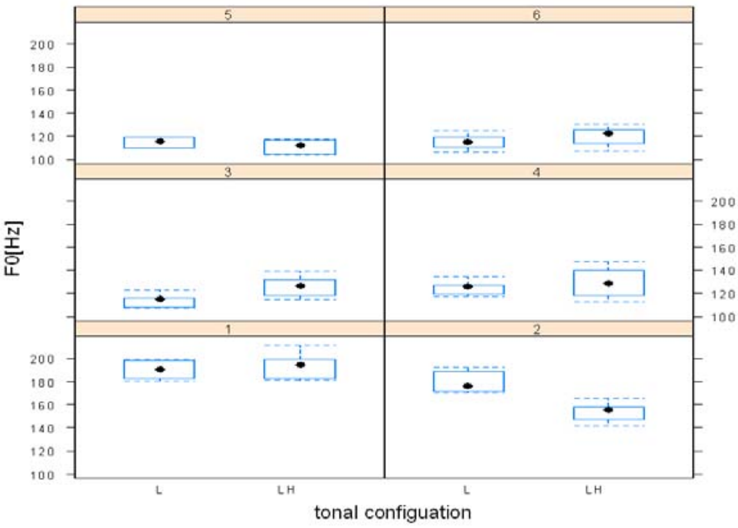


figure 53: Boxplot pitch level of L tones in long sentences, aggregated over repetitions and positions, split by tonal configuration (L = only L)/(LH= alternating) and speakers.

The table 14 further shows that the amount of pitch drop between the L tones is stable, irrespective of the tonal configuration. The figure 54 provides information of the pitch drop in sentences with only L tones (L) and in sentences with alternating tones (LH/HL), aggregated over repetitions and lengths, for each speaker. The majority of speakers show a similar amount of pitch drop in sentences with only L tones and sentences with alternating tones. Only the data of speaker 1 exhibits a greater pitch drop in the alternating environment.

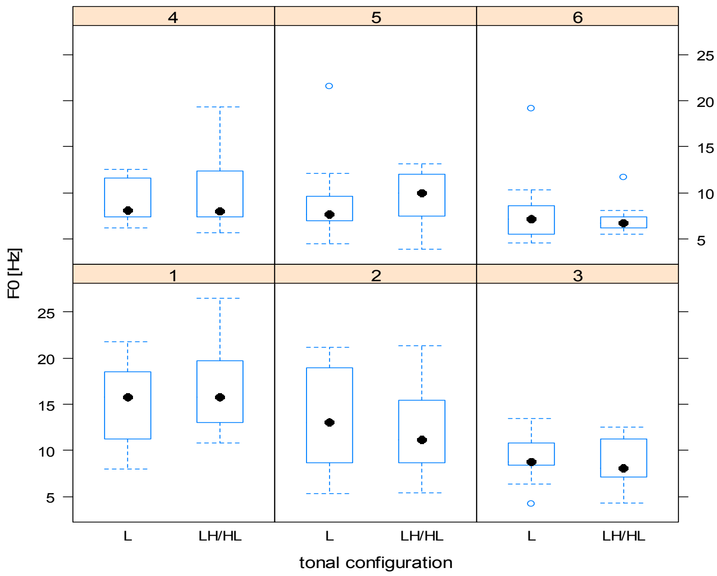


figure 54: Boxplot of pitch drop of L tones, aggregated over repetitions and lengths, split by tonal configuration (L = only L)/(LH/HL = alternating tones) and speakers.

Moreover, table 14 shows that the amount of pitch drop between the L tones gradually decreases with increasing length in both tonal configurations. The figure 55 provides information of the pitch drop in relation to sentence length, aggregated over repetitions and tonal configurations, for each speaker. The majority of speakers show a gradual decrease with increasing length. However, speakers 5 and 6 show a slightly higher pitch drop value for the longest renditions compared to the long ones.

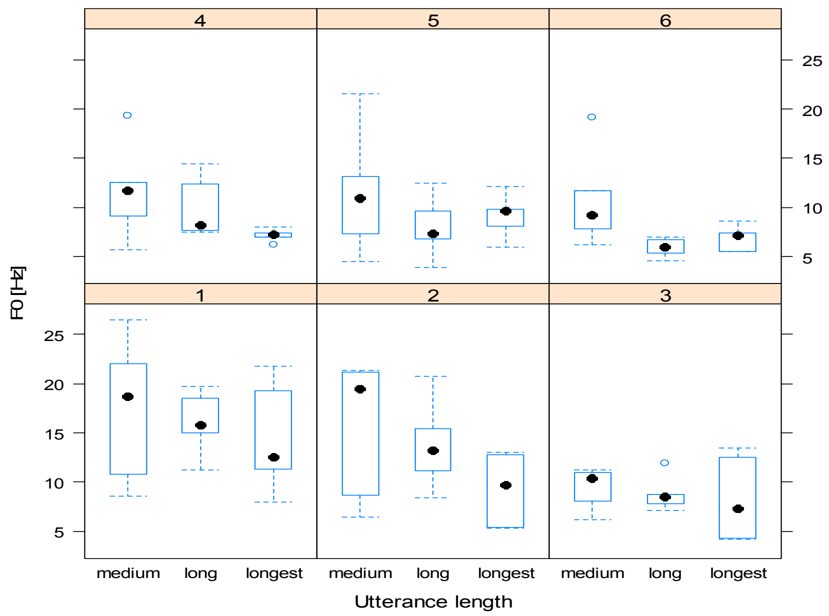


figure 55: Boxplot of pitch drop of L tones, aggregated over repetitions and tonal configuration, split by utterance lengths (medium = 7σ's, long = 10σ's, longest = 12σ's) and speaker.

Turning to the numerical results of the pitch level of the H tones and pitch drop between the H tones, the table 15 presents an overview of the mean F0 values of the selected H tones and the pitch drop between the selected H tones, aggregated over speakers and repetitions, for sentences with only H tones and sentences with alternating (LH/HL) tones; split by sentence lengths.

Tone/ Length	**only H**	**LH**	**HL**
7 σ's	H3 = 161.97 Hz (36) H5 = 143.5 Hz (35) drop = 18.47 Hz (6.4)		H2 = 159.31 Hz (38) H3 = 140.76 Hz (30) drop = 18.55 Hz (9.7)
10 σ's	H4 = 167.88 Hz (44) H6 = 151.29 Hz (46) drop = 16.59 Hz (8.83)	H2 = 157.51 Hz (46) H3 = 141.34 Hz (37) drop = 16.17 Hz (9.36)	
12 σ's	H4 = 162.96 Hz (42) H6 = 150.39 Hz (36) drop = 12.57 Hz (5.83)	H2 = 162.41 Hz (35) H3 = 145.26 Hz (29) drop = 17.15 Hz (8)	

table 15: Mean F0 values for the selected H tones plus standard deviation and mean pitch drop between the selected H tones, aggregated over speakers and repetitions for sentences with only H tones and sentences with alternating (LH/HL) tones; n=5.

The table 15 shows that the pitch level of the earlier H tones is overall higher than that of later H tones in both tonal configurations. This is illustrated in figure 56. The behavior is reminiscent of that observed for the L tones. The figure 56 provides speaker-specific information of the earlier tones H3/H4 in sentences containing only H tones and H2 in sentences containing alternating tones. Position 1 refers to the early tones. Position 2 in figure 56 refers to the later tones H5/H6 in sentences containing only H tones and H3 in sentences containing alternating tones. The values are aggregated over repetitions, lengths and tonal configurations, for each speaker. All speakers realize the later H tones lower than the earlier H tones.

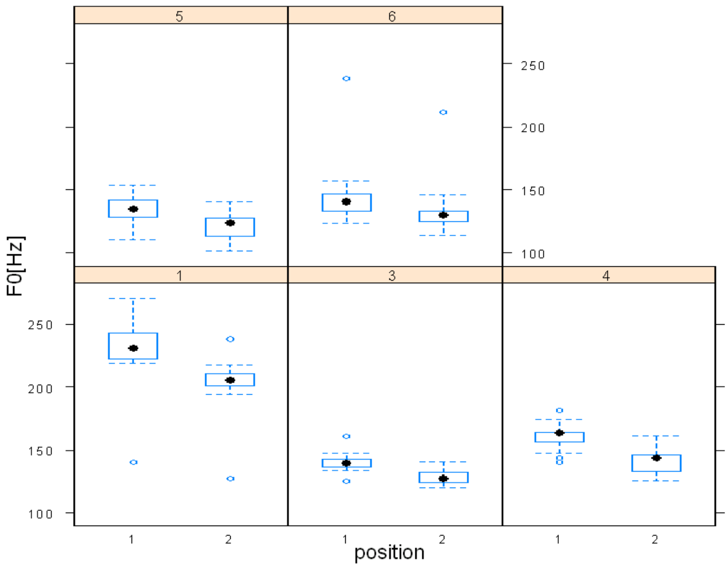


figure 56: Boxplot of pitch level of H tones, aggregated over repetitions, lengths and tonal configurations, split by positions (1 = early)/(2 = late) and speakers.

Furthermore, table 15 reveals that, contrary to the situation found for L tones, the H tones in alternating environments are not overall higher although H raising is present, at least on initial H tones; see chapter 3 section 3.2. As a result, the pitch level for early and late tones does not differ as a function of tonal configuration for the medium and longest sentences. However, for the long renditions the mean values for both tones are about 10 Hz lower in the alternating environment compared to the sentences with only H tones. The figure 57 provides information of the F0 height of the H tones in the long rendition for each tonal configuration. H refers to only H tones and LH refers to alternating tonal configuration, aggregated over repetitions and positions; for all speakers. The effect observed in the means is mainly due to the values of speaker 4 and 5. All other speakers do not show a clearly lower value in the alternating environment.

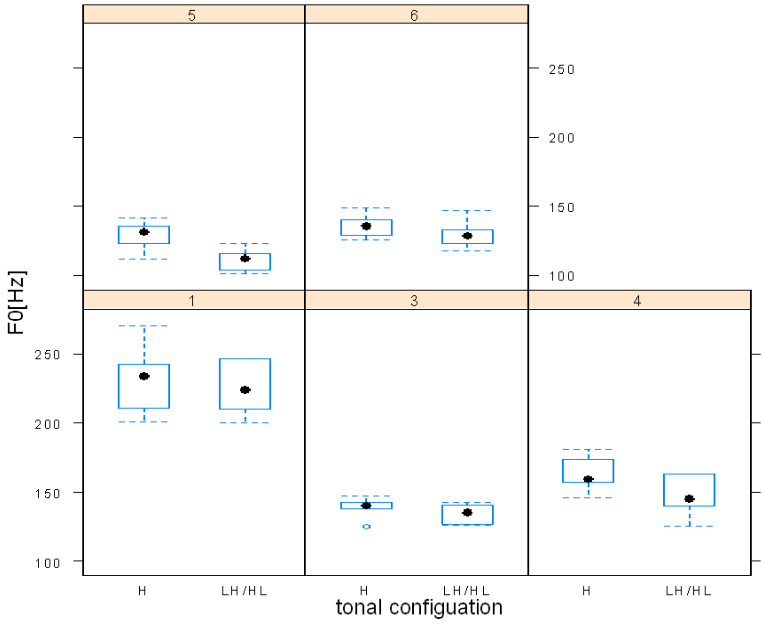


figure 57: Boxplot of pitch level of H tones in long sentences, aggregated over repetitions and positions, split by tonal configuration (H = only H)/(LH= alternating) and speakers.

The table 14 shows that amount of pitch drop between the selected H tones is stable across the conditions (only H tones vs. alternating tones) for the medium and the long renditions. However, the longest sentences show a greater amount of pitch drop for the H tones in alternating environments. The figure 58 compares the pitch drop between the H tones of the medium and long sentences to the longest sentences. It further gives speaker-specific information. The boxplot on the left hand side of figure 58 provides data on the pitch drop in sentences with only H tones (H) and in sentences with alternating tones (LH/HL), aggregated over repetitions and lengths (medium & long), for each speaker. The majority of speakers exhibit a similar amount of pitch drop in sentences with only H tones and sentences with alternating tones. Speaker 1 exhibits an overall greater pitch drop in the alternating environment, whereas speakers 5 and 6 show a tendency into the opposite direction. The boxplot on the right hand side of figure 58 provides information of the pitch drop in sentences with only H tones (H) and in sentences with alternating tones (LH/HL), aggregated over repetitions, of the longest renditions for each speaker. All speakers show a greater pitch drop in the alternating LH environment than in the sentences containing only H.

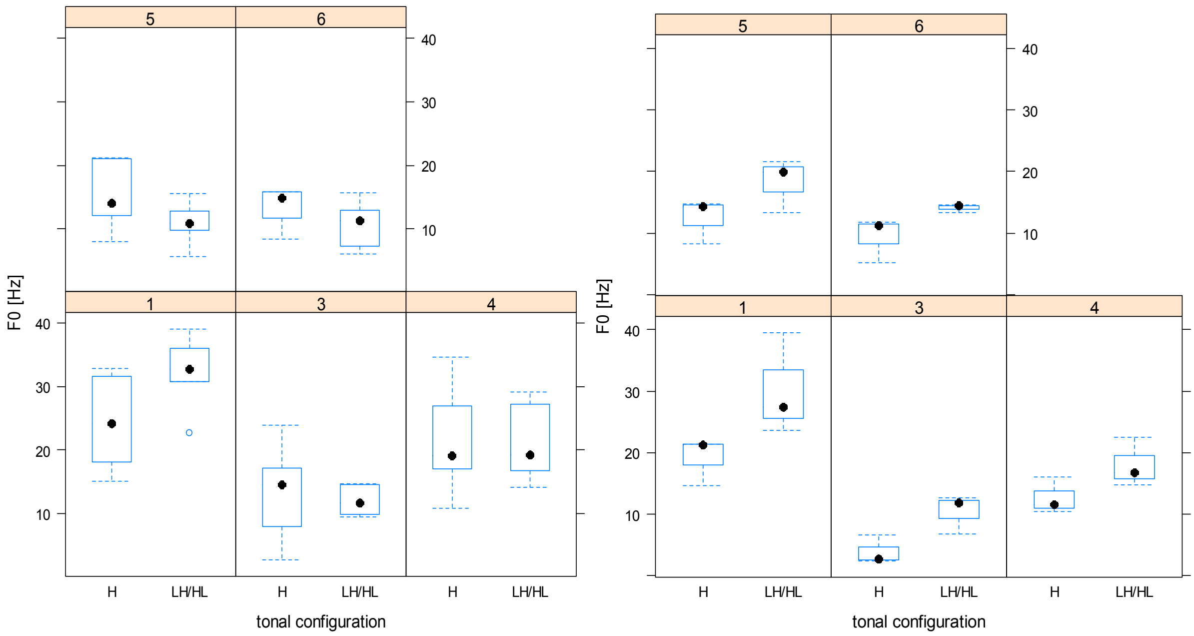


figure 58: Left boxplot: pitch drop of H tones, aggregated over repetitions and lengths (medium & long), split by tonal configuration (H = only H)/(LH/HL = alternating tones) and speakers; Right boxplot: pitch drop of H tones for the longest sentences.

The table 15 further shows that the amount of pitch drop gradually decreases with increasing length for the sentences with only H tones. However, for the alternating environment a decrease can be observed for the mean F0 of the medium length to the long sentences but not for the longest renditions. The figure 59 provides speaker-specific information of the amount of pitch drop between the H tones for each sentence length and each tonal configuration. The boxplot on the left hand side of figure 59 provides information of the pitch drop in relation to sentence length in sentences with only H tones (H), aggregated over repetitions, for each speaker. Except speaker 1, all speakers show a gradual increase in pitch drop with increasing length. The boxplot on the right hand side of figure 59 provides information of the pitch drop in relation to sentence length in sentences with alternating tones, aggregated over repetitions, for each speaker. A gradual decrease with increasing length can be observed for speakers 1 and 3. Speaker 4 shows a decrease in pitch drop between the medium and long/longest rendition. The overall higher mean value for the longest renditions in table 15 seems to be attributed to the higher values of speakers 5 and 6.

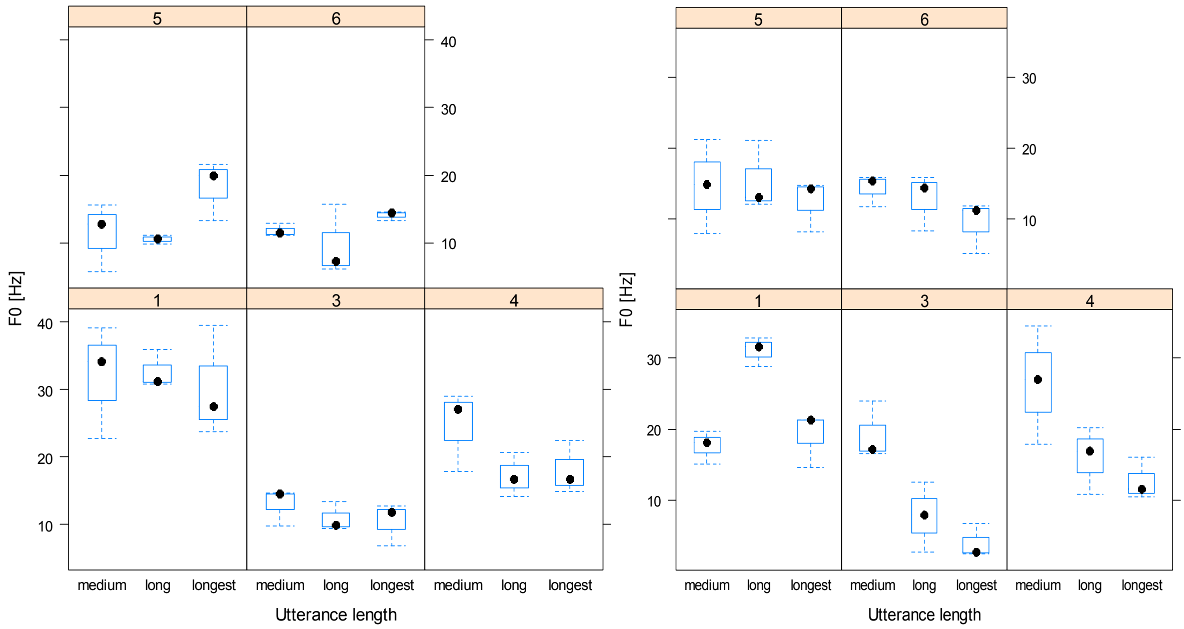


figure 59: Left boxplot: pitch drop of H tones in sentences with only H tones, aggregated over repetitions, split by length (medium = 7σ's, long = 10σ's, longest = 12σ's) and speakers. Right boxplot: pitch drop of H tones in sentences with alternating LH/HL tones.

A linear mixed effects model was calculated on the pitch level. Tone (L/H), position of the tone, tonal configuration (only L, H/alternating LH, HL) and length (medium, serves as baseline for the contrast calculation/long/longest) were treated as fixed factors. Repetitions and speakers were included as random factors. The tonal specification had a significant effect on the pitch level (t = -16.200, pMCMC < 0.0001, 358 observations), which means that H tones are systematically realized higher than L tones.

With regard to the position of the tone, the statistic analysis has revealed that it affects the pitch level systematically (t = 10.027, pMCMC < 0.0001); earlier tones are overall higher than later tones. There was a significant interaction between tone and position (t = - 2.335; pMCMC < 0.05). The interaction is significant because the difference between the tones is more pronounced in the earlier position than in the late position.

The factor tonal configuration affects the pitch level of the tones in a significant way (t = -2.927, pMCMC < 0.005); the pitch level is generally lower in sentences with only H and only L tones compared to the pitch level in sentences containing alternating tones.

No significant result was obtained for the effect of the sentence length on the pitch level. However, the interaction between tonal configuration and the contrast medium vs. long approached significance (t = -1.962, pMCMC = 0.0565); in the medium rendition no difference in pitch level is present, whereas the F0 is slightly higher for the alternating sequences than for the only L/only H sentences in the long renditions. Moreover, the interaction between tonal configuration and the contrast medium vs. longest yielded a significant result (t = 3.614, pMCMC < 0.001); the pattern found in the medium renditions is more pronounced in the longest renditions.

A linear mixed effects model was calculated on the pitch drop. Tone (L/H), tonal configuration (only L, H/alternating LH, HL) and length (medium, serves as baseline for the contrast calculation/long/longest) were treated as fixed factors. Repetitions and speakers were included as random factors. Tone had a significant effect (t = -8.496, pMCMC < 0.0001, 197 observations); the amount of pitch drop is generally greater for H tones than for L tones. The factor tone did not interact with any other factor.

Tonal configuration did not yield a significant result. There were also no significant interactions. The length contrast medium versus long was significant (t = 3.984, pMCMC < 0.0001); the amount of pitch drop is greater for the sentences with medium length than for the long renditions.

All other calculation did not reveal any significant results.

Summary: Generally, the data of the pitch level has shown that later tones are always lower than earlier tones. The effect was robust for L tones in sentences with only L tones and for H tones in sentences with only H tones as well as for both tones in alternating sequences; see figures 51 and 55. This was expected from the results obtained for declination presented in section 4.3.1. The expectation that pitch lowering is generally more pronounced in H tones than in L tones, nourished from the outcome of section 4.3.1, was fulfilled. The amount of pitch drop is overall higher for H tones in sentences with only H tones and sentences with alternating tones than for L tones in the same environments. Irrespective of the lowering process, tonal contrasts are well preserved (Huang, 1985).

The investigation of the pitch level further revealed that L tones in sentence with alternating tones were systematically realized higher in alternating tone environment than sentences with only L tones. The outcome was expected from the findings on L raising presented in chapter 3 section 3.1. Although L raising is present, the lowering between the L tones in alternating sequences, measured as pitch drop, equaled the lowering in sentences with only L tones; see figure 54. The amount of pitch drop between the L tones presentend in table 14 is slightly lower than that reported by Dolphyne (1994); 10 Hz vs. 15 Hz.

The exploration of the pitch level of H tones did not provide evidence for a systematic additional lowering effect (downstep and/or carry-over lowering) in the alternating environment compared to the only H environment; see figure 56. The amount of pitch drop in the alternating environment is comparable to that found in only H sentences for the sentences with medium and long length; see figure 58 left plot. Overall, the pitch drop between the H tones in the data presented above is higher than that reported by Dolphyne (1994), 16 Hz vs. 10 Hz. For the longest rendition the pitch drop was overall greater in alternating environments than in sentences with only H tones; see figure 58 right plot. This is potentially due to a more pronounced declination rate decrease in the longest rendition of the only H sentences; see section 4.3.1 for details.

Furthermore, sentence length had an effect on the pitch drop on both L and H tones; see figure 55 and figure 59. The amount of pitch drop decreases with increasing length, an effect which is reminiscent of that reported for declination rate in section 4.3.1. The effect of sentence length will be elaborated further in chapter 5, section 5.2.

Discussion: The question of this section was whether the lowering of L and H tones in alternating LH/HL sequences is greater than the lowering which can be attributed to declination, i.e. whether declination and downstep are different phenomena in Akan. The question has been empirically tackled by comparing the pitch level and the pitch drop of H and L tones in sentences with only H or only L tones to the pitch level and the pitch drop of H and L tones in sentences with alternating tones; see e.g. (82). The evaluation of the data has shown that the lowering of H tones in sentences with alternating tones equals the lowering of H tones in sentences with only H tones. This is schematically illustrated in (84); the arrow indicates lowering. The illustration is reminiscent of the conclusion drawn in the discussion of section 4.2, inspired by Huang (1985), that the lowering pattern between H tones in alternating environments (downstep) is best characterized as the lowering of an H tone in relation to a preceding H tone.

(84)			
	H L H | | | CVCVCV	=	H H H | | | CVCVCV

Furthermore, the data has revealed that lowering of L tones in sentences with alternating tones equals the lowering of L tones in sentences with only L tones. This is schematically illustrated in (85).

(85)			
	L H L | | | CVCVCV	=	L L L | | | CVCVCV

Thus, the pattern in alternating tonal sequences known as downstep or tone terracing is caused by declination of L and H tones, i.e. no additionally lowering principle is at work. Hence, there is no need to assume a different phonological representation or a phonological trigger to represent downstep; see (69) for a proposal involving register tones which associate to tonal feet. A phonological surface representation of a declarative sentence with register tones which associate to the IP has been proposed in (77). The results of this section have revealed that a sentence containing alternating tones has exactly the same phonological representation as a sentence containing only H or only L tones, see (77), because downstep is not an independent phonological process.

The results for L and H tones, presented in this section, support Liberman et al.'s (1992) observation for Igbo that downstepped H and L tones seem to behave identically and add empirical support to Liberman and Pierrehumbert's (1984:219) speculation that the scaling of L tones is symmetric to that of H tones. Additionally, section 4.1 has revealed that the lowering of the second H tone in an HLH sequence (automatic downstep) equals the lowering of the second H tone in an HL̥H sequence (non-automatic downstep), in which the L tones is not phonetically realized. This is schematically illustrated in (86).

```
(86)   H  L  H    =    H  L̥  H
       |  |  |         |     |
       CVCVCV          CV    CV
```

4.3.4 The implementation of downstep

Turning to the question how alternating tones can be phonetically implemented, the pitch algorithm has to be upgraded in such a way that it is able to cope with the empirical observations illustrated in (84)-(86). In section 4.3.1, it has been shown that declination can be modeled by the formula in (71). The algorithm proceeds from left-to-right with one tone look-ahead (Pierrehumbert, 1980). It calculates F0 values for the second tone (T2) from the first tone (T1) and the third tone (T3) from T2 and so forth as illustrated by the arrows in (87); the reference value of the initial tone serves as input value.

```
(87)   T1 T2 T3 T4 T5 T6
       |  |  |  |  |  |
       CVCV  CVCV  CVCV
```

Let us imagine we want to predict the F0 values for the tonal string in (88). The first tone is an H tone (H1) and receives a default value (H1 = 260 Hz) according to the speakers pitch range.

(88) H1 L1 H2 L2 H3
| | | | |
CVCV CVCV CV

The algorithm calculates the value of T2 on the basis of its input value in reference to *s* (0.76) and *r* (140 Hz).[40] The data has shown that the lowering process applies to tones of the same identity. A later H tone is lowered in relation to a preceding H tone and a later L tone in relation to a preceding L tone. Thus, the algorithm has to be made sensitive to the tonal identity. The algorithm needs to check whether the second tone in the tonal string in (88) is H or L. If T2 is H, the calculated value of H2 is assigned. However, in (88) T2 is L and since it is the first occurrence of its identity, it also receives a default value (L1 = 180 Hz). The algorithm keeps the value of H2 in its memory and calculates the value of T3 on the basis of H2 and the value of L2 on the basis of L1. Hence, for the third tone, the algorithm calculates values for both tonal entities. It checks whether the third tone is H or L and since it is H, H3 is assigned. The procedure is illustrated in (89).

(89)

Tonal string		calculation	assigned F0 value	memory
T1	H	H1 default	260 Hz	
T2	L	L1 default	180 Hz	
		H2-r = s*(H1-r)		H2= 231.2 Hz
T3	H	H3-r = s*(H2-r)	209.3 Hz	
		L2-r = s*(L1-r)		L2= 170.4 Hz
T4	L	L3-r = s*(L2-r)	163.1 Hz	
		H4-r = s*(H3-r)		H4= 192.7 Hz
T5	H	H5-r = s*(H4-r)	180.1 Hz	
		L4-r = s*(L3-r)		L4= 157.6 Hz

This procedure differs from Liberman & Pierrehumbert's (1984) approach for modeling automatic downstep in English. The lowering of a second H tone in an H*L H*L configuration is directly derived from the preceding H tone, see (71), i.e. without parallel computation of two possible F0 values as suggested for Akan. As already pointed out in the introduction, L tones are not explicitly captured by the model. They serve as trigger for downstep in the original proposal developed for English. The language-specific phonetic

[40] I would like to thank Mira Grubic and Himanshu Sharma for discussing the algorithm with me.

implementation model proposed for Akan is able to capture that tone terracing does not occur phrase-initially, see section 4.2, and that the amount of pitch drop for non-automatic downstep is similar to the amount of pitch drop for automatic downstep; see 4.1. Furthermore, it accounts for the similarity of declination and downstep and preserves the two basic ideas: the surface pattern of terracing is triggered by the presence of an intervening L tone between two H tones and "downstep" is relational.

In the following, it will be shown how the upgraded version of the phonetic implementation algorithm performs in predicting the terracing pattern in alternating tonal sequences. The table 16 provides information of the mean initial tone values for the sentences with alternating HL tones, aggregated over repetitions, *s* and *r*, as established in section 4.3.1, and R^2 for obtained vs. predicted F0 values for each speaker.

Speaker	Length	R	S	H1 in Hz	L1 in Hz	R^2
1	1	140	0.76	264.89	189.98	0.98
	2			300.35	206.86	1
	3			275.03	206.78	0.97
	4			302.99	220.87	0.91
2	1	90	0.77	223.75	134.12	0.98
	2			190.68	142.94	0.97
	3			237.01	160.07	0.97
	4			210.10	172.78	0.71
3	1	79	0.75	161.26	126.33	0.92
	2			161.01	128.58	0.91
	3			160.65	126.80	0.91
	4			167.16	139.09	0.84
4	1	75	0.8	161.41	117.02	0.94
	2			183.61	135.29	0.96
	3			176.8	131.66	0.97
	4			166.59	135.23	0.81
5	1	80	0.72	128.38	111.56	0.82
	2			147.51	131.58	0.82
	3			131.98	119.18	0.81
	4			138.5	128.72	0.74
6	1	78	0.7	141.15	121.35	0.94
	2			152.36	124.8	0.85
	3			163.17	140.31	0.84
	4			185.41	155.96	0.76

table 16: Reference values *r* and *s*; mean initial F0 (H1 & L1), aggregated over repetitions, for short, medium, long and longest sentences with alternating HL tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6.

The model is very successful in predicting the terracing pattern in sentences with HL alternating tones in Akan. The R^2 values range from 0.71 to 1. Perfect agreement between obtained and predicted values is obtained for the medium renditions of the female speaker; speaker 1. This is illustrated in figure 60 which presents obtained mean F0 values of the medium sentences (solid line), aggregated over repetitions, and predicted values (dashed line).

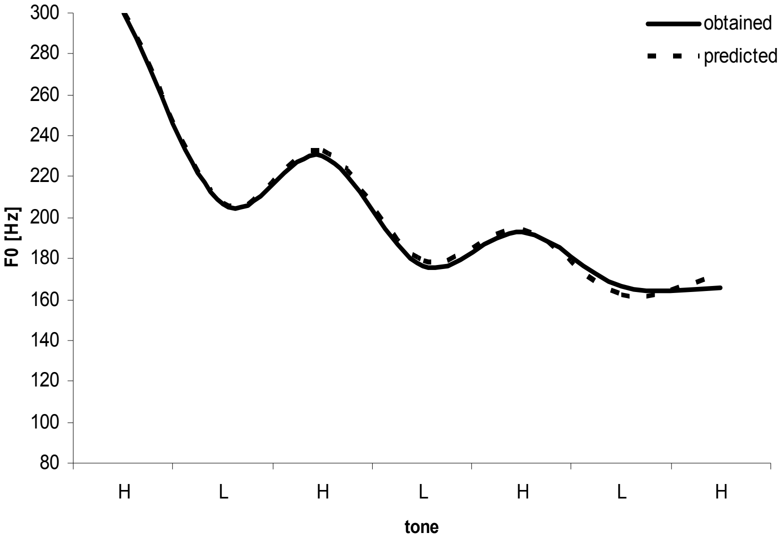


figure 60: Obtained F0 values (solid line) of the medium sentences containing alternating HL tones, *Anane bisa sika*. 'Anane asks for money.', aggregated over repetitions, and predicted F0 values (dashed line); female speaker.

The longest sentences of the male speakers show an overall lower quotient of determination. The figure 61 presents obtained mean F0 values, aggregated over repetitions, of the longest sentences (solid line) and predicted F0 values (dashed line) of speaker 4, who is representative for all male speakers. The predicted values are overall lower than the obtained values. This is presumably caused by two reasons.

First, the initial H tone value was not sufficiently raised in anticipation of the sentences length and hence the distance between initial H and L was not large enough to predict the F0 values with more accuracy. Anticipatory raising will be elaborated further in chapter 5, section 5.2. Second, unexpected changes in the underlying tone pattern occurred on the words *sìká* 'money' and bɔ̀né 'bad'; see Christaller, 1933:38; Dolphyne, 1988:122 and Kotey, 2009:51 for tonal specification. In isolation both words show up with LH tones. The algorithm treated the tonal sequences as an alternating one but infact *sika* was realized low throughout, whereas *bɔne* exhibits high pitch on the initial syllable and low pitch on the second. This is illustrated in (90). It seems that the male speakers interpreted this noun-adjective sequence as a compound, see Dolphyne (1988) and Marfo (2005) for further information on tonal changes in compounds.

(90) sìká – 'money' + bɔ̀né - 'bad' -> sìkà bɔ́né - 'dirty money/loot'

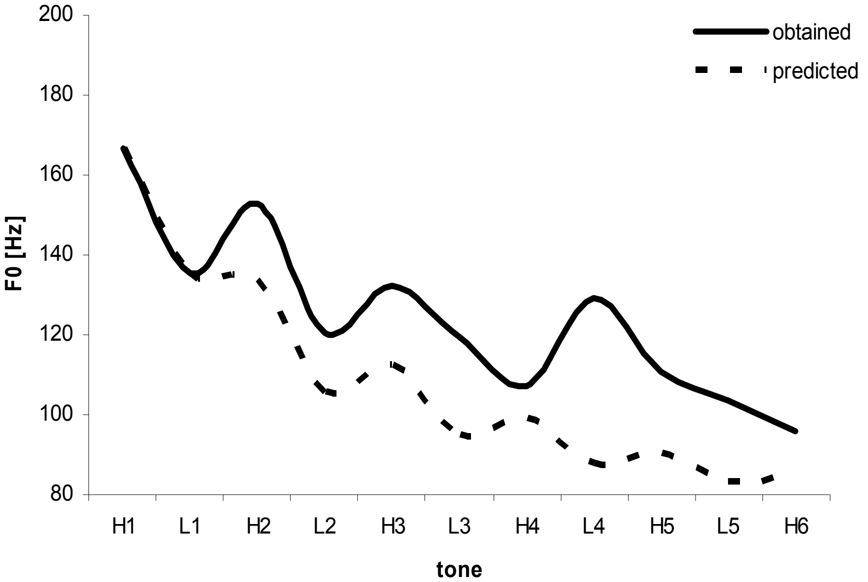


figure 61: Obtained F0 values of the longest sentence with alternating HL tones (solid line), *Anane bisa sika bɔne ɛnɛ*. 'Anane asks for money today.', aggregated over repetitions and predicted F0 values (dashed line); male speaker.

Turning to the prediction of the sentences with alternating LH tones, table 17 provides information of the mean initial tone values for the sentences with alternating LH tones, aggregated over repetitions, and the coefficient of determination for obtained vs. predicted F0 values for each speaker.

Speaker	Length	r	s	L1 in Hz	H1 in Hz	R^2
1	1	140	0.76	191.5	283.05	0.90
	3			208.35	298.86	0.94
	4			190.93	296.00	0.85
2	1	90	0.77	138.29	219.46	0.92
	3			149.25	254.22	0.85
	4			136.53	224.55	0.75
3	1	79	0.75	115.48	151.33	0.73
	3			120.19	164.60	0.79
	4			118.28	160.41	0.75
4	1	75	0.8	114.40	152.08	0.71
	3			111.43	198.49	0.88
	4			122.62	177.14	0.80
5	1	80	0.72	107.75(L2)	128.17	0.74
	3			104.10	144.23	0.72
	4			133.00	159.49	0.72
6	1	78	0.7	117.95	157.77	0.80
	3			122.11	152.47	0.78
	4			121.18	163.45	0.73

table 17: Reference values *r* and *s*; mean initial F0 (L1 & H1), aggregated over repetitions, for short, long and longest sentences with alternating LH tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6[41].

[41] Due to creakiness, no L initial value for speaker 5 could be obtained. L tone values were calculated from L2.

The R^2 values range from 0.71 to 0.94 which means that the model is quite successful in predicting the F0 values in alternating LH sentences. The coefficients of determination are slightly lower than for the HL alternating tone sequences. The figure 62 presents obtained mean F0 values (solid line), aggregated over repetitions, and predicted F0 values (dashed line) of the medium sentences of the male speaker 2, which is representative for all speakers. The L tone values are systematically precdicted lower than they appear in the obtained data. However, this is expected from the standpoint that L tone raising is a coarticulatory phonetic process, which was adopted in chapter 3 section 3.1. Furthermore as pointed out in chapter 3 section 3.1, the initial L tone in sentences with alternating LH tones is frequently realized lower than later L tones. In the alternating HL tones, the input value of the first L tones taken as reference value for the prediction was already raised. Therefore, the L tones in alternating HL sentences were predicted more accurately.

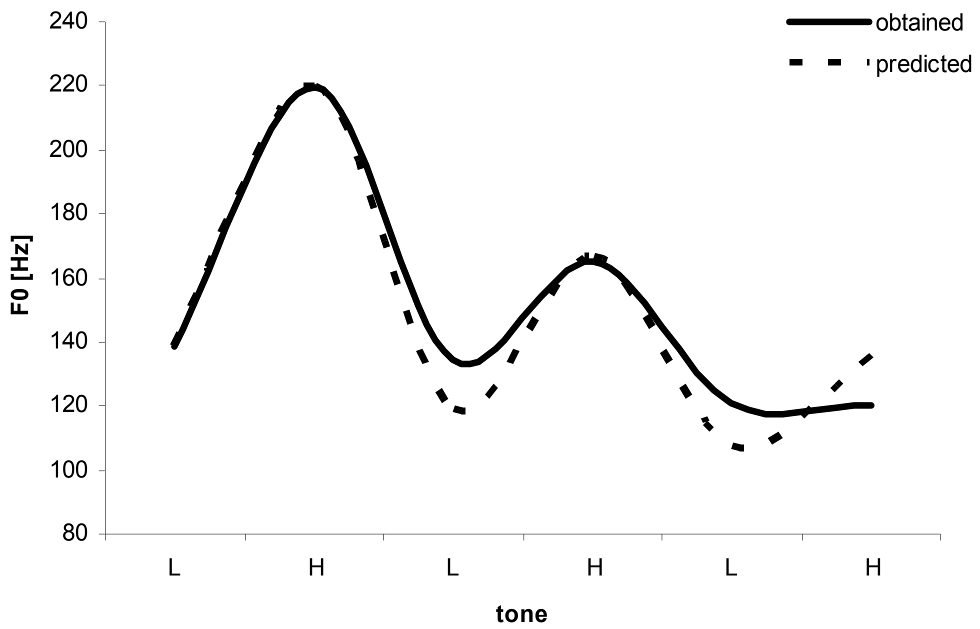


figure 62: Obtained F0 values (solid line), aggregated over repetition and predicted F0 values, for the medium sentence with alternating LH tones, *Papa Kofi kasa*. 'Papa Kofi talks.' and predicted F0 values (dashed line); male speaker.

Turning to the prediction of the sentences containing automatic and non-automatic downstep presented in section 4.1, table 18 provides information of the reference values, *r* and *s*, the mean initial F0 values (L1 & H1), aggregated over repetitions, and the coefficient of determination for obtained vs. predicted F0 values for each speaker.

Speaker	Downstep type	r	s	L1 in Hz	H1 in Hz	R^2
1	Automatic downstep	140	0.76	218.23	350.54	0.63
	Non-automatic downstep			214.27	340.20	0.62
2	Automatic downstep	90	0.77	163.20	273.97	0.69
	Non-automatic downstep			145.87	276.83	0.49
3	Automatic downstep	79	0.75	121.69	174.71	0.57
	Non-automatic downstep			121.67	177.59	0.54
4	Automatic downstep	75	0.8	127.73	186.52	0.71
	Non-automatic downstep			127.72	206.75	0.62
5	Automatic downstep	80	0.72	133.32	165.73	0.65
	Non-automatic downstep			134.68	165.93	0.60
6	Automatic downstep	78	0.7	115.91	175.80	0.55
	Non-automatic downstep			118.00	178.04	0.50

table 18: Reference values *r* and *s*; mean initial F0 (L1 & H1), aggregated over repetitions, for sentences containing automatic and non-automatic downstep and Pearson's R^2 for obtained vs. predicted F0 value; n=6.

The R^2 values reveal that the model is less successful in predicting the F0 values for these renditions. Overall, too low F0 values are predicted. This was expected for the L tones, since the first L tone in these renditions is sentence initial. But also H tones are frequently predicted too low, as illustrated in figure 63. The figure compares the obtained F0 values (solid line), aggregated over repetitions, to the predicted values (dashed line) for the sentences containing downstep (left plot) and for the sentences containing non-automatic downstep (right plot), for the female speaker. Note that for the sentences containing downstep, the tonal string used for the calculation involves an L tone between H2 and !H3 which is phonetically not realized (floating).

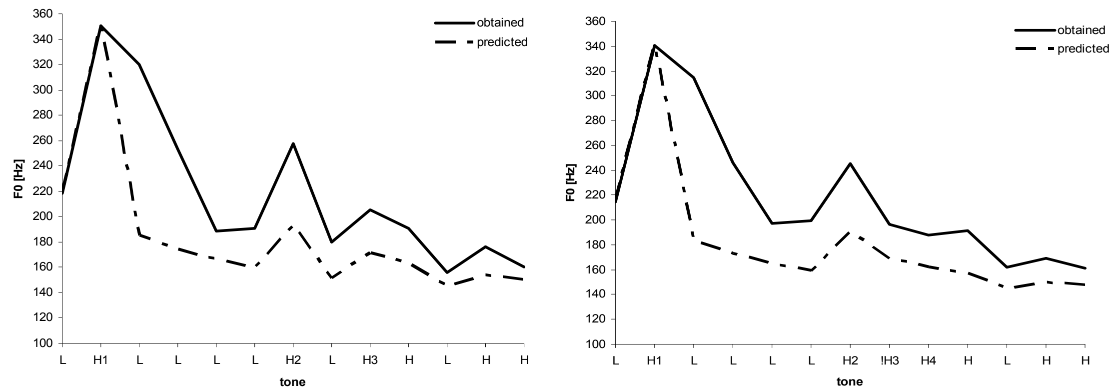


figure 63: Left plot obtained F0 values (solid line), aggregated over repetitions, *Afua hunu Kofi papa anɔpa yi.* 'Afua sees Kofi's father this morning.' and predicted F0 values (dashed line) for the sentences with automatic downstep; Right plot obtained F0 values (solid line), aggregated over repetitions, *Afua hunu Kofi dan anɔpa yi.* 'Afua sees Kofi's house this morning.' and predicted F0 values (dashed line) for the sentences with non-automatic downstep, female speaker.

However, the model predicts the phonetic similarity of automatic and non-automatic downstep which was experimentally established in section 4.1. Further research and testing of the model on different material has to show whether the algorithm always calculates two values for L and H or whether better results can be achieved by putting it on hold under specific circumstances e.g. if the tonal check was negative two times.

In what follows, I will summarize the most important findings of this chapter. In section 4.1 it has been shown with the help of controlled experimental data, that the lowering of the second H tone in a sequence of alternating LHLH tones in which the L tone is phonetically realized equals the amount of lowering that is caused by a floating L tone in the same environment. This finding led to the conclusion that automatic and non-automatic downstep are both automatic in the sense that they are caused by an underyling L tone which appears between two H tones. Furthermore, the lowering of an H tone by a phonetically not realized L tone points to the fact that downstep is not a coarticulatory effect in Akan.

In section 4.2, the absence of downstep in initial position was substantiated. The observation was taken as starting point to scrutinize phonological proposals of downstep that involve register tones that associate with tonal feet (Clements, 1983; Huang, 1985). Further, the findings were taken to add evidence to the proposal that downstep in Akan is not phonetic/coarticulatory and that it is relational. An H tone is lowered in reference to a preceding H tone. If an H tone is the first in a tonal string it is not lowered, irrespective of how many L tones precede the H tone. First tones are reference tones, which receive a default value (Liberman & Pierrehumbert, 1984).

Section 4.3.1 explored declination in sentences containing only H or only L tones. The data has shown that both tonal entities show a considerable amount of declination and that it is intra- and inner-speaker stable. It has been argued that declination is phonologized in Akan. Furthermore, a phonological representation of declarative Akan sentences involving register tones associating with the IP has been proposed, see (77). Section 4.3.2 dealt with the issue of how discretely represented phonological entities, lexical tones and register tones are mapped onto F0 targets in the phonetic implementation module. Declination in sentences containing only H or only L tones has been modelled as an exponential decay towards a non-zero asymptote (Liberman & Pierrehumbert, 1984; Shih, 2000). The lowering quotient and baseline value for each speaker have been established and the goodness of the model has been demontrated.

Section 4.3.3 investigated the relationship between declination and downstep. It has been shown declination and downstep are phonetically similar, by means of carefully selected measurement points. This observation led to the conclusion that downstep is not a separate phonological process but rather a term to refer to the lowering of H tones in sentences with alternating tones. Hence, the phonological representation introduced in (77) also captures the lowering of tones in sentences with alternating tones. In the current section 4.3.4, declination in sentences containing alternating tones, formerly known as downstep, has been modelled using the same quotients that have been established in section 4.3.2. The operation mode of the algorithm (Liberman & Pierrehumbert, 1984) has been upgraded in such a way that it is activated by the presence of the IP associated register tones, it also predicts L tones and that it works tone-sensitive. The goodness of the model has been demontrated.

The next chapter bears on the issue of preplanning i.e. how sentence length and prosodic structure contribute to the scaling of initial tones aka reference values.

5. Chapter
Anticipatory raising, length and prosodic structure

The phonetic implementation algorithm of tones, introduced in the previous chapter, works with a minimal phonetic look-ahead of one tone. Additionally, it has suggested that declination, which is generated by the algorithm, is triggered by post-lexical tones in the phonological representation. Thus, the algorithm is principally able to take higher level information into account e.g. register tones associated to the IP. Global effects on the realization of tones, such as sentence type, cancelation/reduction of downstep and register raising have been reported in a number of languages (Myers, 2004); see chapter 6 for further details. These effects have been explained using the notion of phonological look-ahead (Pierrehumbert, 2000:31), which is however, limited to "…upward search in the tree structure.". It has been a matter of debate what kind of information is available at the stage of phonetic implementation and which planning unit is employed at the phonetic level; see e.g. Rialland (2001). In this chapter, the details of initial F0 height, corresponding to pitch range/register choice (Liberman & Pierrehumbert, 1984) in Akan will be investigated. Specifically, the pitch register choice in relation to sentence length and in relation to the prosodic structure will be explored in section 5.2. Furthermore, anticipatory raising of initial tones in relation to the presence and number of (non-automatic) downsteps will be investigated in section 5.1. It will be shown that sentence length/complexity affects the pitch register choice, i.e. initial F0, of Akan sentences. Furthermore, it will be proposed that syntactic embedding is reflected prosodically in terms of anticipatory raising at the edge of embdedded IPs.

5.1 Anticipatory raising and downstep

In chapter 3, section 3.2, it has been shown that H tones are raised locally before L tones to make room in the tonal space (Chen, 2012). It has been a matter of debate whether H raising appears globally, in anticipation of the number of upcoming downsteps (Rialland, 2001). Stewart (1965) observes that in Akan, the initial H tone is raised in anticipation of a (non-automatic) downstep occurring later in the utterance. This is illustrated in (91).

(91) a. Á!dɛ́ń nà ɔ́-ń-hyɛ́ kàwá?
5 44 L 33 3 L 2
why PART PRO-OPT-wear ring
'Why must he wear a ring?'

b. Á!dɛ́ń nà ɔ́-ń-!hyɛ́ kàwá?
6 55 L 44 3 L 2
why PART PRO-OPT.NEG-wear ring
'Why must he not wear a ring?'

(Stewart 1965:5; gloss adjusted and L tones added)

Abstract tonal heights are given under the segmental string in numbers. A higher number corresponds to a higher pitch. L tones are marked as such. Crucially, the initial H tone is higher in (91)b. than in (91)a. because it contains one more downstep; see also Rialland & Somé (2000) for Dagara; Laniran & Clements (2003) for Yoruba; van Heuven (2004) for Dutch and Hyman (2007) for Amo and Luganda. Schachter rejects Stewart's claim in a comment by stating that "…it may be the case that the speaker will begin a very long sentence at a somewhat higher pitch level than he will a shorter one, it seems to me on the face of it unlikely that the speaker could anticipate the exact number of downsteps…" (Schachter, 1965:32). In chapter 3 section 3.2 I have shown that H raising is strictly local in Akan, in chapter 4 section 4.1 that automatic downstep is phonetically similar to non-automatic downstep and in section 4.3 that downstep is phonetically similar to declination, hence, it seems unlikely to me, that Akan speakers will adjust their pitch register solely on the basis of a floating L tone later in the tonal string. Thus, the following hypothesis will be tested.

Hypothesis: Initial H tones are not sensitive to the number/presence of upcoming downstep(s).

Material: To test the hypothesis, the height of an initial H tone in a sentence with automatic downstep (66)a., that corresponeds to (91)a., will be compared to the height of an initial H tone in a sentence containing non-automatic downstep (66)b., that correcponds to (91)b. The material used in Genzel & Kügler (2010) is illustrated in chapter 4 section 4.1.

Results: The mean F0 value, aggregated over speakers and repetitions, obtained for the initial H tone amounts to 225.43 Hz (85) for the sentence with downstep, and to 224.24 Hz (67) for the sentence with non-automatic downstep; see figure 32 for illustration. The figure 64 displays the F0 values of the initial H tones, aggregated over repetitions and split by downstep type, for each speaker. Speaker 4 shows a slightly higher value for the initial H tones of the sentence containing downstep. Speaker 2 shows the opposite effect. However, there is no systematic raising effect on the initial H tone of the sentence containing downstep for the rest of the speakers.

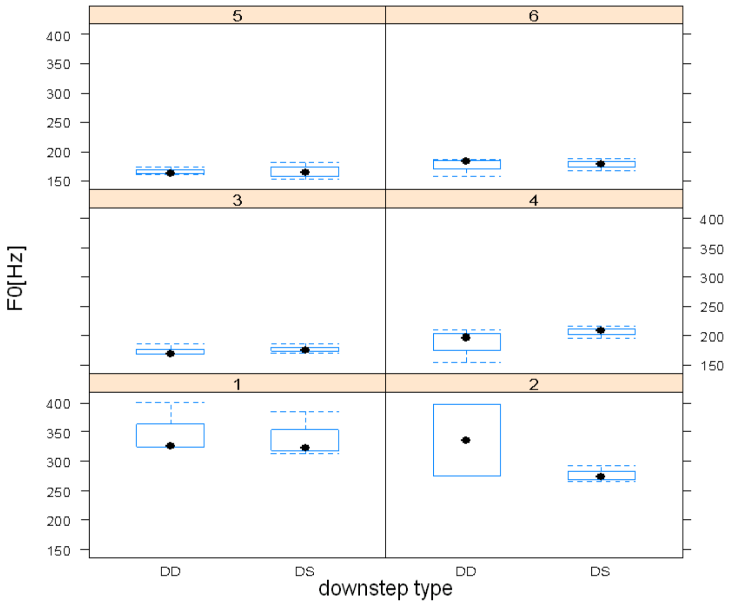


figure 64: F0 values of the initial H tone, aggregated over repetitions[42] and split by speakers and downstep type, DD refers to automatic downstep and DS refers to non-automatic downstep.

A linear mixed effects model was calculated on the height of the initial H tone. Downstep type (automatic/non-automatic) was treated as fixed factor. Repetitions and speakers were considered as random factors. There was no significant effect.

Summary: Akan speakers do not anticipate the initial F0 height according to the number of non-automatic downsteps; see figure 63 for illustration. This finding is in line with Schachter's (1965) observation and speaks against Stewart's (1965) claim. It thus can be concluded that Akan speakers do not anticipate the initial F0 value in reference to the number of tones in the tonal string. The following section will explore the dependence of anticipatory raising and length/complexity of pPs and IPs.

[42] For one repetition of speaker 2 F0 was not used for analysis because of falsetto voice.

5.2 Anticipatory raising and phrase length

Introduction: Already Christaller (1875:183) made the observation that "… in the *beginning* of a longer sentence or complex sentences,… tones, … are *higher throughout*….". The weakest position on preplanning one can take, is that it is a speaker's choice to adapt the pitch register according to the length of a sentence (Liberman & Pierrehumbert, 1984 'soft preplanning'; Prieto et al., 2006), i.e. the height of the first (H) tone depends on the length of a sentence and thus serves as an indicator of preplanning or anticipatory raising.

Studies testing the influence of the factor sentence length on the scaling of initial F0 values have shown mixed results in intonation (Liberman & Pierrehumbert, 1984; Prieto et al., 1996; van Heuven, 2004; Prieto et al., 2006; Thorson, 2007 among others) as well as in tone languages. Lindau (1986) investigated preplanning in Hausa and found no anticipatory raising effect for initial H tones. Snider (1998) reports that initial H tones in sentences with only H tones in Chumburung show an anticipatory raising effect, while initial L tones in sentences with only L tone do not.

Laniran & Clements (2003) show for Yoruba that initial H tones in HL tone sequences are raised in F0 with elongation of the utterance but admit that the effect does not hold for all speakers. The main strategy to economize pitch space, used by Yoruba speakers, is resetting of a later H tone in longer sequences involving downstep (Laniran & Clements, 2003:206). Connell (2003, 2004) studied preplanning in Mambila, and reports that in sentences containing only L and only H tones and HL sequences, the initial tones show no correlation with utterance length. The absence of a correlation between sentence length and initial F0 of H in Hausa and L in Chumburung was interpreted as a language-specific difference. It was accounted for by assuming phonologically specified F0 values (Lindau, 1986; Snider, 1998).

Alternatively, it may also be the case, as discussed earlier in the chapter 4 section 4.3.1 about declination, that the phonology of the language constraints the phonetic implementation in a certain way; see also Connell (2004). However, recent research has shown that speakers prefer to plan, if the speaking situation allows it (see e.g. Ferreira & Swets, 2002; Levelt, 2002; Wagner, Jescheniak & Schriefers, 2010) and the working memory capacity is sufficient (Petrone et al., 2011), which bears on the issue of inter-speaker variation reported in many experiments. Petrone et al., (2011) also investigated planning effects in German from a phonetic point of view. The relevant factor in determining initial F0 height, in German, is the length of the subject and not the prepositional length i.e. the planning scope seems to be rather local in the sense that the first pP is the domain of planning; see also Fuchs et al. (2013) for German and Scholz (2012) for Wenzhou Chinese. The figure 65 illustrates with data from

Wenzhou Chinese (Scholz, 2012:134), which is also an SVO language, that the height of the first H tone of the initial prosodic constituent (subject) is higher in relation to the number of prosodic words in the first pP and not in relation to the number of prosodic words of the second pP (object).

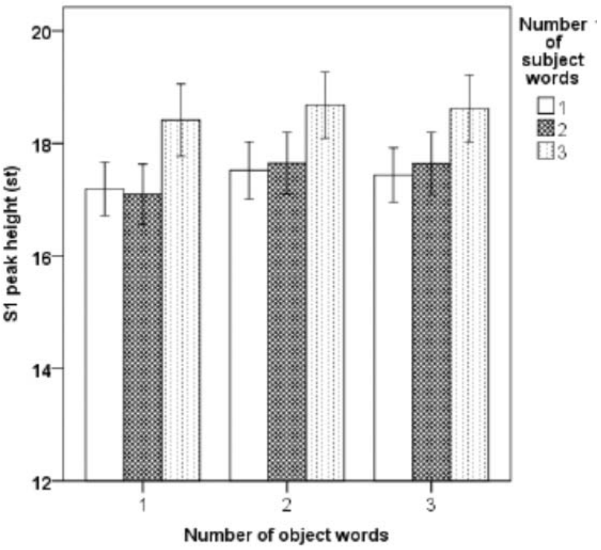


figure 65: F0 maxima (st) on the first subject peak in Wenzhou Chinese, broken down by constituent length, averaged across speakers (19). T-bars = ±2 SE; reproduced from Scholz (2012:134).

Apart from constituent length, syntactic complexity of a sentence has been found to influence F0 scaling (e.g. Christaller, 1875). Most studies investigating the issue are concerned with prosodic phrasing and its phonetic correlates; see chapter 1 section 1.4 for further details and also e.g. Ladd, (1986), (1996) for English; van den Berg, Gussenhoven & Rietveld, (1992) for Dutch; Truckenbrodt, (2002), (2007) and Truckenbrodt & Féry, (2003), Féry & Truckenbrodt, (2005) and Féry & Schubö, (2010) for German. Since the influence of syntactic complexity per se will not be the focus of attention here, it is sufficient to note that, assuming recursion on the prosodic level (e.g. Ladd, 1986, 1996; Selkirk, 2009; Féry, 2010 and Féry & Schubö, 2010), syntactically embedded structures (e.g. complementizer clauses) are also prosodically embedded into their matrix clause i.e. declination is continuing; see also Scholz (2012) for details in Wenzhou Chinese.

Given Christaller's (1875) and Schachter's (1965) observations, together with the fact that the experimental situation allows the speakers to plan, anticipatory raising is expected to take place. Based on the assumption that the information of register tones associated to the IP node is available for the phonetic implementation, see chapter 4 sections 4.3.2 and 4.3.4, it is further expected that preplanning in Akan is not limited to the first pP (subject) but rather takes the size of the whole IP into account. Moreover, chapter 4 section 4.3.3 has motivated the assumption that initial tones receive reference values. On the basis of these reference values, the pitch implementation algorithm calculates F0 values for each identity separately. Hence, it is expected that initial L and H tones in alternating (LH/HL) sequences anticipate the length of the sentence individually. The following hypotheses will be tested.

Hypotheses:

i. Initial tones are scaled higher in longer sentences.
ii. The domain of planning is the IP.

Material: The first hypothesis (i.) will be tested by comparing the initial F0 value in a short sentence to the initial F0 value in a longer sentence. The material is represented in table 19. I will refer to it as data set 1, from now on. The data was grouped for the result presentation. The length categories, see table 19, short/medium were merged to short and long/longest to long. The short sentences for each tonal environment are illustrated in (92). The short sentence containing only H tones is displayed in (92)a. and the short sentence containing only L tones in (92)b. The short sentence with alternating HL tones is displayed (92)c. and the short sentence with alternating LH tones in (92)d. The longest sentences for each tonal environment are illustrated in (93). The longest sentence containing only H tones is displayed (93)a. and the longest sentence containing only L tones in (93)b. The longest sentence with alternating HL tones is displayed (93)c. and the longest sentence with alternating LH tones in (93)d. The complete list of materials is introduced in chapter 2, section 2.2, blocks A and B.

Length	**H**	**Length**	**L**	**Length**	**HL**	**Length**	**LH**
short (5 σ's)	(42)a.	short (7 σ's)	(43)a.	short (5 σ's)	(45)a.	short (6 σ's)	(44)a.
medium (7 σ's)	(42)b.	medium (8 σ's)	(43)b.	medium (7 σ's)	(45)b.		
long (10 σ's)	(42)c.	long (10 σ's)	(43)c.	long (9 σ's)	(45)c.	long (10 σ's)	(44)b.
longest (12 σ's)	(42)d.	longest (12 σ's)	(43)d.	longest (11 σ's)	(45)d.	longest (12 σ's)	(44)c.

table 19: Material for testing anticipatory raising, data set 1.

(92) a. Kúkúó-bá nó. (42)a.
pot-DIM DEF
'The small pot.'

b. Yàẁ fì Àkyèm̀fò. (43)a.
proper name originate.HAB proper name
'Yaw comes from Akyemfo.'

c. Ánàné bìsá. (45)a.
proper name ask.HAB
'Anane asks.'

d. Pàpá Kòfí kàsá. (44)a.
father proper name talk.HAB
'Father Kofi talks.'

(93) a. Kúkúó-bá pápá páá nó bɔ́ dáá. (42)d.
pot-DIM good very DEF break.HAB daily
'The very good small pot breaks everyday.'

b. Wɔ̀fà Àdò Àsàrè fì Àkyèmfò. (43)d.
uncle proper name proper name originate.HAB proper name
'Uncle Ado Asare comes from Akyemfo.'

c. Ánàné bìsá sìká bɔ̀né ɛ̀nɛ́. (45)d.
proper name ask.HAB money bad today
'Anane asks for bad money today.'

d. Pàpá Kòfí kàsá kyèrɛ́ nè bá bíò. (44)c.
father proper name talk.HAB point_to.HAB PRO child again
'Father Kofi talks to his child again.'

To explore the second hypotheses (ii), the initial F0 in a sentence with a complex first pP (subject), as in sentences with only H/only L tones, see (92)a., (92)b., (93)a. and (93)b., will be compared to the the initial F0 in a sentence with a complex second pP (VP), as in sentence with alternating LH/HL; see (92)c., (92)d., (93)c. and (93)d. Additionally, complementizer clauses containing only H or only L tones varying in length will be analyzed. The material is represented in table 20. I will refer to it as data set 2. An example of a complex sentence that contains a short H toned complementizer clause is displayed in (94). The complete list of materials is introduced in chapter 2, section 2.2, block C.

Length	**H**	**Length**	**L**
short (5 σ's)	(47)a.	short (7 σ's)	(48)a.
medium (7 σ's)	(47)b.	medium (8 σ's)	(48)b.
long (11 σ's)	(47)c.	long (10 σ's)	(48)c.

table 20: Material for testing anticipatory raising, data set 2.

(94) Nàná kà-à ɛ̀-nórà sɛ́ kúkúó-bá bɔ́. (47)a.
proper name say-PST NOM-yesterday COMP pot-DIM break.HAB
'Nana said yesterday that the small pot breaks.'

The initial F0 of the matrix clause and the complementizer clause will be compared to further explore the second hypothesis (ii.). The matrix clause (InfP1) contains the verb *to say* which is followed by the complementizer *sɛ*. The tonal specification of *sɛ* is unclear. According to Amfo (2007:144), the complementizer is associated with an H tone, whereas Dolphyne (1988:65) observes, that it has a falling pitch (HL), which is part of the lexical structure of *sɛ*; see discussion of this section for further details. The complementizer introduces a clausal complement (InflP2). The assumed syntactic structure is displayed in (95), reproduced from Boadi (2005:51).

(95) $[_{\text{InfP1}}$... kà-à $[_{\text{CP}}$sɛ́ $[_{\text{InfP2}}$....]]]
say-PST COMP
'… said that ….'

(Boadi; 2005:51)

Regarding the prosodic structure, I will follow Selkirk's (2009:49) proposal that the sentence in (95) exhibits the prosodic structure in (96), consisting of two recursive IPs; see also chapter 1 section 1.4.2. The complements of every Comp^0 and every Force^0 are mapped onto their own IP. Selkirk assumes, following Rizzi (1997), that the Force phrase carries the illocutionary force which constitutes the speech act of the sentence. Evidence in favor of this prosodic structure will be provided in the discussion of this chapter.

(96) $_{\text{ForceP}}$[…[Force^0[……..[Comp^0[……….]]]]]
ι(…kàà…sɛ́ ι(……….)ι)ι

(Based on Selkirk, 2009:49)

Again, speaker 2 was excluded from the analysis since he shows falsetto voice. The data of speaker 5 was discarded from parts of the analysis because initial L tone values could not be obtained due to creakiness in the short sentences with alternating LH tones.

Results data set 1: The figure 66 gives an overview of anticipatory raising in sentences with only H tones, left hand side, and only L tones, right hand side, uttered by a female speaker. Short sentences are represented by a solid line and long sentences by a doshed line. For both tones, initial values are realized higher in anticipation of the length of the sentence. Anticipatory raising is more pronounced for the H tones. Despite of anticipatory raising, tonal contrasts are well preserved. A raised initial L tone does not reach the height of an initial H tone.

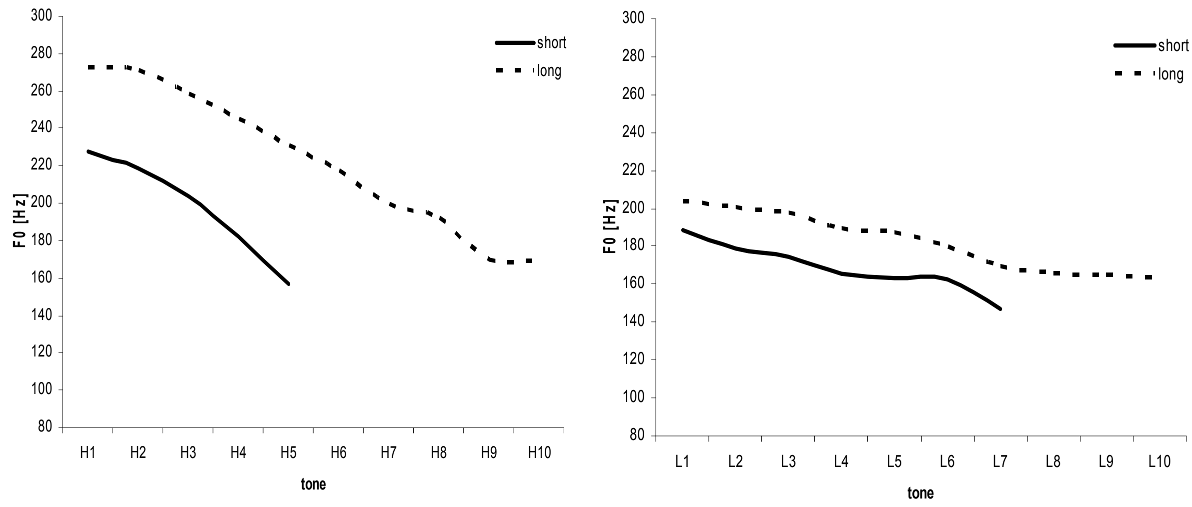


figure 66: Left plot: Mean F0, aggregated over repetitions, for short (solid line), *Kukuoba no*. 'The small pot.' and long sentences, *Kukuoba papa no bɔ daa.* 'The good small pot breaks everyday.' (dashed line) consisting of only H tones; Right plot: Mean F0, aggregated over repetitions, for short, *Yaw fi Akyemfo*. 'Yaw comes from Akyemfo.' and long sentences, *Wɔfa Asare fi Akyemfo*. 'Uncle Asare comes from Akyemfo.' consisting of only L tones; female speaker.

The figure 67 gives an overview of anticipatory raising in sentences with alternating HL tones (left hand side) and alternating LH tones (right hand side), uttered by a female speaker. Short sentences are represented by a solid line and long sentences by a dotted line. In both configurations initial values are realized higher in anticipation of the length of the sentence.

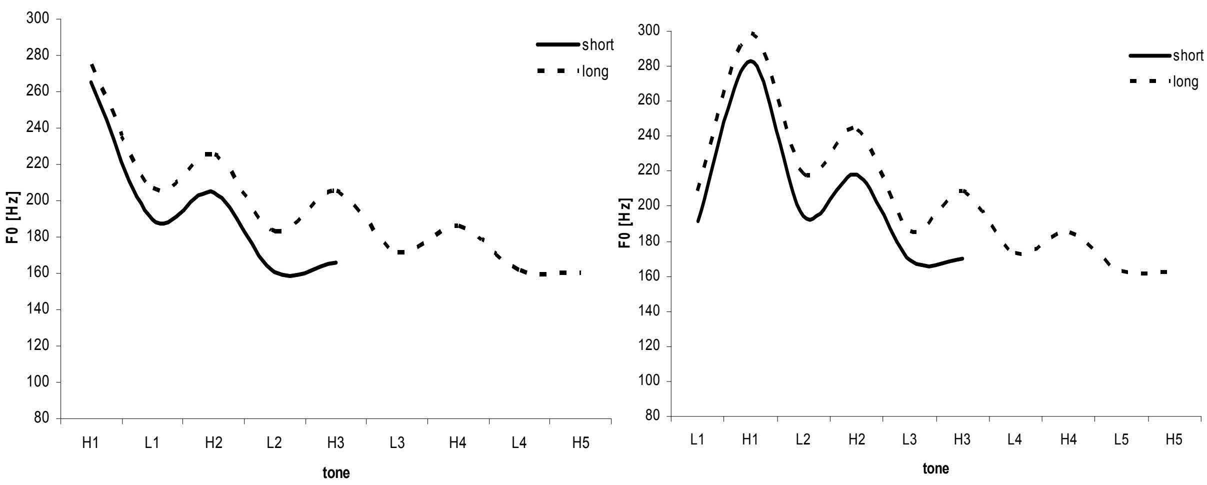


figure 67: Left plot: Mean F0, aggregated over repetitions, for short (solid line), *Anane bisa*. 'Anane asks.' and long sentences (dashed line), *Anane bisa sika ɛnɛ*. 'Anane asks for money today.', consisting of alternating HL tones; Right plot: mean F0, aggregated over repetitions, for short, *Papa Kofi kasa*. 'Father Kofi talks.' and long sentences, *Papa Kofi kasa kyerɛ ne ba*. 'Father Kofi talks to his child.', consisting of alternating LH tones; female speaker.

Turning to the results for the grouped whole sample, the mean F0 values for the initial H tones in sentences with only H tones, aggregated over speakers and repetitions, amount to 168.02 Hz (40) for the sentences of the short and to 179.27 Hz (42) for the sentences of the long group. The figure 68 displays the mean F0 values of the initial H tones, aggregated over repetitions and split by speakers and length group. All speakers show a higher initial F0 for the long renditions compared to the short ones.

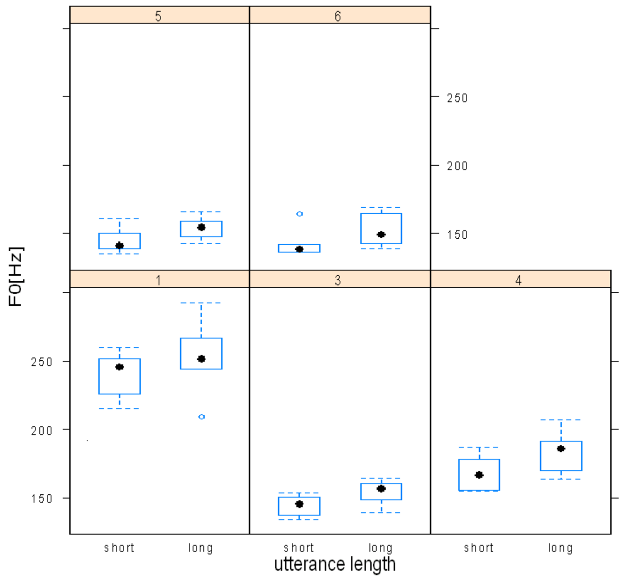


figure 68: Mean F0 values of the initial H tone in sentences with only H tones, aggregated over repetitions and split by speakers and length group.

Turning to the F0 of the initial L tones in sentences with only L tones, the mean values**,** aggregated over speakers and repetitions, amount to 143.01 Hz (30) for the short and to 144.22 Hz (35) for the long group. The figure 69 displays the F0 values of the initial L tones, aggregated over repetitions and split by speakers and length group. The anticipatory effect is only observable in the data of speakers 1 and 2. All other speakers do not show a higher initial F0 value for the long group compared to the short one.

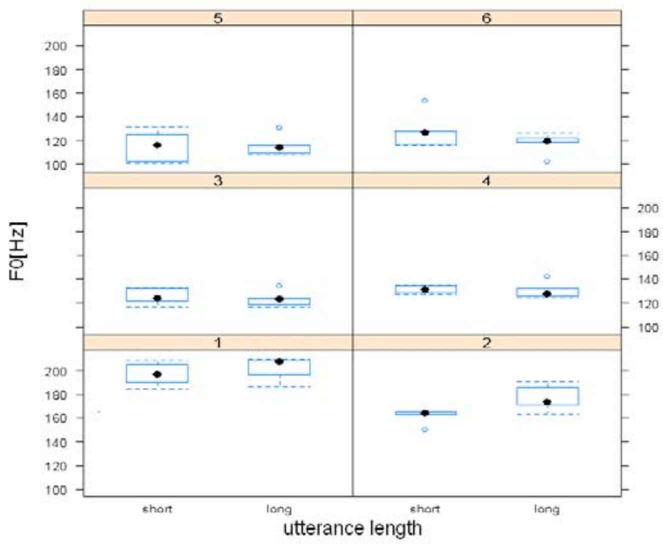


figure 69: Mean F0 values of the initial L tone in sentences with only L tones, aggregated over repetitions and split by speakers and length group.

Turning to the initial F0 in sentences with alternating HL tones, the mean values for the initial H tone, aggregated over speakers and repetitions, amount to 187.33 Hz (54) for the short, and to 195.23 Hz (57) for the long group. The mean F0 values for the first L tone, aggregated over speakers and repetitions are at 141.22 Hz (31) for the short group and at 150.83 Hz (32) for the long group. The figure 70 displays the F0 values of the initial H (left hand side) and L tones (right hand side), aggregated over repetitions split by speakers and length group. Apart from speaker 5, all speakers show an anticipatory raising effect for the H tone as well as for the L tone.

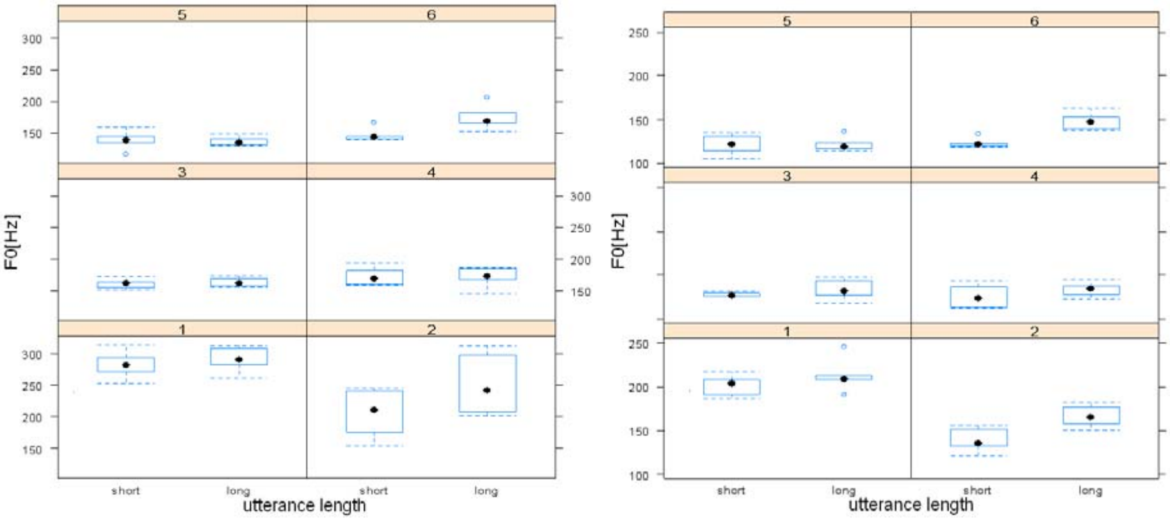


figure 70: Left boxplot: F0 values of the initial H tone in sentences with alternating HL tones, aggregated over repetitions and split by speakers and length group; Right boxplot: F0 values of the initial L tone in sentences with alternating HL tones, aggregated over repetitions and split by speakers and length group.

Turning to the initial F0 in sentences with alternating LH tones, the mean values for the initial L tone, aggregated over speakers and repetitions, amount to 135.52 Hz (31) for the short group, and to 140.69 Hz (34) for the long group. The mean values for the first H tone, aggregated over speakers and repetitions, are at 181.98 Hz (57) for the short group and at 199.94 Hz (58) for the long group. The figure 71 displays the F0 values of the initial L (left hand side) and H tones (right hand side), aggregated over repetitions split by speakers and length group. All speakers show an anticipatory raising effect for the L tone as well as for the H tone.

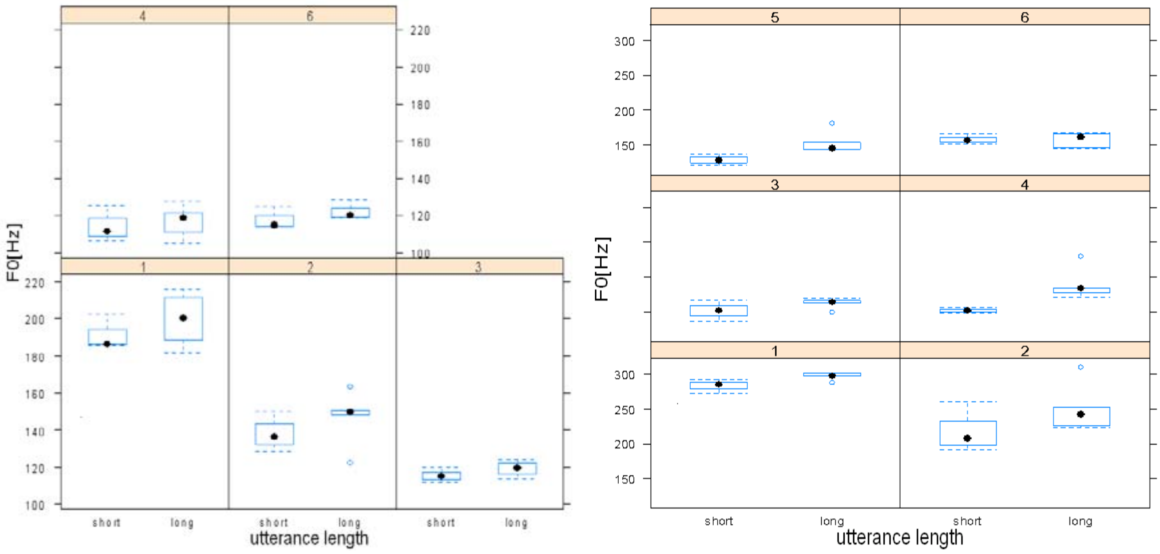


figure 71: Left boxplot: F0 of the initial L tone in sentences with alternating LH tones, aggregated over repetitions and split by speakers and length group; Right boxplot: F0 of the initial H tone in sentences with alternating LH tones, aggregated over repetitions and split by speakers and length group.

A linear mixed effects model was calculated on the initial F0. Tonal configuration (only L, H/ alternating HL, LH), tone (L/H) and length (short/long) were treated as fixed factors. Repetitions and speakers were considered as random factors. Recall that the long sentences with only H and only L tones contain a longer first pP (subject) whereas the long sentences with alternating tones exhibit a longer second pP (VP). Tonal configuration showed a significant main effect (t = -2.591, pMCMC < 0.05, 379 observations); tones in alternating environments are overall higher than in sentences with the same tonal identity. Furthermore, tone had a significant main effect (t = -20.581, pMCMC < 0.0001); as expected, L tones are generally lower than H tones and vice versa. There was a significant interaction between the two factors (t = 3.644, pMCMC < 0.0005) pointing to the fact that the effect of tonal configuration is limited to the H tones i.e. H tone raising is at work; see also chapter 3 section 3.2. Moreover, the factor length affected the initial tones significantly (t = -4.627, pMCMC < 0.0001); initial tones in both configurations are raised in longer sentences. There was no significant interaction between tonal configuration & length, tone & length and tonal configuration & tone & length.

Summary: The data at hand provided evidence for Christaller's (1875) and Schachter's (1965) claim that speakers begin longer sentences with a higher initial F0, which corresponds to a pitch range/register increase (Liberman & Pierrehumbert, 1984:191). Both, L and H tones can be raised in anticipation of the length of the sentence; see figure 66 for illustration. L tones are raised to a lesser degree than H tones. Furthermore, the data has shown that in alternating sequences both initial tones, L and H, are raised irrespective of which tone is the first; see figure 67. Tonal contrasts are preserved. Complexity/length increase in the data containing only L or only H tones was created by insertion of one or two pws to the first pP containing the subject. This is illustrated in (97)a. with the short sentence *Yaw fi Akyemfo.* 'Yaw comes from Akyemfo.' and the longest sentence *Wɔfa Ado Asare fi Akyemfo.* 'Uncle Ado Asare comes from Akyemfo.', containing only L tones. Anticipatory raising is indicated by the upward arrow (↑). Complexity/length increase in the data containing alternating LH/HL tones was created by insertation of one or two pws/pPs to the pP wrapping the VP. This is illustrated in (97)b. with the short sentence *Papa Kofi kasa.* 'Papa Kofi talks.' and the long sentence *Papa Kofi kasa kyerɛ ne ba.* 'Father Kofi talks to his child.'. Unlike in Wenzhou Chinese (Scholz, 2012) anticipatory raising in Akan is not only sensitive to the length/complexity of the first maximal pP (subject), but also to length/complexity of the second maximal pP (VP).

(97) a. ((Yaw)φ (fi(Akyemfo) φ)φ)ι

((↑Wɔfa Ado Asare)φ (fi(Akyemfo)φ) φ)ι

b. ((Papa Kofi)φ (kasa)φ)ι

((↑Papa Kofi)φ (kasa kyerɛ (ne ba) φ)φ)ι

Results data set 2: The figure 72 and figure 73 give a first impression of anticipatory raising in complementizer clauses differing in length. The complementizer clauses containing only L tones are displayed in figure 72 and the complementizer clauses containing only H tones in figure 73. Short sentences are represented by the solid line, medium sentences by the dashed line and long sentences by the gazed line in both figures. Generally, the initial L tone of the matrix clause (L1) does not show any raising effect in anticipation of the length of the embedded complementizer clause. However, the first H tone (H1), which is the second tone in the matrix clause, is realized slightly higher if the embedded complementizer clause is long. The initial L tone (L1) of the complementizer clause, in the medium and long sentences is realized higher than in the short ones, see figure 72.

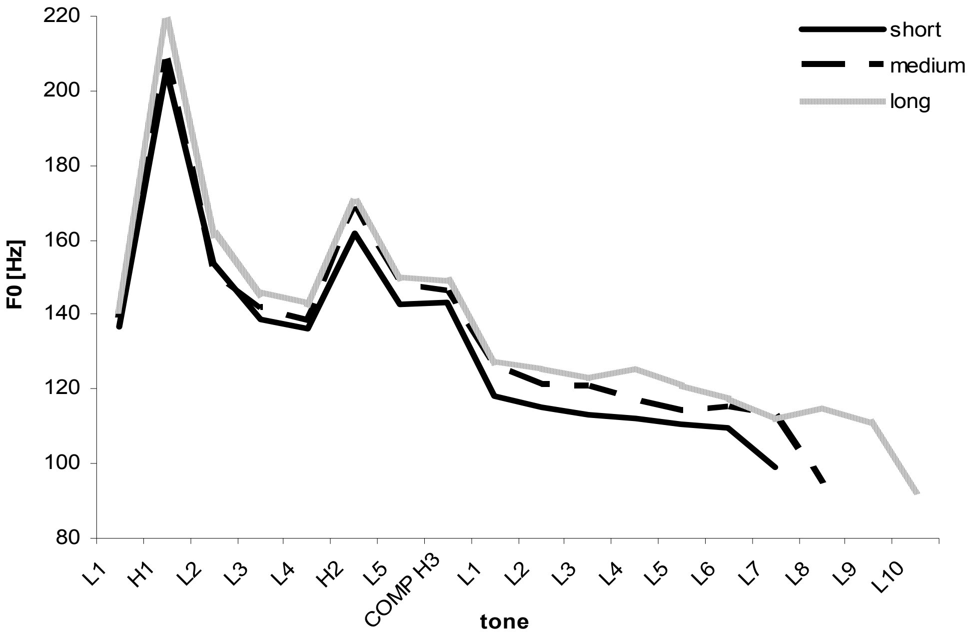


figure 72: Mean F0 values, aggregated over speakers and repetitions, of complex sentences with complementizer clauses containing only L tones; short complementizer clause (solid line), *Nana kaa ɛnora sɛ Yaw fi Akyemfo.* 'Nana said yesterday that Yaw comes from Akyemfo.', medium complementizer clause (dashed line), *Nana kaa ɛnora sɛ Asare fi Akyemfo.* 'Nana said yesterday that Asare comes from Akyemfo.' and long complementizer clause (gazed line), *Nana kaa ɛnora sɛ wɔfa Asare fi Akyemfo.* 'Nana said yesterday that uncle Asare comes from Akyemfo.',; n=6.

The anticipatory effect is even more pronounced for the initial H tones (H1) of the complentizer clause, see figure 73.

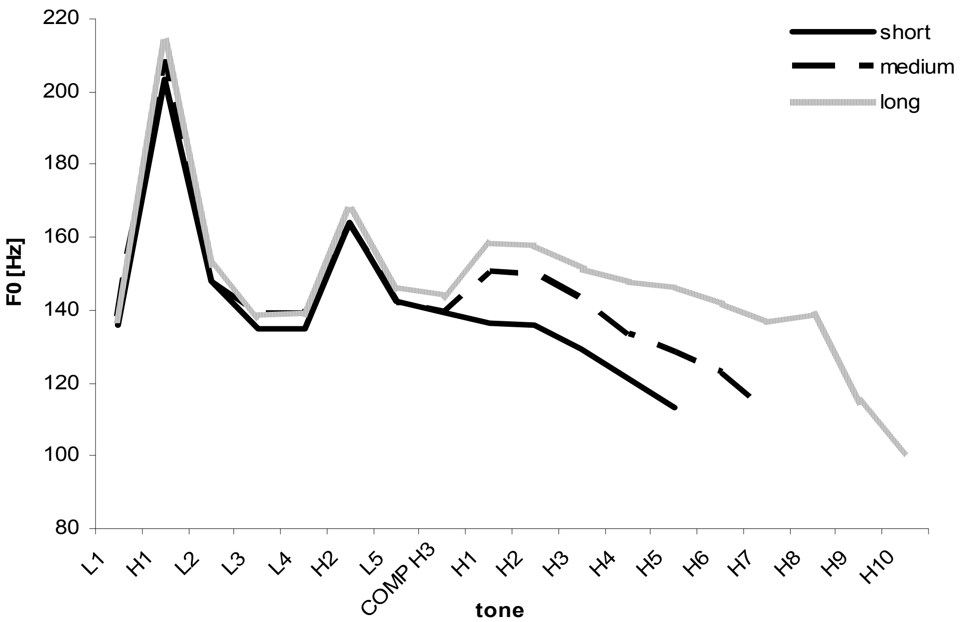


figure 73: Mean F0 values, aggregated over speakers and repetitions, of complex sentences complementizer clauses containing only H tones; short complementizer clause (solid line), *Nana kaa ɛnora sɛ kukuoba bɔ.* 'Nana said yesterday that a small pot breaks.', medium complementizer clause (dashed line), *Nana kaa ɛnora sɛ kukuoba papa bɔ.* 'Nana said yesterday that a good small pot breaks.' and long complementizer clause (gazed line), *Nana kaa ɛnora sɛ kukuoba papa no bɔ daa.* 'Nana said yesterday that the good small pot breaks everyday.'; n=6.

Turning to the numerical results of anticipatory raising of initial tones of the matrix clause, the mean F0 values for the initial L tone, aggregated over speakers and repetitions, amount to 136.33 Hz (31) for sentences containing the short complementizer clause, to 138.86 Hz (34) for sentences containing the medium complementizer clause, and to 139.10 Hz (31) for sentences containing the long complementizer clause. The means for the initial H tone (second tone) of the matrix clause, aggregated over speakers and repetitions, are at 204.45 Hz (56) for sentences containing the short complementizer clause, at 208.76 Hz (54) for sentences containing the medium complementizer clause and at 216.05 Hz (62) for sentences containing the long complementizer clause.

The figure 74 displays the F0 values for the initial L (left hand side) and H tones (right hand side) of the matrix clause, aggregated over repetitions and split by speakers and length of the complementizer clause. A gradual increase, for the initial L tones of the matrix clause, with increasing length of the complementizer clause, can be observed for speaker 4 only. Speaker 3 shows a slightly raised value for sentences containing the long complementizer clause. The two other speakers do not show a systematic anticipatory effect.

The initial H tone of the matrix clause, speakers 1, 2, 3 and 6 exhibit a higher F0 for the sentences containing the long complementizer clause. All other speakers do not exhibit a systematic anticipatory effect.

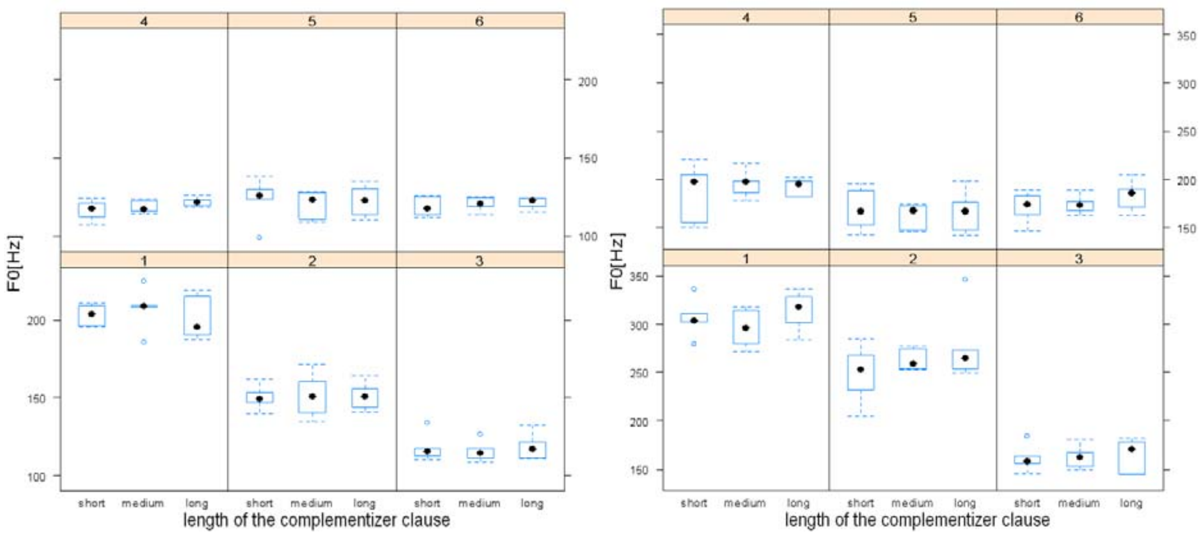


figure 74: Left boxplot: F0 of the initial L tone of the matrix clause, aggregated over repetitions and split by speakers and lengths of the complementizer clause; Right boxplot: F0 of the initial H tone of the matrix clause, aggregated over repetitions and split by speakers and lengths of the complementizer clause.

A linear mixed effects model was calculated on the height of the initial tone of the matrix clause. Tone (L/H) and length (short as reference category/medium/long) were treated as fixed factors. Speakers and repetitions were included as random factors. The factor tone had a significant effect (t = -26.769, pMCMC < 0.0001, 216 observations), which is redundant. All other comparisons and interactions did not yield significant results.

Turning to the initial tones of the complementizer clause, the mean values for the initial L tone, aggregated over speakers and repetitions, amount to 118.01 Hz (23) for the short complementizer clause, to 126.56 Hz (31) for the medium complementizer clause and to 126.97 Hz (25) for the long complementizer clause. The means for the initial H tone, aggregated over speakers and repetitions, are at 139.03 Hz (31) for the short complementizer clause, at 150.72 Hz (40) for the medium complementizer clause and at 158.16 Hz (39) for the long complementizer clause. The figure 75 displays the F0 values of the initial L (left hand side) and H tones (right hand side) of the complementizer clause, aggregated over repetitions split by speakers and lengths. Only speaker 3 shows gradual increase of the initial L tone with increasing length of the complementizer clause. Speakers 1, 2 and 6 show a higher value for the medium and the long complementizer clause in comparison to the short one and speakers 4 and 5 only for the long complementizer clause in comparison to the medium and short one. For the initial H tone, a gradual increase with increasing length of the complementizer clause can be observed for speaker 2, 3, 4 and 6. Speakers 1 and 5 show an increase in F0 for the medium length but not for the long instances.

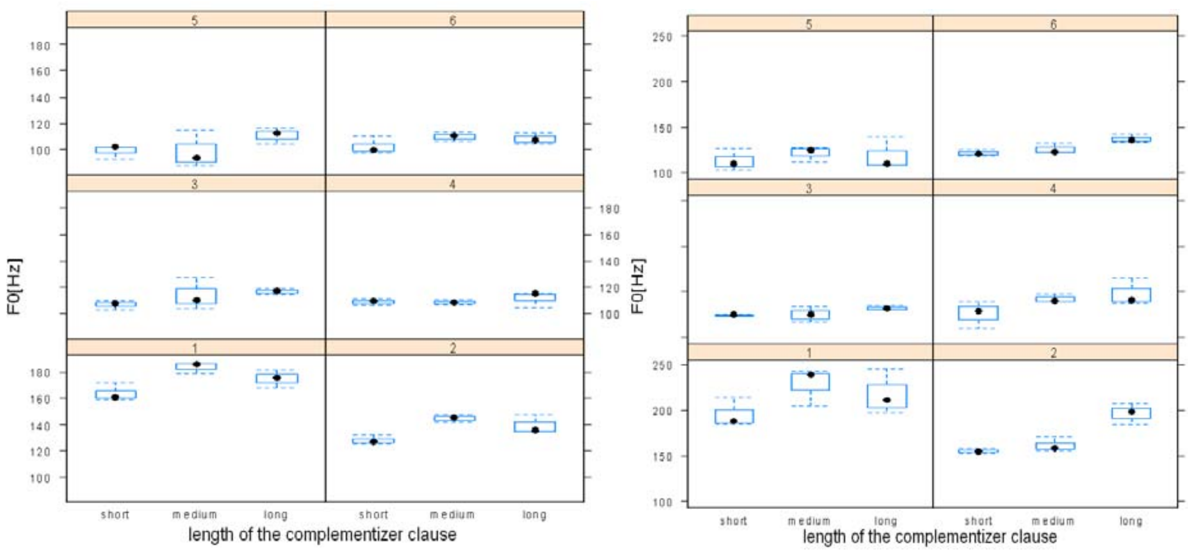


figure 75: Left boxplot: F0 of the initial L tone of the complementizer clause containing only L tones, aggregated over repetitions and split by speakers and lengths; Right boxplot: F0 of the initial H tone of the complementizer clause containing only H tones, aggregated over repetitions and split by speakers and lengths.

A linear mixed effects model was calculated on the height of the initial tone of the complementizer clause. Tone (L/H) and length (short as reference category/medium/long) were treated as fixed factors. Speakers and repetitions were considered as random factors. The factor tone had a significant effect (*t* = -11.569, *p*MCMC < 0.0001, 108 observations), which is redundant. The length contrast between short and long was significant (*t* = 3.849, *p*MCMC < 0.005); the initial F0 is higher in the long instances than in the short ones, for both tones. There was no significant interaction between tone and length. The length comparison between short and medium did not yield a significant result.

Summary: The analysis of dataset 2 has provided further insights into the domain of preplanning in Akan. The length/complexity of the matrix clause has been kept constant while the length/complexity of the complementizer clause was varied by insertion of one or two pws to the subject of the complementizer clause. The complementizer clause contained either only L tones or H tones. The intial tones of the matrix clause are insensitive to the length/complexity of the complementizer clause, no significant anticipatory raising effect was observed. The preplanning effect was only found on the initial tone of the complementizer clause; see figure 72 and figure 73. Anticipatory raising in complex sentences in Akan is illustrated in (98). (98) shows the complex sentence containing the short complementizer clause *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' or the long complementizer clause, *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' with only L tones. The upward arrow indicated raising.

(98) ((Nana)φ (kaa(ɛnora)φ (sɛ)φ)φ ((Yaw)φ (fi(Akyemfo) φ)φ)ι)ι
((Nana)φ (kaa(ɛnora)φ (sɛ)φ)φ ((↑wɔfa Asare)φ (fi(Akyemfo) φ)φ)ι)ι

Discussion: In section 5.1 of this chapter, it was investigated whether the initial F0 height is sensitive to the number of upcoming non-automatic downsteps, following Stewart's (1965) observation. The analysis of the data has shown that the intial H tone in sentences containing non-automatic downstep, i.e. a floating L tone in the tonal string, exhibits the same height as the intial H tone in sentences without a floating L tone. This observation is in line with Schachter (1965).

The investigation of data set 1, presented in section 5.2, has revealed that anticipatory raising does not depend on the size of the first constituent (subject) of a sentence. Unlike in German (Petrone et al., 2011) and Wenzhou Chinese (Scholz, 2012), Akan speakers take the length of the whole IP into account to pre-plan the height of the initial tone(s). The analysis of preplanning in more complex structures (data set 2) has provided further details on the domain of planning. It has been shown that the height of the initial tone(s) of the matrix clause, whose length was kept constant, are not subject to anticipatory raising in adaption to the length of the complementizer clause. However, the information that an embedded IP is coming up is available. The height of the first H tone of the matrix clause, which is eight syllables long (including COMP), is systematically scaled higher, 209.8 Hz (57)[43], than the first H tone, 189.23 Hz (56), in the simple sentence with alternating HL tones consisting of seven syllables ($t = 4.079$, pMCMC < 0.001, 125 observations). Thus, preplanning in Akan is global in the sense that general information of the upcoming higher level prosodic structure (IP) is available and local in the sense that specific information of the length of the IP affects the initial F0 height when the relevant IP starts. The experimental results have revealed that the initial tone(s) of the complementizer clause is subject to anticipatory raising.

In sentences with alternating tones both initial tones, L and H, are raised in anticipation of the length of the IP, irrespective of the identity of the first tone. The majority of studies investigating pitch range/register effects in relation to sentence length (e.g. Liberman & Pierrehumbert, 1984; Scholz, 2012) equate preplanning with the height initial H tones (peaks) in a sentence. The standard assumption based on English is that "…increasing the pitch range increases the F0 value of an initial peak…." (Liberman & Pierrehumbert, 1984:191); see also Pierrehumbert & Beckman (1988:179) for Japanese. This entails the idea that the scaling of L tones depends on the pitch range choice which is reflected in the height of initial H tones. In Akan, however, as already argued in chapter 4 section 4.3.3, L tones are independent phonological entities. The experimental results presented in this chapter have shown that pitch range/register is chosen for both tones independently, which may be interpreted as evidence that speakers "know" that the pitch implementation algorithm calculates F0 values for each identity separately; see chapter 4 section 4.3.3 for details.

Furthermore, the investigation of the scaling of initial H tones in alternating tone sequences (data set 1) has shown that H raising, which is most pronounced in the first H tone immediately preceding an L tone, see chapter 3 section 3.2, and anticipatory raising are cumulative; see Wang & Xu (2011) for cumulative effects on initial tones in Mandarin Chinese. H tones are raised by about 10 Hz in anticipation of the length of the sentence, which is comparable to the amount of H raising established in chapter 3 section 3.2. Unlike tonal dissimilation (H raising), anticipatory raising does not serve to facilitate perception of tones of the opposite identity but to make speech production more convenient for the speaker. It prevents or postpones reset, which was, unlike in Yoruba (Laniran & Clements, 2003), not present in the data presented here. The use of anticipatory raising ensures that the declination pattern can be maintained in longer and complex sentences to signal coherence; see also

[43] The mean value is aggregated over speaker, repetition and length of the complementizer clause.

chapter 4 section 4.3.1. Thus IP nodes and associated tonal specification (h, l) are available for the phonetics and are taken into consideration when reference values are planned before the pitch implementation algorithm starts.

The final part of the discussion will be concerned with evidence in support of the prosodic structure of the complex sentences introduced in (96) and repeated in (99), based on Selkirk (2009).

(99) $_{\text{ForceP}}$[…[Force^0[……..[Comp^0[………..]]]]]
ι(…kàà…sέ ι(………..)ι)ι
(Based on Selkirk, 2009:49)

First of all, it is interesting to note that syntactic embedding was also reflected prosodically. Declination was present over the whole IP. Evidence for the recursive IP structure, in which the complementizer is phrased together with the matrix clause as in (99), comes from pause distribution. If pauses occurred, they were located after the complementizer *sɛ*. Altogether, eight pauses occurred before the complementizer clause[44] containing only L tones. One appeared in the data of speakers 1 (54 ms) & 6 (317 ms), and six, in the data of speaker 4 (70 ms).

Furthermore, Selkirk's approach predicts no IP boundary at the right edge of the matrix clause, see (99). Supporting evidence comes from co-articulatory segmental effects observable on *sɛ*. Under the heading of extrinsic vowel duration, it has been noticed in the literature (e.g. Delattre, 1962; Keating, 1985) that vowels are shorter before voiceless obstruents. In the present data, *sɛ* exhibits a mean duration of 102 ms (45) when it is followed by a voiceless plosive, and a mean duration of 244 ms (67) when it is followed by a sonorant (t = 16.33, pMCMC < 0.0001, 108 observations). Moreover, Selkirk's approach predicts that the complementizer clause constitutes its own IP. The data on anticipatory raising, once more illustrated in figure 76, has provided evidence supporting this view.

[44] Since the complementizer clauses consisting of only H tones always started with a voiceless plosive and articulation of voiceless plosives involves a closure of the airstream passage, which is indistinguishable from a pause, pauses were only investigated for the complementizer clauses containing only L tones.

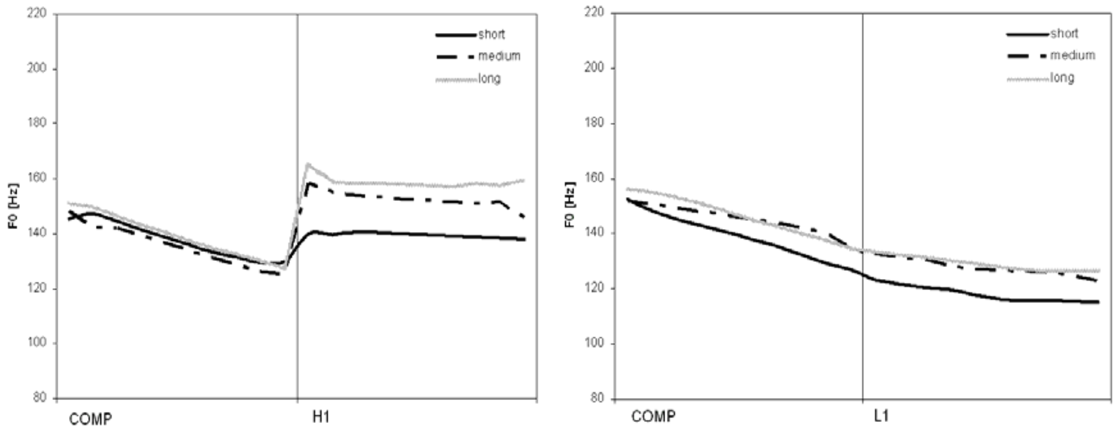


figure 76: Left plot: mean F0 (10 points of measurement per vowel), aggregated over speakers and repetitions, for the complementizer (COMP) and first H tone (H1) of the short (solid line), medium (dashed line) and long (gazed line) complementizer clauses containing only H tones; Right plot: mean F0, aggregated over speakers and repetitions, for the complementizer (COMP) and first L tone (L1) of the short (solid line), medium (dashed line) and long (gazed line) complementizer clauses containing only L tones; n=6.

Turning to the tonal behavior of the complementizer, figure 76 shows that *sɛ* exhibits a falling pitch movement, which is in line with Dolphyne's (1988) observation. The falling pitch movement on *sɛ* seems to be triggered by an additional L tone which is according to Dolphyne part of the lexical entry. Impressionistically, the vowel of the complementizer sounds prominent, which speaks in favor of assuming an additional tonal target as suggested by Dolphyne (1988:65); see chapter 6 for further evidence for increased articulatory effort on the final syllable of Yes – No questions. Thus, the final fall does not resemble final lowering at an IP boundary which is usually accompanied by a decrease in intensity (Vaissière, 1983:58); see following chapter 6 for further information.

In this section two different data sets have been analyzed to explore on which basis Akan speakers pre-plan the initial F0 of a sentence. The investigation of data set 1, consisting of simple sentences with only H or L tones and simple sentences with alternating tones differing in complexity/length has revealed that Akan speakers anticipate the length of the IP and not only the length of the first pP (subject) in simple (SVO) structures, as in German (Petrone et al., 2011; Fuchs et al., 2013) or Wenzhou Chinese (Scholz, 2012). Furthermore, it has been shown that anticipatory raising is present on initial L tones and on initial H tones, irrespective of which one is the first. This can be interpreted as evidence that pitch range is chosen for both tones (L and H) individually, which supports the tone-sensitivity idea of the phonetic implementation algorithm proposed in the previous chapter.

The analysis of data set 2, consisting of complex sentences (SVAdvCompSVO) with a matrix clause of constant length followed by a complementizer clause exhibiting only L or H tones differing in length, provided further insights into preplanning in Akan. It has been shown that the information that an embedded complementizer clause is coming up is available i.e.

information on the level of the IP, since the height of the first H tone of the matrix clause was systematically scaled higher in the complex structures compared to the simple SVO sentences. However, the height of the first H tone of the matrix clause was not raised in anticipation of the length of the complementizer clause, which was analysed as recursive embedded IP following Selkirk, 2009. The preplanning effect in the complex sentences was located only on the initial tone of the complementizer clause. This fact was interpreted as evidence that specific information of the length of an embedded IP is not available at the beginning of the sentences. Hence, preplanning in Akan is global and local.

Preplanning (anticipatory raising) is an important process at the level of pitch implementation. In Akan, it serves to ensure that declination can be maintained throughout the IP. It prevents pitch resetting; see also chapter 4 section 4.3.4.

6. Chapter
The intonation of Yes – No questions

Questions that lack any syntactic and/or morphological marking are said to be the best candidate for studying intonation (Haan, van Heuven, Pacilly & van Bezooijen, 1997). Yes – No questions, also referred to as polar questions, can be answered with Yes or No and therefore serve to request "…information about the truth of a certain proposition." (Nordhoff, 2009:697). They can be generated without wh-element or particle and are thus string identical with statements, as illustrated in (100) with an example from Akan.

(100)	a.	Kòfí	kɔ́.	
		proper name	go.HAB	
		Kofi goes.		
	b.	Kòfí	kɔ́?	
		proper name	go.HAB	
		Kofi goes?		(Dolphyne, 1988:69)

In this chapter, in depth phonetic analysis of controlled data of the kind in (100) will be undertaken to explore the intonation of Yes – No questions and to determine what has to be regarded as an intonational morpheme in Akan. Furthermore, it will be investigated which of the observed effects are a by-product of the phonetic implementation of the intonational morpheme. This investigation will deepen the understanding of the interplay of phonological intonational tones with phonetic cues and will lead to a refinement of the classification of Akan along the lines of the prosodic typology of African languages proposed by Rialland (2007, 2009). It will be shown that sentences type affects the F0 pattern of Akan sentences locally, in terms of boundary tone insertion at the right edge of the IP, and globally, in terms of register raising.

Introduction to the question prosody in African languages: Rialland (2007, 2009) presents a preliminary typology of the prosodic marking of Yes – No questions in African languages. She provides an excellent overview, mostly based on impressionistic observations documented in the literature, and differentiates two main categories: languages with H-pitched, see (101)a.-e., and non-H-pitched, see (102)a.-f., question intonation. The former category includes languages that use at least one of the following phonetic cues:

(101) a. Final H tone or final rise
b. HL melody
c. Raising of the last H tone(s)
d. Register expansion and cancelation/reduction of downstep
e. Cancelation/reduction of final lowering (Rialland, 2007:35)

Final rising intonation is, according to Rialland, quite common but not evenly distributed in African tone languages mentioned in her survey. A few languages of the Kru and Gur family and a fair amount of Bantoid languages show a terminal rise. Rialland (2007:42f.) points to the difficulty of distinguishing final (H) tones and final boundary tones (H%) and provides insights into implementational differences. Post-lexical boundary tones associate with boundaries and lexical tones associate to TBUs. This difference unfolds on the surface in a distinct scaling; H% is assumed to be realized higher than final Hs; see also chapter 1 section 1.5.2 (33). Duration may be another phonetic cue to differentiate between final H and final H%. Rialland (2007:42) notes that "…a lexical tone is typically associated with a tone-bearing unit, while a boundary tones is linked to a boundary. The addition of a tone with its tone-bearing unit introduces a degree of lengthening determined by the nature of the tone-bearing unit (mora or syllable), while any lengthening triggered by a boundary tone is unrelated to the duration of the tone-bearing unit of the given language.".

Final post-lexical HL melody is quite uncommon. Only four languages in the database belonging to different language families (Gur, Bantu and Cushitic) exhibit this tonal marking.
Raising of the last Hs, which is regarded to be a local process, has to be distinguished from register expansion, which is regarded to be a global effect. This local process may involve raising of either the last H (also in non-final position), as illustrated in figure 77e. below, or of all Hs in the last phrase, as illustrated in figure 77e'. It has been reported for some Bantu, Chadic and Eastern Sudanic languages.

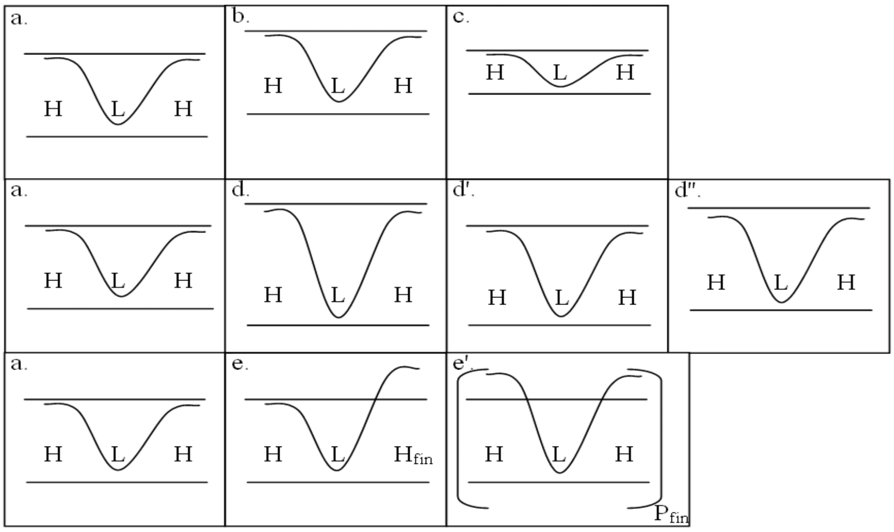


figure 77: Global (b-c) and local (d) variations in pitch range: a. neutral realization (baseline) of an HLH sequence in a hypothetical tone language; b. higher register (H and L raising); c. bottomline raising (L raising); d. span expansion (H raising and L lowering), d'. bottomline lowering (L lowering), d''. topline raising (H raising), e. raising of the last H, e'. raising of the Hs in the final phrase.

Expansion of the register (pitch range expansion) usually affects all tones of an utterance. The terminology used in Rialland's survey is not fully conclusive to me. Rialland (2007:39) writes "'Register expansion' refers to the expansion of the pitch range within which tones are realized. This expansion results mainly in the raising of H tones." Since Rialland notes that the magnitude of the register effect seems to be language-specific, it may be useful to differentiate precisely which registerline is manipulated. The figure 77, inspired by Ladd (1996), who was first to distinguish between range, register, level and span, and Gussenhoven (2004:77), shows possible register effects. Subpart a. of figure 77 serves as a baseline for comparison. The term register expansion (Rialland, 2007) may be described as lowering of the bottomline and raising of the topline. This will be referred to as span expansion, inspired by Gussenhoven (2004:77). The effect is illustrated in figure 77a. vs. figure 77d. It is in principle possible that only one reference line is involved in the creation of a greater pitch span; see figure 77d'. and figure 77d''. Rialland's statement that the register effect is mainly found in H tones suggests that the term topline raising seems to be more appropriate; see figure 77a. vs. figure 77d''. Other register effects might occur. These are illustrated in figure 77b. and figure 77c. I will refer to the former type as higher register, since both registerlines are raised in the tonal space relative to the baseline condition, a. The latter constitutes a subtype of a higher register since only the bottomline is raised. Rialland (2007:39) claims that span expansion i.e. topline raising (register expansion in her terminology) is generally affecting the downstep pattern. It is reduced or even canceled. The strength of this effect seems to be language-specific. Interestingly, register effects have been reported for a number of language families (Atlantic, Gur, Mande, Kwa (Ga), Ijoid, Buene-Congo (non-Bantoid and Bantu), Chadic, Cushitic, Eastern Sudanic and Khoisan). Note that register effects and effects on the downstep pattern are not distinguished in Rialland's survey.

Closer examination of the literature providing phonetic data reveals that languages differ in the domain in which the downstep is reduced or canceled. Myers (1996) provides an in depth analysis of the phenomenon in Chichewa. The figure 78 shows the intonational difference between a string identical Yes – No question and a statement.

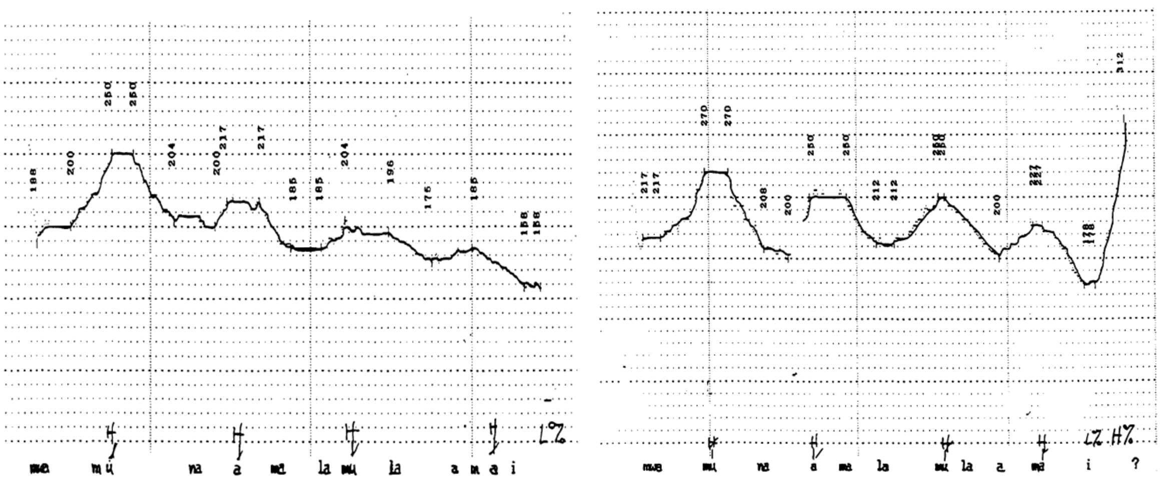


figure 78: Time normalized course of F0 for the sentences: *Mwamuna amalamula amayi.*/? 'The man bosses around the woman./?'; utterance final; one phrase; reproduced from Myers (1996:33f.).

Myers (1996) interprets the difference in terminal pitch excursion as manifestation of boundary tones. The statement exhibits an L%, whereas the Yes – No question is marked by an additional H%. The tones on the right hand side of figure 78 give the impression that the register is raised in Yes – No questions but since Myers (1996) only provides an analysis of the H tones, it could also be a topline raising effect; see figure 77d". Additionally, the downstep is significantly smaller for Yes – No questions. The downstep reduction spans the whole IP. He further addresses the questions whether to model the effects in the phonology (e.g. Inkelas & Leben, 1990) or in the phonetic implementation and convincingly shows that the observations are best captured with a phonetic implementation model (Liberman & Pierrehumbert, 1984); see (71) or (74). Manipulation of the baseline quotient (r) accounts most adequately for the reduction of downstep and the register effect i.e. the presence of H% in the phonological representation affects the phonetic implementation algorithm, specifically r. This view is reminiscent of what has been proposed for Akan in chapter 4 section 4.3.2. The presence of register tones associated to the IP node affects the lowering quotient (s).

The last cue under the heading high pitched Yes – No questions markers is reduction/cancelation of final lowering. This effect seems to be relatively rare. It is reported for a few languages from different families (Kwa (Ga), Buene-Congo Bantu, Chushitic and Eastern Sudanic).

Turning to the category of non-H-pitched question markers, Rialland (2007, 2009) subsumes languages which show at least one of the following characteristics:

(102) a. Final L tone or final fall
b. Polar tone or mid tone
c. Extra final lengthening
d. Breathy termination
e. Cancelation of penultimate lengthening
f. Insertion of a final [open] vowel (Rialland, 2007:35)

Thirty-six out of seventy-eight languages, belonging to different language families (Kru, Gur, Mande, Ijoid, Bantuid and non-Bantuid Benue-Congo, Chadic and Eastern Sudanic), in Rialland's database militate against the widespread assumption that questions are characterized by high or rising terminal pitch (e.g. Bolinger, 1978; Ohala, 1984, 1994; Gussenhoven, 2002, 2004). These languages utilize a final low or falling pitch to distinguish Yes – No questions from statements, including the Kwa languages Akan, Adiokrou, Baule, Ewe, Fon and Gun. Rialland (2007:44) notes that "The only difference between final Low tones and final falling intonations may lie in the timing of the pitch movement, intonational movements being generally associated with greater vowel lengthening". Unfortunately, I am not aware of any phonetic in depth study of an African tone language which uses L% as intonational cue of Yes – No questions; but see Byrd (1992) for optional use of L% in Nchufie; a Bantu language spoken in Cameroon.

Polar tones, or M in three-tone systems, refer to the use of H% after an L tone and L% after a H tone to signal sentence type (Clements & Rialland, 2008:76). This question marker is very rare and only reported for four languages in the database, which belong to different families (Mande, Kwa (Ga) and Bantu).

Twenty-three languages in the database use extra final lengthening as question marker from which two, Ntani (Gur) and Wobé (Kru), use it as the only cue. Rialland (2007:45) notes that lengthening is usually used in conjunction with other markers e.g. L% or H%.

Breathy termination is reported for some Gur languages and Hausa (Chadic). It can be associated with falling intonation as in Ncam (Gur) or with final lengthening as in Moba (Gur). Rialland notes that it may have gone unnoticed in other languages. Salffner (2010) observes for Ikaan (Buene-Congo) that Yes – No questions show a gradual final intensity decrease and breathy termination, whereas statements exhibit an abrupt intensity decrease and a glottal stop. Note that none of the Kwa languages has been reported to use breathy termination and/or final lengthening.

Languages that exhibit penultimate lengthening (see Hyman, 2009 for an overview), such as Zulu (Buene-Congo, Bantu) or Southern Sotho (Buene-Congo, Bantu), may cancel it to signal Yes – No questions. Rialland (2007:49) notes that it is sometimes associated with register expansion.

The last question marker to be introduced is insertion of an [open] vowel. It can be used as the only cue as in Vata (Kru) or Tikar (Buene-Congo, non-Bantu bantoid). According to Rialland, it is generally associated with a final L tone or falling intonation and is spread across different language families (Kru, Gur, Kwa (Adiakrou, Ewe, Fon and Gun), Buene-Congo (Bantoid and non-Bantoid), Chadic and Eastern Sudanic).

Non-H-pitched question markers are of particular interest from the standpoint of phonetic universals. According to the Frequency code, introduced in chapter 1 section 1.6.2, questions are expected to be realized with raised or high pitch. Gussenhoven (2004:79f.) remarks that in the unmarked case "…languages have structural, i.e. morphologically encoded, meanings at their disposal which are the same as the universal meanings …. in absence of any motivation to the contrary, the intonational morphemes of a language will reflect the universal form-function relations.". Rialland's (2007, 2009) survey has shown that grammatical forms that militate against the universal phonetic form are by no means rare. It is likely that the phonology of a language plays a crucial role in determining the form of an intonational morpheme. Gussenhoven (1999:302) notes that "…languages frequently also have intonation patterns with meanings that go against Ohala's "frequency code", notably falls signalling interrogatives and rises signalling statements…" and further claims that in those cases the phonetic implementation component will substitute for it. This substitution may be reflected in the third category of question markers mentioned by Rialland (2009) under the heading of "hybrid" question prosody.

"Hybrid" question prosody languages show characteristics from both sets of cues, such as Baule (Kwa) which uses L% and (maybe) register expansion, Bambara (Mande) which uses final -a/-wa and rising intonation, Ikaan (Salffner, 2010) which uses a raised/higher register, breathy termination and vowel epenthesis and Izon (Ijoid) which uses final L and H tone raising. Other Kwa languages which seem to combine characteristics from H-pitched and non-H-pitched question intonation are Dangme and Ga.

Kropp Dakubu (1986) reports for Yes – No questions in Dangme (not in Rialland's database), which is also spoken in Ghana, that the register is raised: "H is higher than in statements…especially utterance-finally. M is at about the pitch that H has in statements, L is at Mid pitch and level, except finally, when it glides down to the bottom of the normal register." (Kropp Dakubu, 1986:160). She reports for Ga that the register is raised and final tones show a terminal fall (downglide in her terminology). Note that Rialland classifies Ga as exhibiting a register expansion, cancelation of final lowering and polar tone in Yes – No questions, based on Kotei (1969).

Previous works on Akan question intonation: Sentence type (statement vs. question) in Akan may be marked syntactically, morphologically and/or intonationally. Combinations of grammatical devices are also possible. Since the focus of attention lies on the intonation of Yes – No questions here, the interested reader is referred to Christaller, (1875); Saah, (1988); Boadi, (1990); Marfo & Bodomo, (2005), Kobele & Torrence, (2006).

Dolphyne (1988:55f.) observes that the basic tonal pattern in Yes – No questions is maintained and that examples such as (100)a. & (100)b. above can be distinguished by an intonational difference. Several authors identify the terminal pitch excursion as the disambiguating force between Yes – No questions and statements in Akan (e.g. Dolphyne, 1988; Boadi, 1990; Abakah & Koranteng, 2007; Rialland, 2007, 2009). Though the identity (boundary tone or "grammatical" floating tone) and the phonetic implementation of the terminal pitch movement has not been subject to a controlled investigation, there is agreement on the observation that utterance-final H tones exhibit a sharply falling pitch movement in Yes – No questions. Concerning final L tones, two contrasting observations have been made. According to Dolphyne (1988:69), final H tones in Yes – No questions are realized as fall to the bottomline, whereas final L tones which are usually realized with a slight falling pitch in declaratives (final lowering), do not show this drop in questions (reduction/cancelation of final lowering). Abakah & Koranteng (2007:80) also identify the falling pitch movement on the final TBU as the relevant feature of Yes – No questions and state that it occurs irrespective of the underlying tonal specification. They derive the terminal pitch movement in Akan by assuming a suffixial L question tone which is unassociated, hence floating. L̥ co-anchors on the final TBU and causes final H and L tones to be realized with a falling pitch movement. However, Abakah & Koranteng (2007) observe that this type of grammatical tone behaves differently than "usual" grammatical tones and describe the association to an already L toned TBU as follows: "…the floating L of an interrogative morpheme docks leftwards…the linked tone of any TBU absorbs an identical tone when it docks or spreads to it, but here the process of tone absorption is put on hold even as the docking L co-shares the sentence-final TBU with its pre-associated L. This inevitably causes the final TBU to be produced on an extra L that results in a falling tone.” (Abakah & Korateng, 2007:81).

Berry & Aidoo (1975:18) and Dolphyne (1988:55) observe that Yes – No questions in Akan show a global F0 effect. Both L and H tones are scaled higher (higher register) in Yes – No questions compared to their statement counterparts. In addition to that, Hyman (2001) tentatively states that Yes – No questions in Akan show a suspension or reduction of downstep; see chapter 1 section 1.5.2 for further cues such as extra final lengthening (Christaller, 1875:97; Boadi, 1990:72) and extra voicing (Boadi, 1990:72).

Classification of Akan: Akan has been classified by Rialland (2007, 2009), along with other languages belonging to the Niger-Congo phylum, except Atlantic, Bantu and Kordofanian languages, as showing “lax” question prosody as opposed to “tense” question prosody which is associated with H-pitched question markers, as introduced in (101) above. “Lax” question prosody has been proposed to be an areal feature of the African Sudanic belt (Rialland, 2009). Characteristics of “lax” question prosody are according to Rialland, a falling intonation, lengthening, breathy termination and insertion of an open vowel e.g. [a] as question marker, as introduced in (102). A language showing the property of “lax” question prosody may use all or just a subset of the established markers. Rialland (2009:929) states: “We view falling intonation and breathy termination as resulting from a form of laryngeal relaxation, occurring at the end of questions….lengthening, might be viewed as a strategy for making these prosodic characteristics more salient, i.e. as a form of enhancement.”.

The literature overview on the topic has shown that Akan seems to combine phonetic features from H and L-pitched question markers (Rialland, 2007), which makes it an interesting test case also from a typological point of view. The table 21 summarizes the observations.

L-pitched “lax” prosody cues	**H-pitched “tense” prosody cues**
L% or terminal falling F0 (H) (e.g. Dolphyne, 1988; Abakah & Koranteng, 2007)	Reduction/cancelation of final lowering (L) (Dolphyne, 1988)
Extra final lengthening (Christaller, 1875; Boadi, 1990)	Reduction/cancelation of downstep (Hyman, 2001; Gussenhoven, 2004)
	Register raising (Berry & Aidoo, 1975; Dolphyne, 1988)
	Extra voicing (intensity increase?) (vowel space expansion?) (Boadi, 1990)

table 21: Summary of prosodic cues, found in the literature, marking Yes – No questions in Akan.

The phonetic details, the interplay of the phonetic cues and their implementation have not been subject to a quantitative experimental study. This enterprise will be undertaken here. It will lead to a refinement of Rialland's (2007, 2009) typology and provide methodological tools to investigate all of the reported cues; see table 21. The phonetic cues F0, duration, intensity and vowel quality will be examined to test the following hypotheses, derived from the literature summarized in table 21.

Hypotheses:

i. Yes – No questions exhibit a higher pitch register than statements; a higher pitch register results in a reduction or cancelation of downstep.
ii. Yes – No questions ending in an H tone are characterized by a low terminal excursion; Yes – No questions ending in an L tone lack final lowering.
iii. The final vowel in Yes – No questions is lengthened.
iv. The final vowel in Yes – No questions shows an intensity increase.
v. The final vowel in Yes – No questions undergoes a quality change.

Material & Measurements: Seven string-identical Yes – No question statement pairs constitute the data base for the exploration of the prosodic marking of sentence type in Akan. The material is explained in detail in chapter 2 section 2.2, block D, (49)a. - g.

To investigate the first hypothesis (i.), the initial F0 values in Yes – No questions and statements will be compared. Initial F0 values are assumed to be sufficient to detect a potential register effect (Huang, 1985; Liberman & Pierrehumbert, 1984). To examine the reduction/cancelation of downstep, the drops in pitch will be calculated between H_n-H_{n+1} (ΔDD) in Hz, except for the last H tone of items 4, 5 and 7. The Yes – No questions items 4 and 5 are displayed together with their statement counterparts in (103)a. and (103)b.

The ΔDD obtained for Yes – No questions will be compard to the ΔDD in the corresponding statement renditions.

(103)	a.	Pàpá	Kòfí	kàsá./?				(44)a. & (49)d.
		father	proper name	talk.HAB				
		'Father Kofi talks./?'						
	b.	Pàpá	Kòfí	kàsá	kyèrɛ́	nè	bá./?	(44)b.& (49)e.
		father	proper name	talk.HAB	point_out.HAB	PRO	child	
		'Father Kofi talks to his child./?'						

To explore the second hypothesis (ii.), the F0 values corresponding to the terminal pitch excursion will be obtained. A tonal marker was set at the last reliable extractable pitch point

(*T*). *T* in Yes – No questions will be compared to *T* in the corresponding statements. The lack of final lowering will be examined with the use of linear regression lines (Swerts et al., 1996), calculated in R. They simulate a linear function. Under the assumption that the F0 declines linearly (Maeda, 1976), the linear regression lines will be fit to Hz values of item 7, extracted every 10 ms with the help of a Praat script over the course of the final vowel. The difference (ΔL) between obtained and predicted value will be calculated for the values obtained over the last 30 ms (30ms will be referred to as antepenultimate position, 20 ms as penultimate position and 10 ms as ultimate position). ΔL obtained for Yes – No questions will be compared to ΔL in the corresponding statements. If the final L tones in Yes – No questions lack lowering, the obtained F0 value should be equal or higher than the predicted one; inspired by Liberman and Pierrehumbert (1984) and Prieto et al. (1996).
Two different measures of length of the final vowel were taken to investigate the third hypothesis (iii.). Since breathy termination, abbreviated BT, may occur, the duration will be measured from the beginning to the end of the formant structure (F2) and from the end of the formant structure to the end of phonation (Turk et al., 2006). The difference will be expressed as percentage (%L). The length of the final vowel in Yes – No questions will be compared to the length of the final vowel in the corresponding statements.
To explore the fourth hypothesis (iv.), the intensity (Int) maximum of the final vowel (MaxIntfinV) and the maximum overall Int (MaxIntU) of the utterance will be measured in db. A relative measure of Int (Int_{rel}) will be presented, inspired by Remijsen & van Heuven (2005)[45]. Relative Int will be obtained by subtracting the maximal Int of the final vowel from the maximal Int of the corresponding utterance. The formula is presented in (104).

(104) $Int_{rel} = MaxIntU - MaxIntfinV$

This procedure provides a relative measure of the Int (Int_{rel}) on the final vowel. Item 1 was discarded from the analysis since both values overlap. Int_{rel} obtained for the Yes – No questions will be compared to Int_{rel} in the corresponding statements. If Int is higher on the final vowel, Int_{rel} is expected to be smaller for Yes – No questions than for statements.
To test whether the final vowel in Yes – No questions is subject to a vowel quality change (hypothesis v.), the first three formants (F1, F2, F3) will be measured in Hz with the help of a Praat script at the mid of the final vowel (e.g. Remijsen & van Heuven, 2005; Asu, Schötz &

[45] Remijsen & van Heuven (2005:217) calculated the mean Int of a vowel and normalized it with the mean Int of the utterance; the mean overall Int was subtracted from the mean Int of the vowel.

Kügler, 2009), using the LPC method. All formant values (F) will be checked for implausible frequencies, in such cases re-measured and manually corrected. F1, F2 and F3 obtained for the Yes – No questions will be compared to F1, F2 and F3 in the corresponding statements.
The short renditions with alternating LH tones of speaker 5 (item 4) were discarded from the analysis of the initial F0 because F0 could not be reliably tracked, due to creakiness in the statement condition.

Results: The figure 79 gives a first impression of the results. The mean F0, aggregated over speakers and repetitions, for a sentence with alternating LH tones is displayed. The Yes – No question is represented by the dashed line. It is realized in a higher register than the statement, which is represented by the solid line. The register raising does not affect the terracing pattern. Moreover, both F0 curves cross each other in the final H toned vowel. The Yes – No question terminates lower than the statement.

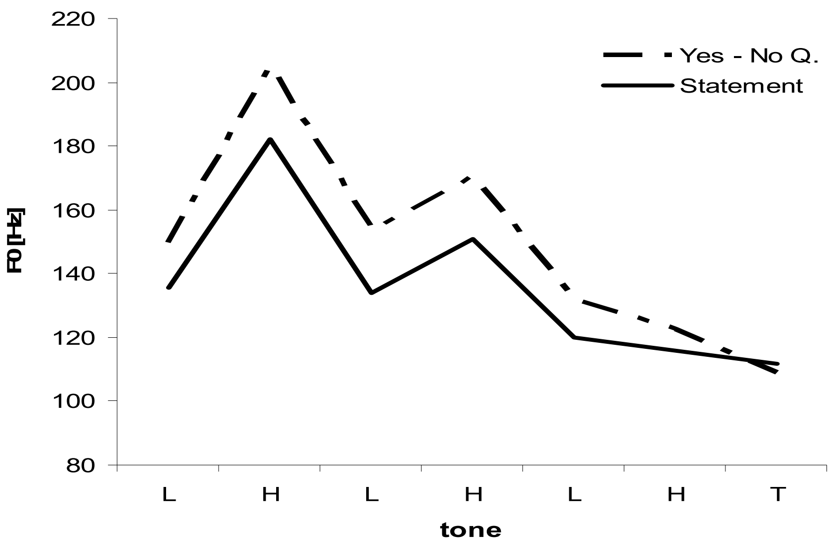


figure 79: Mean F0, aggregated over speakers and repetitions, of the sentence *Papa Kofi kasa.*/? 'Father Kofi talks./?', statement solid line, Yes – No question dashed line*, T* terminal F0; n=6.

Turning to the details of the observed effects, the mean F0 values for the initial tones reflecting the pitch register, averaged over items, speakers and repetitions, split by tones amount to 144.89 Hz (40) for the statements with an initial L tone, and to 152.31 Hz (42) for the Yes – No questions with an initial L tone. The mean F0 values for the sentences with an initial H tone are at 227.03 Hz (62) for the statement, and at 248.48 Hz (86) for the Yes – No questions. The table 22 presents the mean initial F0 values, averaged over speakers and repetitions, for all items split by sexes, tones and sentence type.

Item	Sex	Tone	Statement	Y/N Question
1	♀	L	187.55 Hz	191.43 Hz
1	♂	L	106.50 Hz	114.52 Hz
2	♀	H	275.23 Hz	311.06 Hz
2	♂	H	169.20 Hz	173.38 Hz
3	♀	L	197.41 Hz	205.48 Hz
3	♂	L	106.55 Hz	111.43 Hz
4	♀	L	191.50 Hz	208.63 Hz
4	♂	L	121.53 Hz	142.68 Hz
5	♀	L	208.35 Hz	212.86 Hz
5	♂	L	118.19 Hz	132.58 Hz
6	♀	L	199.66 Hz	215.88 Hz
6	♂	L	114.93 Hz	121.17 Hz
7	♀	L	190.93 Hz	215.26 Hz
7	♂	L	124.52 Hz	129.79 Hz

table 22: Initial F0 values, aggregated over speakers and repetitions and split by sexes, tones and sentence type.

A linear model was run on initial F0. Sentence type (Yes – No question/statement) and tone (L/H) were included as fixed factors. Speakers, items and repetitions were considered as random factors. The initial F0 values showed an effect of sentence type (t = 3.454, *p*MCMC <0.01, 185 observations), so does tone. This information is redundant. Importantly, there was no significant interaction between tone and sentence type. Thus both, L and H tones are raised in Yes – No questions.

The figure 79 has suggested that the downstep pattern is not affected by the register raising. This is also supported by the means obtained for the pitch drops of items with alternating LH tones. The first mean pitch drop (H1-H2), averaged over items, speakers and repetitions, amounts to 33.21 Hz (23) for the statements and to 34.01 Hz (21) for the Yes – No questions. The mean value for the second pitch drop (H2-H3), aggregated over items, speakers and repetitions, is at 20.44 Hz (12) for the statements and at 21.49 (12) for the Yes – No questions. The third mean pitch drop (H3-H4), averaged over items, speakers and repetitions, amounts to 14.65 Hz (7) for the statements and to 18.88 Hz (10) for the Yes – No questions. Finally, the fourth mean pitch drop (H4-H5) for item 7, aggregated over speakers and repetitions, is at 7.98 (4) for the statements and at 6.70 Hz (5) for the Yes – No questions. The table 23 presents the pitch drop values in Hz, averaged over speakers and repetitions, for all items, split by sexes, sentence type and positions. PoD refers to position of the drop.

Item	Sex	PoD	Statement	Y/N Question
4	♀	1	64.63 Hz	68.67 Hz
4	♂	1	26.84 Hz	25.61 Hz
5	♀	1	56.31 Hz	58.38 Hz
5	♂	1	29.36 Hz	31.39 Hz
5	♀	2	32.63 Hz	27.53 Hz
5	♂	2	20.04 Hz	20.54 Hz
5	♀	3	23.88 Hz	29.95 Hz
5	♂	3	13.68 Hz	15.32 Hz
7	♀	1	50.92 Hz	62.54 Hz
7	♂	1	28.29 Hz	29.86 Hz
7	♀	2	30.23 Hz	26.77 Hz
7	♂	2	17.26 Hz	20.18 Hz
7	♀	3	20.51 Hz	31.60 Hz
7	♂	3	12.37 Hz	17.43 Hz
7	♀	4	7.33 Hz	6.46 Hz
7	♂	4	8.31 Hz	6.57 Hz

table 23: Pitch drop for items 4, 5 & 7, aggregated over speakers and repetitions and split by positions, sentence type and sexes; PoD refers to position of the drop (1 = H1-H2; 2 = H2-H3; 3 = H3-H4; 4 = H4-H5).

A linear model was run on the pitch drop. Sentence type (Yes – No question/statement) was included as fixed factor. Speakers, repetitions, PoDs and items were considered as random factors. There was no significant main effect of sentence type; the downstep pattern is not systematically affected by the higher register in Yes – No questions.

Turning to the details of the terminal excursion, figure 80 presents the time normalized course of F0, aggregated over speakers and repetitions, for a final H toned vowel. Both curves differ in the first 30 ms, the Yes – No question represented by the dashed line has a higher F0 than the statement (solid line). During the last 30 ms of the Yes – No question the F0 falls towards the end.

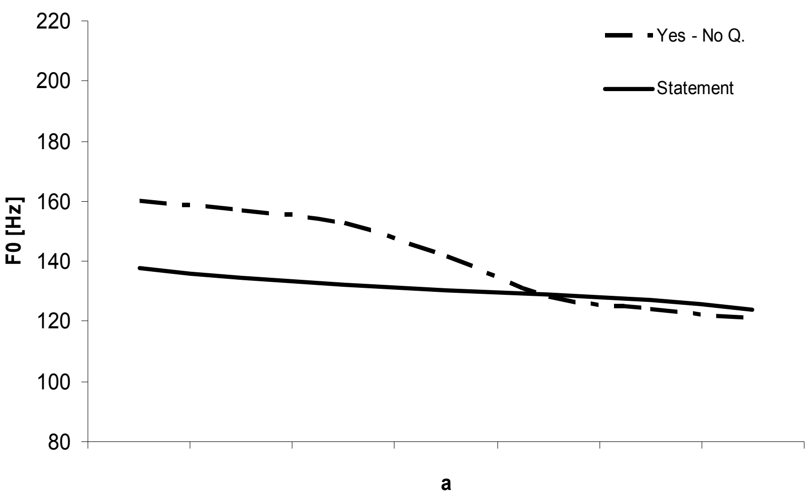


figure 80: Mean F0, aggregated over speakers and repetitions, taken from the last vowel of the sentence *Papa Kofi kasa*./? 'Father Kofi talks./?' 7 points of measurement: F0 extracted at the first 30 ms, mid of vowel and last 30 ms, statement solid line, Yes – No question dashed line*,* n=6.

The means for the terminal F0, averaged over items, speakers and repetitions, for the H tone amount to 125.75 Hz (35) for the statement and to 115.21 Hz (30) for the Yes – No questions. The figure 81 illustrates the time normalized course of F0, aggregated over speakers and repetitions, for the final L toned vowel. Similar to the final H tone, the F0 of the final L tone is higher for the Yes – No question than for the statement (dashed line) in the first 30 ms. However, in opposition to the final H tone, no fall is detectable in the last 30 ms of the Yes – No question.

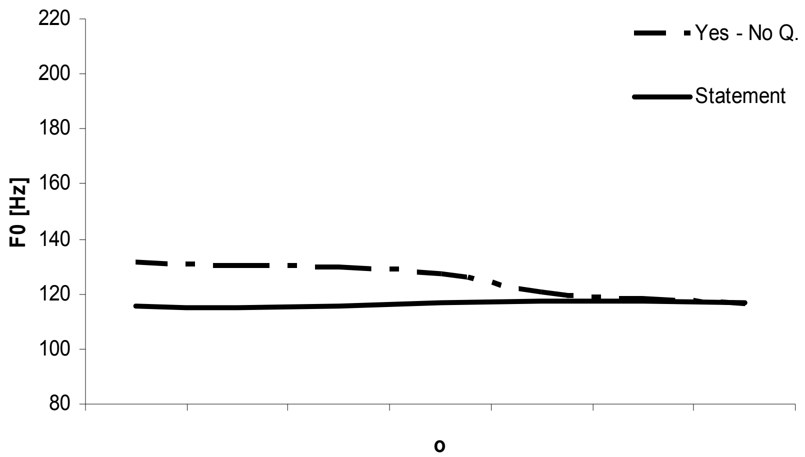


figure 81: Mean F0, aggregated over speakers and repetitions, taken from the last vowel of the sentence Papa Kofi kasa kyerɛ ne ba bio./? 'Father Kofi talks to his child again./?' 7 points of measurement: F0 extracted at the first 30 ms, mid of vowel and last 30 ms, statement solid line, Yes – No question dashed line*,* n=6.

The mean value, averaged over speakers and repetitions, for the terminal F0 of the L tone amounts to 101.27 Hz (24) for the statement and to 97.83 Hz (24) for the Yes – No questions. The table 24 presents the F0 values for the final element of all items in Hz, averaged over speakers and repetitions, split by items, sexes, tones and sentence type.

Item	Sex	Tone	Statement	Y/N Question
1	♀	H	185.79 Hz	154.92 Hz
1	♂	H	116.22 Hz	99.80 Hz
2	♀	H	155.81 Hz	141.19 Hz
2	♂	H	100.61 Hz	95.37 Hz
3	♀	H	167.48 Hz	150.01 Hz
3	♂	H	104.80 Hz	91.21 Hz
4	♀	H	159.62 Hz	148.74 Hz
4	♂	H	99.59 Hz	97.82 Hz
5	♀	H	163.59 Hz	157.52 Hz
5	♂	H	98.07 Hz	93.01 Hz
6	♀	H	162.21 Hz	145.83 Hz
6	♂	H	97.88 Hz	90.46 Hz
7	♀	L	151.36 Hz	147.46 Hz
7	♂	L	91.26 Hz	87.90 Hz

table 24: Final mean F0 values, aggregated over speakers and repetitions, split by items, sexes, tones and sentence type.

A similar model like the one for initial F0 values was calculated for the final F0. Final F0 values show an effect of sentence type (t = -3.568, pMCMC = 0.001, 196 observations); final tones are realized lower in Yes – No questions than in statements. Tone did not show a significant effect. The interaction between tone and sentence type approached significance (t = 1.808, pMCMC = 0.0722); final H tones in Yes – No questions exhibit a terminal falling F0, whereas final L tones do not.

Turning to the results for the lack of final lowering, the mean ΔL describes the difference for the final 30 ms between the linear regression line which was fit to Hz values of the final vowel of item 7 (predicted value) and the obtained values of the final vowel of item 7. Remember that, if the final L tones in Yes – No questions lack lowering, the obtained F0 value should be equal or higher than the predicted one. The mean ΔL for the statements, averaged over speakers and repetitions, for the antepenultimate position amounts to 0.05 Hz (3), for the penultimate position to -0.66 Hz (3) and for the ultimate position to -1.63 Hz (3). The mean difference (ΔL) for the Yes – No questions, averaged over speakers and repetitions, amounts to 0.35 Hz (2), to 0.06 Hz (3), and to 0.35 Hz (2) for the antepenultimate, penultimate, and ultimate positions respectively. The table 25 presents ΔL values for all three final positions, averaged over speakers and repetitions and split by sexes and sentence type. Position 1 refers to the antepenultimate value, position 2 to the penultimate value and position 3 to the ultimate value.

Position	Sex	Statement	Y/N Question
1	♀	-3.04 Hz	3.36 Hz
1	♂	0.71 Hz	-0.26 Hz
2	♀	-2.56 Hz	2.89 Hz
2	♂	-0.25 Hz	-0.51 Hz
3	♀	-1.20 Hz	1.03 Hz
3	♂	-1.72 Hz	0.21 Hz

table 25: Mean ΔL values, aggregated over speakers and repetitions, split by positions, sexes and sentence type. Position 1 refers to the antepenultimate value, position 2 to the penultimate value and position 3 to the ultimate value.

A linear model was run on ΔL. Sentence type (Yes – No question/statement) and position of the tone (antepenultimate/penultimate/ultimate) were included as fixed factors. Speakers and repetitions were considered as random factors. Sentence type showed a significant main effect (t = 1.988, pMCMC = 0.05, 105 observations); final L tones in Yes-No questions lack final lowering. The position of the tone did not have a significant effect and there was no interaction.

Turning to the durational results, figure 82 displays the mean duration, aggregated over items, speakers and repetitions, for the final vowel measured from the beginning of phonation till the end of F2 (V-F2) and from the end of F2 till the end of phonation (BT) for both sentence types. The Yes – No questions are represented by the white bar and the statements by the black bar.

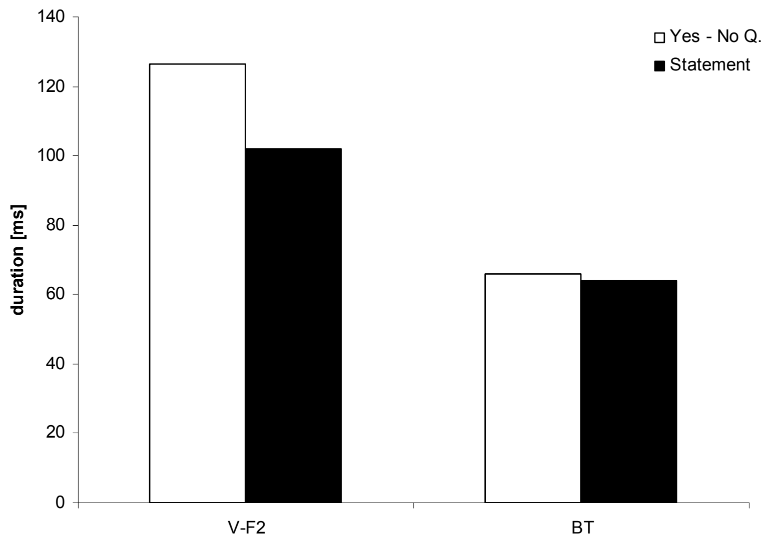


figure 82: Mean duration (V-F2), aggregated over items, speakers and repetitions, of the final vowel measured from the beginning of phonation (V) till the end of F2 and from the end of F2 till the end of phonation (BT), BT stands for breathy termination; the Yes – No questions are represented by the white bar and the statements by the black bar; n=17.

In the case of BT no difference between the sentences types can be observed. The mean value, aggregated over items, speakers and repetitions, amounts to 63.94 ms (32) for the statement and to 66.08 ms (34) for the Yes-No question. To ensure that BT does not play a significant role, a linear model was run on BT. Sentence type (Yes – No question/statement) was included as fixed factor. Speakers, repetitions and items were considered as random factors. There was no significant effect of sentence type; BT (145 observations[46]) does not differ as a function of sentence type. Therefore, it is suitable to work with the total length of the vowel (V-F2+BT). The means, averaged over items, speakers and repetitions, for the total length of the vowel amount to 145 ms (43) for the statements, and to 180 ms (58) for the Yes – No questions. This corresponds to a lengthening of 19 % of the vowel in Yes – No questions. The table 26 presents the durational values, averaged over speakers and repetitions, split by items, tones and sentence types. The L toned vowel is also lengthened by about 19 % on average.

[46] The whole data set contains 196 realizations. Altogether, 51 realizations did not show any sign of BT. Thereof were 19 Yes – No questions.

Item	Condition	Tone	Total V length
1	S	H	115 ms
1	Q	H	157 ms
2	S	H	144 ms
2	Q	H	177 ms
3	S	H	116 ms
3	Q	H	132 ms
4	S	H	164 ms
4	Q	H	233 ms
5	S	H	179 ms
5	Q	H	216 ms
6	S	H	137 ms
6	Q	H	130 ms
7	S	L	133 ms
7	Q	L	165 ms

table 26: Mean duration of the final vowel, aggregated over speakers and repetitions, split by items, tones and sentence types.

A linear model was run on the total length of the vowel. Sentence type (Yes – No question/statement) and tone (L/H) were included as fixed factors. Speakers, repetitions and items were considered as random factors. The length of the final vowel showed a significant effect of sentences type (t = 4.762, pMCMC = 0.0001, 196 observations). There was no significant interaction; the final vowel is lengthened in Yes – No questions irrespective of its tonal specification.

Turning to the results for the intensity of the final vowel, the means, aggregated over items, speakers and repetitions, for the relative intensity amount to 11 db (5) for the statements and to 8 db (5) for the Yes – No questions. The table 27 presents the mean values, aggregated over speakers and repetitions, split for items, tones and sentence type, for MaxIntU, MaxIntfinV and the difference between the two (Int_{rel}).

Item	Condition	Tone	MaxIntU	MaxIntfinV	Int_{rel}
2	S	H	80 db	74 db	6 db
2	Q	H	80 db	76 db	4 db
3	S	H	82 db	73 db	9 db
3	Q	H	83 db	72 db	11 db
4	S	H	79 db	69 db	10 db
4	Q	H	79 db	74 db	4 db
5	S	H	79 db	69 db	10 db
5	Q	H	79 db	74 db	5 db
6	S	H	80 db	69 db	11 db
6	Q	H	82 db	71 db	11 db
7	S	L	79 db	64 db	15 db
7	Q	L	80 db	67 db	14 db

table 27: Mean intensity, aggregated over speakers and repetitions, split by items, tones and sentence type; Maximal intensity of the utterance (MaxIntU), maximal intensity of the final vowel (MaxIntfinV) and the difference between the two (Int_{rel}).

Generally, the intensity of the L toned vowel is lower than for the H toned vowels. Int_{rel} is also lower for the Yes – No questions.

A linear model was run on Int_{rel} specified like the ones for initial and final F0. Sentence type (t = -3.395, pMCMC < 0.005, 174 observations) and tone (t = 2.544, pMCMC < 0.05) showed a significant effect, there was no significant interaction. Both, final L and H tones show an intensity increase in the final vowel. Final L tones exhibit a generally lower intensity than final H tones.

Turning to vowel quality measured in terms of formant structure, figure 83 gives an overview of the results for the final /a/ and final /o/, /ɔ/ vowels. All three formants are plotted, split by sentence type and vowel quality. Generally, the low vowel has a higher F1 and F2 than the non-low vowels, but a lower F3. Sentence type affects the formants but vowel quality seems to play a role. For the low vowel, F1 and F2 are higher for the Yes – No questions, but F3 is slightly lower, whereas for the non-low vowels, only F3 is realized higher for the Yes – No questions.

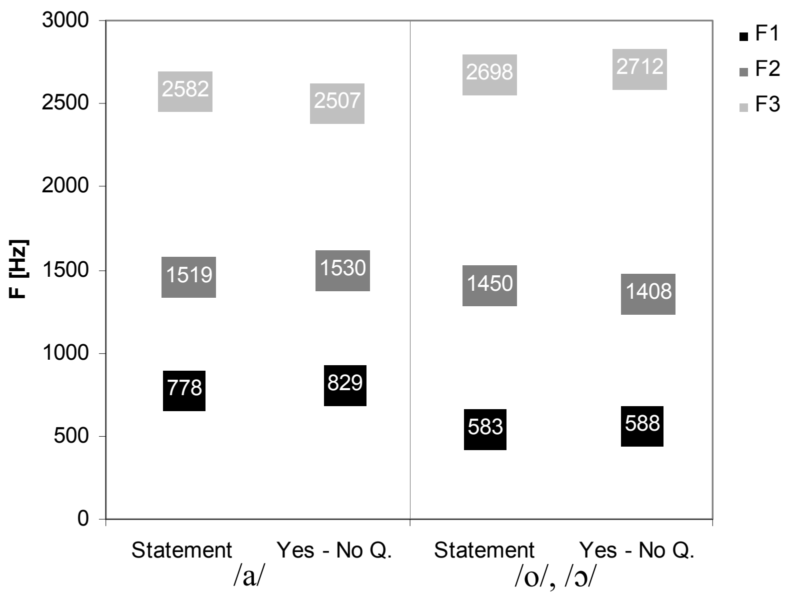


figure 83: Mean vowel formants (F1, F2, F3) of the last vowel, aggregated over speakers, items and repetitions, split by sentence type and vowel quality; /a/ left panel; /o/, /ɔ/ right panel; F1 black quadrate, F2 grey quadrate, F3 light grey quadrate; n=17.

The mean value, averaged over vowels, items, speakers and repetitions, for F1 is at 654.88 Hz (205), at 1408.01 Hz (414) for F2 and at 2634.57 Hz (374) for F3 in the statement condition. The mean F1 value of the Yes – No questions amount to 671.79 (208), to 1393.45 (433) for F2 and to 2619.82 Hz (388) for F3. The table 28 presents the mean formant values for each item, aggregated over speakers and repetitions, split by sentence type, tones and position in the frequency band.

Item	Condition	Tone	Vowel	F1	F2	F3
1	S	H	ɔ	585.78 Hz	847.04 Hz	2512.92 Hz
1	Q	H	ɔ	545.55 Hz	880.48 Hz	2559.68 Hz
2	S	H	ɔ	596.01 Hz	977.98 Hz	2587.05 Hz
2	Q	H	ɔ	544.86 Hz	900.81 Hz	2494.64 Hz
3	S	H	o	611.76 Hz	1449.08 Hz	2647.51 Hz
3	Q	H	o	541.68 Hz	1444.82 Hz	2529.07 Hz
4	S	H	a	744.17 Hz	1533.46 Hz	2571.73 Hz
4	Q	H	a	786.38 Hz	1522.45 Hz	2479.06 Hz
5	S	H	a	812.01 Hz	1505.40 Hz	2592.95 Hz
5	Q	H	a	871.12 Hz	1537.97 Hz	2535.79 Hz
6	S	H	o	554.65 Hz	1753.76 Hz	2697.25 Hz
6	Q	H	o	583.42 Hz	1593.34 Hz	2723.96 Hz
7	S	L	o	574.25 Hz	1554.41 Hz	2796.20 Hz
7	Q	L	o	646.08 Hz	1580.93 Hz	2949.65 Hz

table 28: Mean formant values (F1, F2, F3), aggregated over speakers and repetitions, split by items, vowels, tones and sentence type.

A linear model was run on F. Sentence type (Yes – No question/statement), position in the frequency band (F1/F2/F3) and vowel quality ([+LOW], [-LOW]) were included as fixed factors. Speakers, repetitions and items were considered as random factors. Sentence type did not show a significant effect. A significant effect was observed for the contrasts among the positions of F in the frequency band (F1, F2, F3). This information is redundant. Vowel quality did not show a significant effect. However, there was a significant interaction between F1 and vowel ($t = 3.35$, *p*MCMC < 0.001, 588 observations) and F2 and vowel ($t = 2.68$, *p*MCMC < 0.01); low vowels have a higher F1 and F2 than non-low vowels. All other comparisons did not yield a significant result.

Summary: Regarding the first hypothesis (i.), the analysis of the data has shown that Yes – No questions are uttered in a higher pitch register than statements. This is in line with observations made by Berry & Aidoo (1975) and Dolphyne (1988); see figure 79 for illustration. Both, initial L (~10Hz) and H tones (~20Hz) are raised. The investigation of the terracing pattern has disclosed that the register raising effect does not affect declination/ downstep in Akan, contrary to Hyman's (2001) and Gussenhoven's (2004) claim.

The investigation of final lexical tones (hypothesis ii.) has revealed that final H tones in Yes – No questions exhibit a lower terminal excursion than in statements. The lowering is about 10 Hz. This finding is in line with observations made by several authors; e.g. Dolphyne, 1988; Boadi, 1990; Abakah & Koranteng, 2007; Rialland, 2007, 2009; see e.g. figure 80 for visualisation. Final L tones in Yes – No questions lack final lowering; see table 24. This observation is in line with Dolphyne (1988).

With regard to the third hypothesis (iii.), the data has shown that the final vowel is lengthened by about 19% in Yes – No questions irrespective of its tonal specification; supporting Christaller's (1875) and Boadi's (1990) claims. Two measures have been presented: first, from the beginning of phonation to the end of F2 and second, from the end of F2 till the end of phonation. Breathy termination is not systematically used to mark Yes – No questions; see figure 82 for illustration.

Furthermore, Boadi's (1990) observation that the final element in Yes – No question exhibits extra voicing (hypotheses iv. & v.) was explored. Contrary to Rialland (2009), the data revealed that the final vowel in Yes – No question is subject to an intensity increase; measured in relation to the maximal intensity of the whole utterance. The effect appeared irrespective of the tonal specification though final L tones exhibited generally less intensity than final H tones which is expectable (Zee, 1978); see table 27 for details.

With regard to the vowel space, measured as formant structure, a slight but non significant tendency for F1 and F2 in final low vowels of Yes – No questions to be realized higher than in statements and the same tendency for F3 in final non-low vowels has been observed; see figure 83.

As a closing visualization of the results in reference to Rialland's (2007, 2009) typology, table 21 is repeated with checkmarks for confirmation. The majority of phonetic cues have been confirmed with the help of an experimental controlled setting.

L-pitched "lax" prosody cues	**H-pitched "tense" prosody cues**
L% or terminal falling F0 (H)✓ (e.g. Dolphyne, 1988; Abakah & Koranteng, 2007)	Reduction/cancelation of final lowering (L)✓ (Dolphyne, 1988)
Extra final lengthening✓ (Christaller, 1875; Boadi, 1990)	Reduction/cancelation of downstep (Hyman, 2001; Gussenhoven, 2004)
	Register raising✓ (Berry & Aidoo, 1975; Dolphyne, 1988)
	Extra voicing (intensity increase?)✓ (vowel space expansion?)✓ (Boadi, 1990)

Discussion: In the discussion, the underlying phonological trigger of the falling movement located on the final element of Yes – No questions will be characterized. Although this is relatively uncharted territory, Rialland (2007) presents two cues, scaling and duration, to distinguish a final tone from a final boundary tone.

The difference between a final lexical H tone and a final H% is distinguishable by its scaling. H% is realized higher than H (Rialland, 2007:42). L% may not be realized lower than L. However, one might expect that L% following a final lexical H tone is characterized by a falling F0; see also chapter 1 section 1.5.2 (33) for an overview. That is what the current data shows. Hence, it will be proposed that the intonational morpheme marking Yes – No questions in Akan is a low boundary tone (L%) which associates to the right edge of the IP (e.g. Pierrehumbert & Beckman, 1988). It thus meets the description of a "lax" prosody language (Rialland, 2007, 2009) in terminating low.

The definition of a "lax" prosody language is crucially based on the view that L% results "…from a form of laryngeal relaxation, occurring at the end of questions." (Rialland, 2009:929). However, it will be argued that Akan does not meet that part of the description because L% is phonetically implemented by means of increasing the articulatory effort. Departing from the phonological analysis, it will be explored whether the register raising effect can be attributed to the phonological or the phonetic component. Proposals offering a phonological analysis, involving a left edge H boundary (Sosa, 1999 cited in Prieto, 2004 for Latin-American Spanish) or insertion of an H register tone (Inkelas & Leben, 1990 for Hausa), are based on languages which are characterized by a final rising pitch movement (H%). This is also true for proposals attributing the register effect to the phonetic implementation (Myers, 1996 for Chichewa; Haan, 2002 for Dutch). The grammaticalization of L% as question marker clearly contradicts the universal form-meaning relation expected from the Frequency code (Ohala, 1994; Gussenhoven, 1999, 2002, 2004). The higher register will be analyzed as substitution of the Frequency code at the phonetic level; alternative analyses will be discussed and rejected.

As pointed out above, the durational results may play a role in determining the nature of the terminal falling pitch movement. Generally, durational differences that go along with lexical tones are related to the TBU of a language, whereas the lengthening that may accompany a boundary tone is unrelated to it (Rialland, 2007:42f.). The data has shown that the final vowel in Yes – No questions is lengthened about 19% irrespective of the tonal specification; see table 26 for details. Following Rialland's (2007) suggestion that a lengthening accompanying a boundary tone is unrelated to the TBU of a language, the following figure 84 provides information of the amount of lengthening brought about by adding a TBU. The figure 84 compares the duration of a final /a/, uttered as a statement (black bar on the left side), to the duration of a final /aa/, also uttered as a statement (black bar on the right side), see (42)c. The mean value, aggregated over speakers and repetitions, amounts to 164 ms (55) for /a/ and to 312 ms (81) for /aa/. The duration of a final /a/, uttered as Yes – No question (white bar), is at 233 ms (59), which is shorter than /aa/, uttered as a statement. A paired samples *t*-test was carried out in R. The difference between the lengthened /a/ in Yes – No questions and /aa/ in statements is significant ($t(5)= -3.8905$, p-value = 0.001).

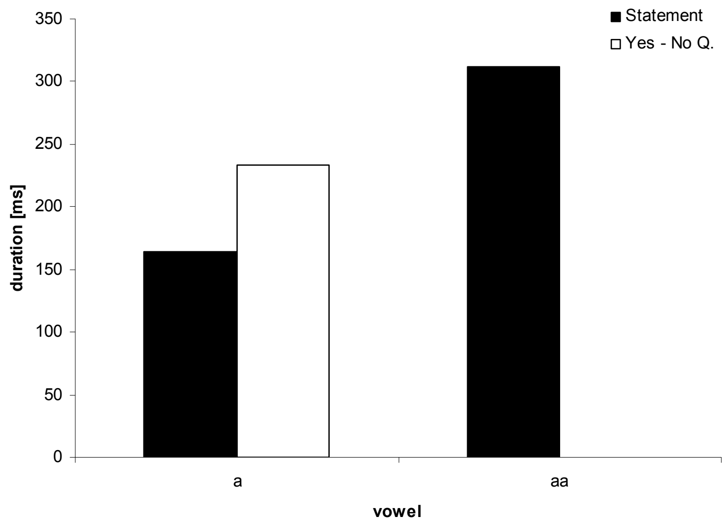


figure 84: Mean duration, aggregated over speakers and repetitions, measurement for /a/ taken from the final vowel of the statement/Yes- No question *Papa Kofi kasa*./? 'Papa Kofi talks./?'; measurement for /aa/ taken from the final vowel of the statement *Kukuoba papa no bɔ daa*. 'The good small pot breaks everyday.'; n=6.

The lengthening, observed in Yes – No questions, is hence not related to the TBU in Akan. Thus, the durational facts speak in favor of assuming an additional intonational low target (L%) at the right edge of Yes – No questions.[47]

Scaling is the other cue helping to determine whether L% is present. Intensity will be included in the discussion because F0 and intensity are intertwined (e.g. Vaissière, 1983). The data presented for Akan has shown that final H tones in Yes – No questions in Akan show a final falling F0; see e.g. figure 80. Before concluding that the pitch movement is triggered by L%, one might consider the possibility that the reason for the lower termination is the longer duration i.e. there is more time for the F0 to decline in the end of a sentence (Byrd, 1992:19).

[47] Note that the connection between boundary tone and lengthening is not absolute. Smith (2002:163) shows that for French the boundary tone is not responsible for the lengthening but the sentence type. She concludes that lengthening might be more required in questions than in statements to mark the end of a conversational turn; see also Callier (2011:5). He observes for Mandarin Chinese that extra final lenthening is used to signal turn-yielding and concludes that "…lengthened utterances have something of an "other-oriented" quality, either yielding the floor or soliciting a sympathetic reaction." (Callier, 2011:16).

However, since final lowering is generally connected to a decrease in intensity (Vaissière, 1983:58), and the data has shown that the terminal element in Yes – No questions is subject to an intensity increase[48], final lowering can be excluded as possible analysis for the final falling F0. Additionally, the behavior of final L tones in Yes – No questions speaks against a final lowering analysis. They lack final lowering, although they are significantly lengthened. Thus, the extra lengthening can not be responsible alone for the abrupt decline of the final H tones in Yes – No questions.

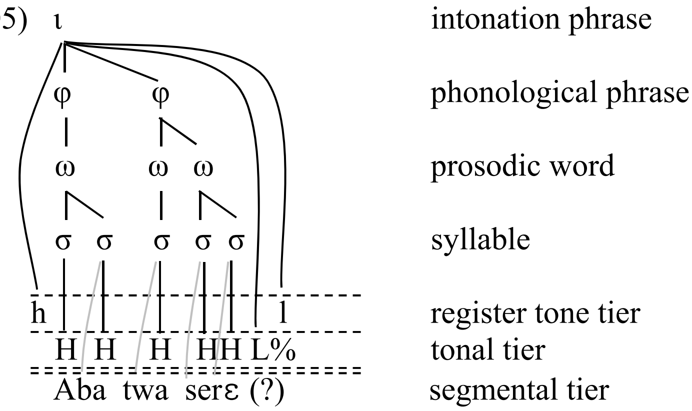
Alternatively, Hyman (p.c.) suggested that the difference between final lexical L and H may be due to a polar tone at the right edge of Yes – No questions in Akan. Recall that Rialland (2007) classified the lack of final lowering as H-pitched question marker. A polar tone would manifest itself as "…a H tone after a L tone and as a L tone after a H tone." (Clements & Rialland, 2008:76). Two arguments speak against this analysis. Firstly, a final H toned question marker should have a noticeable local pitch effect e.g. terminal rise after the lexical L tone. This is not what the present data shows. Secondly, H% could potentially affect the register and thereby reduce or cancel the downstep pattern, which is not observable in the data. It rather seems that the speaker already reaches his/her bottomline when producing a final lexical L tone and there is thus no space to produce an additional fall.

The longer duration of final L toned vowels suggests that there is an additional tonal target (L%); see also Mücke & Grice (2006) for lengthening effects due to tonal crowding in German and Fedden (2007:52) for Mian. But the higher intensity on the final vowel prevents the final L tone to undergo additional lowering. Thus, irrespective of the tonal specification of the final vowel, L% is present at the right edge of Yes – No questions in Akan. Analyzing the terminal excursion in Yes – No questions in Akan as L% elegantly solves Abakah & Koranteng's (2007:80) problem, that their L̥ question morpheme behaves differently than "usual" grammatical tones. Usually, grammatical tones delete the underlying lexical tone(s) in a certain domain. Since intonational tones convey meanings that apply to phrases or utterances as a whole (Ladd, 1996:7f.) and their presence may have local and/or global scaling effects, it has been suggested that they associate to the IP node in the prosodic tree (Pierrehumbert & Beckman, 1988). Thus, I propose that the mental construct of a Yes – No question in the mind of an Akan speaker can be represented as in (105). (105) shows the sentence *Aba twa serɛ*? 'Aba cuts grass?' consisting of lexical H tones, since it is a question

[48] Intensity is one of the lesser studied prosodic cues. According to Vaissière (1983:62), higher intensity often correlates with longer duration.

(indicated by (?) on the segmental tier)[49] an L% is associated with the right edge of the IP; see figure 43 (right hand side) for a pitch track of this sentence. The figure 43 shows that declination is also present in Yes – No questions. Thus the register tone specification is also present in the phonological representation of questions in Akan.

(105)



The proposal makes several predictions. First, since L% is represented on the tonal tier, it should influence the phonetic implementation of the lexical tone associated to the final vowel. Experimental evidence has been presented; e.g. lack of final lowering in the case of a lexical L, and the vowel itself, which is reflected in a longer duration and higher intensity. Second, tones on the register tier should not be affected by the presence of L% on the tonal tier. Hence, it is expected that declination is present in Yes – No questions to the same degree as in statements. The data at hand provided evidence that this is the case. The terracing pattern in Yes – No questions is similar to that in statements. Third, since L% associates to the IP, its presence should be available for the phonetic implementation; see also chapter 4 section 4.3.1 and chapter 5. As pointed out above, early effects of boundary tones are usually reported for languages which employ H% as a question marker. However, the data at hand has shown that a higher pitch register goes along with L%.

Turning to the discussion of the observed register raising effect, a possible phonological account for the higher register in Yes – No questions could be to assume an initial or left edge H boundary tone, as Sosa (1999) suggested for Latin-American Spanish. Three reasons speak against this analysis. First, as for final H%, one would expect an initial H% to be scaled higher than an initial lexical H tone (Rialland, 2007). The pitch on the first TBU of a Yes – No question should be falling, and even more sharply if the initial TBU is associated with a lexical L tone. The data for Akan has not revealed such an effect; see figure 77 for illustration.

[49] The questions mark is added for illustrative purpose only. It is not assumed that it is part of the phonogical representation.

Second, the register effect in Yes – No questions resembles the anticipatory raising effect elaborated in chapter 5 section 5.2, which is clearly a phonetic effect. This is illustrated in figure 83.

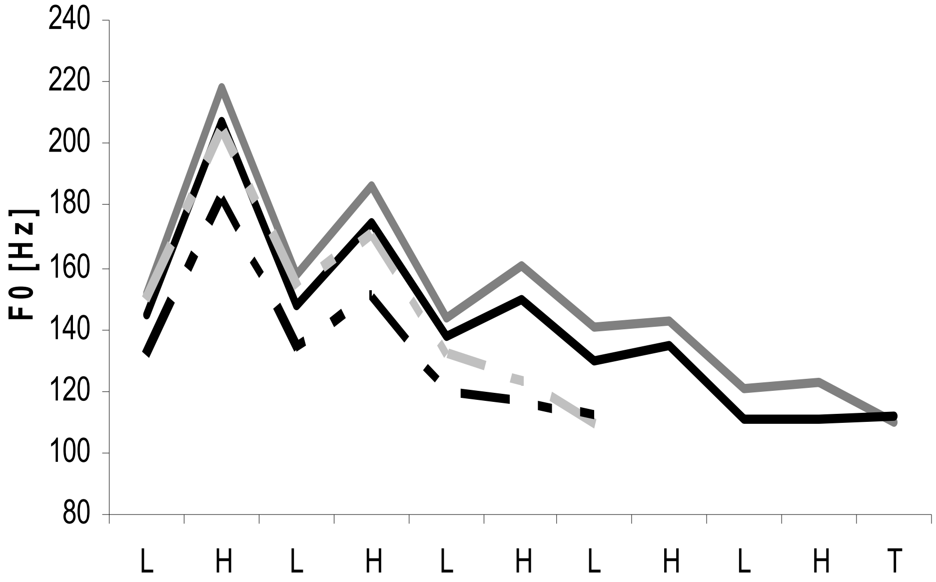


figure 85: Time normalized course of F0, aggregated over speakers and repetitions, of the sentence *Papa Kofi kasa*./? 'Father Kofi talks./?', Yes – No question grey dotted line; statement black dotted line and *Papa Kofi kasa kyerɛ ne ba*./? 'Papa Kofi talks to his child./?'; Yes – No question grey solid line; statement black solid line; n=6.

The grey dotted line refers to a short Yes – No question and the black dotted line to a string-identical statement with the same length (6 syllables). The black solid line refers to a longer statement (10 syllables) and the grey solid line to a string-identical Yes – No question with the same length.

Comparing the initial values of the short Yes – No question (grey dotted line) to the initial values of the longer statement (black solid line), it can be observed that they nearly lie upon each other. Since sentence length is not a grammatical factor such as sentence type, it seems not appropriate to postulate an intonational morpheme (initial H%) to account for the observed effect in Yes – No questions. Third, if the register effect would be due to a phonological trigger it should be perceptually relevant. However, preliminary results from a forced choice task, involving cross-spliced stimuli (higher register i.e. Yes – No question beginning + final fall i.e. statement ending and vice versa), in which listeners were asked to choose whether they heard a question or a statement revealed that the final element is the perceptual relevant area. Furthermore, preliminary results from a gating experiment showed that listeners confidently decide for one sentence type when they heard the final element.

However, this connection seems not to be an absolute one. If listeners are familiarized with the pitch range of a speaker, they are able to take a confident decision earlier.

Having rejected a phonological analysis of an H%, it is likely that the register effect in Akan arises during phonetic implementation. One possibility is that it is a by-product of an overall higher intensity in Yes – No questions. Myers (2004) shows for Chichewa that loudness, which correlates with a higher intensity, raises the F0 globally. Heeren & van Heuven (2011) report an overall higher intensity for Yes – No questions in Dutch. For Akan, some of the maximal intensity values are slightly higher in Yes – No questions; see table 26. Intensity has been shown to be dependent on the nature of lexical tone (Zee, 1978) and on the quality of the TBU. Zee (1978) investigated the relationship of tone and intensity in Taiwanese, which has five tonal oppositions: H[50], mid (M), L, high-falling and low-raising. He shows that intensity correlates with tonal height. The H tone exhibits higher intensity than the M tone and the M tone exhibits a higher intensity than the L tone; see table 26 for data on Akan. Apart from F0, sonorous elements have been found to exhibit a higher intensity (Clark & Yallop, 1990:282); [-HIGH] vowels are more sonorous than [+HIGH] vowels (Parker, 2002:240).

To test whether the higher register is triggered by an overall higher intensity in Yes – No questions in Akan, a subset of the data was compiled to calculate the correlation between the two variables. It consists of 41 polar questions[51]. To minimize the effect of tone and vowel quality, intensity in db and F0 in Hz (later converted into st), were measured at the temporal mid of the initial L toned vowel /a/ of items 4, 5 and 7. The figure 86 illustrates the relationship between intensity and F0.

[50] For further details about the relationship between F0 and intensity of contour tones; see Zee (1978:215f.).

[51] Thirteen renditions were excluded from the analysis because they either did not show a higher F0 on the initial vowel or Intensity or F0 could not be tracked reliably. Thanks to Ruben van de Vijver for suggesting the correlation method to me.

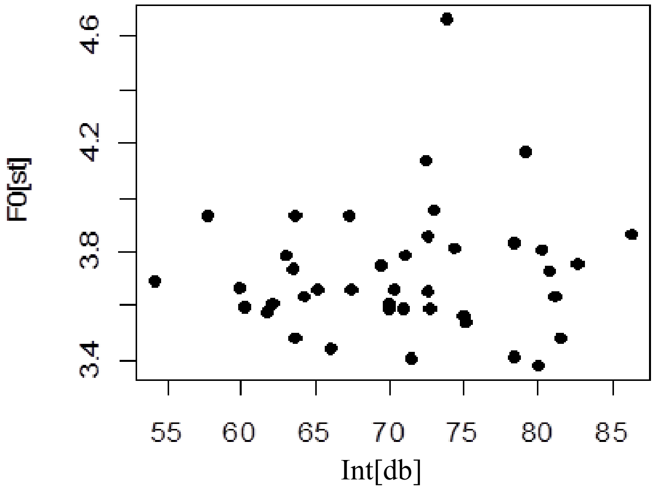


figure 86: Scatterplot displaying F0 (st) and intensity (db) values, measured at the mid of the initial vowel of the Yes – No questions *Papa Kofi kasa*? 'Father Kofi talks?', *Papa Kofi kasa kyerɛ ne ba*? 'Father Kofi talks to his child?' and *Papa Kofi kasa kyerɛ ne ba bio*? 'Father Kofi talks to his child again?'; n = 41.

No linear relation between F0 and intensity can be observed. This is also expressed by the statistical analysis; Pearson's r = 0.07. Pearson's correlation quotient is statistically not different from 0. The two parameters do not correlate. The F0 raising thus occurs independently of the intensity.

Having rejected the possibility that the higher register is triggered by a higher overall intensity in Yes – No questions, it may be the case that it reflects a phonetic compensation in order to satisfy a "pre-linguistic", biological and universal constraint of questions/interrogatives to involve some kind of high(er) pitch (Gussenhoven, 1999:302). Interrogativity should correlate with high(er) pitch because "Asking a question amounts to making an appeal to a listener. This presupposes dependence on the part of the speaker who, after all, relies on the listener's willingness to oblige with an answer." (Haan, 2002:148). The interrogator is thus in an inferior position. Since the question morpheme in Akan has developed a phonological form (L%) which goes against the Frequency code, the higher register may be analyzed as a phonetic compensation (Gussenhoven & Chen, 2000; Gussenhoven, 1999, 2002; Haan, 2002). The very prominent terracing pattern of Akan analyzed as declination, see chapter 4, may have given rise to the development of phonologized intonational forms that involve F0 lowering instead of raising.

As a concluding remark, I would like to take up on Rialland's (2007, 2009) typological classification. At a first glance, Akan seems to fall into the group of hybrid systems since it involves H-pitched question markers, such as lack of final lowering and a higher register, and L-pitched question markers, such as extra final lengthening and terminal falling F0. However, it has been argued here that the lack of final lowering and the higher register are not due to an initial boundary tone H% i.e. are not H-pitched question markers in Akan. Phonologically, Yes – No questions in Akan are marked with a right edge L% tone; see (105). The presence of an additional tonal target at the right edge on the tonal tier leads to an intensity increase which causes the lack of final lowering and extra final lengthening. The phonetic component of the grammar employs the higher register to compensate for a phonological form which goes against the Frequency code.

Rialland (2007, 2009) classified Akan as a “lax” prosody language. However, the investigation of Yes – No questions in Akan presented in this chapter has shown that its intonation is characterized by “tense” prosody. The right edge boundary tone L% is phonetically implemented by means of an increase in articulatory effort (higher intensity and vowel space expansion). Furthermore, L% is not accompanied by breathy termination. Hence the term “lax” seems not to be appropriate. I therefore suggest to extend Rialland’s typology by a new category called “low tense” question prosody languages. “Low tense” languages show a final falling F0 movement, but the implementation of the fall does not involve “lax” characteristics due to relaxation of the larynx.

This chapter has shown that the type of a sentence can be distinguished by the absence or presence of a post-lexical low boundary tone (L%). Akan speakers, thus, have intonational means at their disposal to express pragmatic meaning, despite the fact that tone is already used to express lexical and grammatical distinctions. The following chapter will further draw on the issue of pragmatic meaning. The details of focus marking will be explored.

7. Chapter
The expression of focus

The focus of a sentence, e.g. the element in an answer replacing the wh-word of a preceding question, can be seen as the most informative part of the answer (e.g. Baumann, Grice & Steindamm, 2006; Krifka, 2007). In this chapter, it will be explored which grammatical means Akan speakers have at their disposal to highlight a focused element, i.e. to make it more salient to the listener. Section 7.1 will be concerned with the syntactic marking of (narrow corrective) focus shedding light on the question of the frequency of the ex-situ focus construction with the help of a situation description task (Genzel & Kügler, 2010). The prosodic marking of focused in-situ and ex-situ objects will be investigated in section 7.2. A reanalysis of already published data (Kügler & Genzel, 2012) will be provided and interesting new insights from semi-spontaneous data will be presented. It will be shown that focus affects the F0 pattern of Akan sentences only marginally, if the focus is "strong", i.e. contrastive/corrective. However, it will be demonstrated that focus may play a role in the determination of the surface characteristics of the acoustic signal. Prosodic boundaries delimiting the focused object may be enhanced by pause and/or glottalstop insertion.

7.1 Syntactic focus marking

Introduction: To gain insights into the categoriality of the use of the ex-situ construction, see chapter 1 section 1.4.2 for further details, with reference to narrow informational focus and narrow corrective focus, Genzel & Kügler (2010) conducted a situation description task inspired by the 'Focus Cards' of the Questionaire on Information Structure (Skopeteas et al., 2006). According to scholars, who consider the use of the ex-situ strategy as an expression of a separate linguistic category (focus type) as e.g. Drubig & Schaffar (2001), Marfo & Bodomo (2005) & Ermisch (2006), corrective focus is expected to be consistently marked syntactically. Under more liberal accounts assuming one category of focus and attributing realizational differences to the paralinguistic notion of emphasis (e.g. Dolphyne, 1988; Saah, 1988; Hartmann, 2008) or degrees of strength (Féry, 2012), more variation is expected.

Material & Procedure: Pictures[52] illustrating situations, displayed in figure 87 were presented to the participants of the 'Ghana' group on paper.

figure 87: Pictures for the description of a single situation task. Target sentence left picture: *Agyeman boaa Ado a-nɔpa yi*. 'Agyeman helped Ado this morning.'. Target sentence right picture: *Anum tɔɔ amango anɔpa yi*. 'Anum bought mango(s) this morning.'.

They were asked to answer the following pre-recorded questions in a natural way, using full sentences. The participants were informed that the whole situation happened *this morning* and were instructed to use the temporal information when answering the questions. The first question, represented in (106)a., was used for both pictures to make sure that the participants understood the situation displayed.

(106)	a.	Deɛbɛn	na	wo-hunu	wɔ	saa	m-foni	yi
		what	PART	PRO-see.HAB	be_at_a_place.HAB	this	NOM-picture	this
		mu.						
		in						
		'What do you see on this picture?'						

	b.	Hwan	na	Agyeman	boa-a	a-nɔpa	yi?	(51)a.
		who	PART	proper name	help-PST	NOM-morning	this	
		'Whom did Agyeman help this morning?'						

	c.	Agyeman	boa-a	Anum	a-nɔpa	yi?	(51)c.
		proper name	help- PST	proper name	NOM-morning	this	
		Did Agyeman help Anum this morning?'					

(106)b. displays a wh-question asking for the object of the left picture. It serves to elicit narrow informational focus. The third question (106)c., which is a Yes – No question, places a narrow corrective focus on the object of the left picture by contrasting it with another name in the question. Corresponding questions for the right picture were also pre-recorded; see

[52] I would like to thank Steffen Schuster for the artistic interpretation of the test sentences.

Genzel & Kügler (2010) for details. Altogether, forty-four answers were analyzed, twenty-two answers to a wh-question and twenty-two answers to a Yes – No question.

Results: In the case of narrow informational focus, all twenty-two answers were realized in-situ. Most of the participants produced the structures displayed in (107)a. to answer the wh-question asking for the object in the left picture, and the structures displayed in (107)b. to answer the wh-question asking for the object in the right picture.

(107)	a.	Agyeman	boa-a	Ado	a-nɔpa	yi.
		proper name	help- PST	proper name	NOM-morning	this
		‘Agyeman helped Ado this morning.’				
	b.	Anum	tɔ-ɔ	a-mango	a-nɔpa	yi.
		proper name	buy-PST	NOM-mango	NOM-morning	this
		‘Anum bought mangos this morning.’				

There was remarkably little variation in the data. One single word utterance, (‘Ado.’; speaker 7), repeating only the questioned object (ellipsis) and deleting the previously mentioned (given) material, and one case of pronominalization of the subject produced by speaker 4; see (108). Two other constructions, produced by speakers 7 and 10, showed no sign of fronting and/or morphological focus marking. Speaker 7 uses a topic construction; see e.g. Saah (1994), Ermisch (2006), Amfo (2010b) for an analysis of *deɛ* as topic marker. Doubling of *deɛ* is common if alternatives are available in the discourse (Duah, p.c.) as illustrated in (109). Note that Boadi (1974) analyzes *deɛ* as a focus marker. In the sample of speaker 10, the questioned object appears in the left periphery of the sentence, without *na,* followed by a locative construction, as displayed in (110).

(108)	ɔ-tɔ-ɔ	a-mango.					
	PRO-buy-PST	NOM-mango					
	‘He bought mango.’						
(109)	Deɛ	me-hunu	yi	deɛ	ɛ-n(o)	ɛ-yɛ	a-mango.
	TM	PRO-see.HAB	this	TM	PRO-DET	PRO-be.HAB	NOM-mango
	‘As for what I see, it is a mango.’						
(110)	A-mango	pii	nso	gu	pono	no	so.
	NOM-mango	many	also	on	table	DET	top
	‘Many mangos on top of the table too.’						

Turning to the results for the context eliciting corrective focus, nineteen out of twenty-two answers were realized in-situ; see (107) for illustration. All in-situ instances were preceded by *daabi* ("no"). Some variation occurred. Speakers 4 and 7 used pronominalization of the subject, as illustrated in (111) with an example from speaker 4.

(111) Daabi! ɔ-boa-a Ado.
no PRO-help-PST proper name
'No! He helped Ado.'

A combined usage of pronominalization and negation was also present. Two speakers (4 and 6) exhibited this structure as exemplified in (112); produced by speaker 6.

(112) Mepawokyɛw! Daabi! ɔ-m-boa-a Anum.
please no PRO-NEG-help-PST proper name
'Please! No! He did not help Anum.'

Three instances, produced by speakers 2, 7 and 11, were realized as ex-situ construction involving the "focus marker" *na*. (113) gives an example from speaker 2; see chapter 1 section 1.4.2 (32) and figure 14 for acoustic data of speaker 11.

(113) Daabi! Ado$_i$ na Agyeman boa-a no$_i$ a-nɔpa yi.
no proper name PART proper name help-PST PRO NOM-morning this
'No! It was Ado who Agyeman helped this morning.'

In the sample of speaker 7, *ɛ-yɛ* is preceding the focus-marked element, as illustrated in (114).

(114) Daabi! Daabi! ɛ-yɛ Ado$_i$ na ɔ-boa no$_i$.
no no PRO-be.HAB proper name PART PRO-help.HAB PRO
'No! No! It is Ado whom he helps.'

Summary: The data has shown that answers to a wh-question were frequently realized in-situ. This finding is in line with Ermisch's (2006) observation that narrow informational focus is not (morpho) syntactically marked in Akan. Furthermore, we have seen that narrow corrective focus can be realized ex-situ, supporting observations by Saah (1988), Marfo & Bodomo (2005) and Ermisch (2006). However, speakers coping with the situation description task, more frequently choose to leave the focus-marked element in-situ.

Discussion: The situation description task showed that the in-situ strategy is clearly the preferred one with narrow informational focus and also with narrow corrective focus. Thus, focus fronting is not obligatory with narrow corrective focus in Akan; see Hartmann & Zimmermann (2007) for comparable results on Hausa. The results are problematic for accounts which assume a one to one correspondence between focus "type" and choice of the grammatical form (e.g. Drubig & Schaffar, 2001; Marfo & Bodomo, 2005; Ermisch, 2006) and may be taken as evidence that corrective focus (contrast) does not have the status of a grammatical category (Hartmann, 2008). The ex-situ strategy is an option available in the grammar of Akan which allows the speaker to express extra emphasis (Dolphyne, 1988; Saah, 1988) on the focused element (Hartmann, 2008). The fact that in-situ realizations were much more frequent with narrow informational and narrow corrective focus may be explained by the factor "…*hearer expectation* or *discourse expectability* of the focused content in a given discourse situation." (Zimmermann, 2007:147). Féry (2012:43) observes that if the focus of a sentence is deducible from the context, it does not have to be marked by grammatical devices: "If, in a conversation, one of the protagonists corrects a preceding proposition, it is immediately clear which part of the sentence is corrected. The same holds when the focus is the answer to a question."; see also Hartmann (2008) for a comparable reasoning. It thus seems to be the case that the three participants, who used the ex-situ construction, decided to signal that they perceived the question as an unexpected discourse move (e.g. Zimmermann, 2007) by giving more emphasis to the focus-marked element when correcting a preceding proposition, while all others did not.

As a final remark, I would like to add that there are several issues, related to the ex-situ construction, which are beyond the scope of the thesis but interesting in their own right. The first issue is concerned with subject/object asymmetries (e.g. Zerbian, 2007) in focus marking, also called subject und non-subject asymmetries (e.g. Fielder & Schwarz, 2005). It has been observed in a number of languages that focused non-subjects (NSF) frequently remain unmarked (syntactically, morphologically and/or prosodically), whereas focused subjects (SF) must be marked (e.g. Fiedler, Hartmann, Reineke, Schwarz & Zimmermann, 2010). Akan, however, shows only minimal, if any, differences between SF and NSF. As pointed out in chapter 1 section 1.4.2, focused subjects as well as non-subjects can be fronted in Akan (Boadi, 1974; Kobele & Torrence, 2006) and are obligatorily followed by the particle *na* (Marfo & Bodomo, 2005). Moreover, focused subjects and non-subjects can remain in-situ under narrow informational focus (Ermisch, 2006; Genzel & Kügler, 2010) and also under narrow corrective focus. This has been shown in the current section; see also Genzel & Kügler (2010). Fiedler & Schwarz (2005) observe that Akan exhibits a subject und non-subject asymmetry and present the examples reproduced in (115) and (116) to illustrate it. The sentence involving a fronted focused subject is reproduced in (115)b. The canonical sentence is displayed in (115)a.

(115) a. À-brèwá nó dì-ì à-dùá nó.
NOM-old.woman DEF eat-PST NOM-beans DEF
'The old woman ate the beans.'

b. ɛ̀-yɛ̀ [à-brèwá $nó_i]_F$ nà $ɛ̀_i$-dí-ì à-dùá
PRO-be.HAB NOM-old.woman DEF PART PRO-eat-PST NOM-beans
nó.
DEF
'It is the old woman who ate the beans.'

(Fiedler & Schwarz, 2005:116; F-marking and co-indexing was added to the original. Tones and glosses were slightly changed.)

The sentence illustrating NSF is reproduced in (116)b. The canonical counterpart was added supplementary by me; see (116)a.

(116) a. Pàpá nó súà nè krámán.
father DEF carry.HAB PRO dog
'The father carries his dog.'

b. [Nè $krámán_i]_F$ nà pàpá nó súà $nó_i$.
PRO dog PART father DEF carry.HAB PRO
'It is his dog that the father carries.'

(Fiedler & Schwarz, 2005:116; F-marking and co-indexing was added to the original. Tones, glosses and translation were slightly changed.)

Fiedler & Schwarz (2005:114) note that the subject und non-subject asymmetry in Akan is "…less obvious at first sight." than in e.g. Dagbani. They observe that "In SF, an expletive subject pronoun for 3rd person referents *(ɛ̀-)* is characteristically used, although it might be replaced by the normal pronominal form." (Fiedler & Schwarz, 2005:115). Fiedler & Schwarz do not specify what the "normal pronominal form" would be. Dolphyne (1988:87) list *ɔ-no* as 3rd person singular pronoun and *ɛ-no* as impersonal form. If the resumptive pronoun refers to the subject, only the prefix is affixed to the verb (Cleland, Gyang, Imbeah & Imbeah, 2009:8), as in (115)b. With regard to NSF, Fiedler & Schwarz (2005:114) observe that "…there is no general need for an object pronoun that is coreferent with the constituent in focus, but rather the selection of the object pronoun underlies semantic criteria.". They refer to the "semantic criterion" of animacy; see also chapter 1 section 1.4.2. Saah (1988:24) claims that "…all NP's can be pronominalized in object position, the substituted pronoun is overt if the pro-nominalized NP is animate but covert if it is inanimate."; see (116)b. for pronominalization of an animate object NP. However, this dependency seems not to be absolute, as shown in chapter 1 section 1.4.2 (32). To me, it seems that there is no subject und non-subject asymmetry in Akan, at least not in the original sense of the term; see Fielder et al. (2010). The use of the resumptive pronoun in relation to animacy needs further investigation.

The second issue relates to the identity of the ex-situ construction. Drubig (2003) elaborates on the structural similarities between focus fronting and wh-word fronting in Akan; see also Saah (1988). He establishes a relationship to relative clauses and analyzes the ex-situ construction as a "reduced cleft"; see also Boadi (1974). Kobele & Torrence (2006) suggest that the focus construction is a cleft construction. Fiedler & Schwarz (2007:267) note that "…these focus constructions should be analyzed as extra-clausal structures in which the non-focal information is presented as a narrative clause.".

The third issue is concerned with tonal/prosodic aspects that come along with the ex-situ strategy, already touched upon in chapter 1 section 1.4.2. Boadi (1974:19) observes that "If the sentence is *na*-focussed, the whole predicate is raised one step higher tonally, if it was not high originally. At the same time, the focussed constituent is raised high if it was low.". He further notes that "It would probably not be accurate to describe the change as comprising from syllabic lows to highs."; see also Ameka (2010). The classic example involves fronting of a pronoun as illustrated in (117). The canonical sentence is displayed in (117)a. and its ex-situ counterpart in (117)b. Similar tonal processes on the verb have been observed with relative constructions (Fiedler & Schwarz, 2005 and references therein).

(117)	a.	Mè-bà-à			há.
		PRO-come-PST			here
		I came here.			
	b.	Mé	nà	mé-bá-á	há.
		PRO	PART	PRO-come-PST	here
		'It is I who came here.'			

(Boadi, 1974:19)

The figure 88 presents acoustic data of the material presented in (117). The sentence displayed in (117)a. is visualized in the left track and the sentence in (117)b. in the right track. In line with Boadi's observation, the fronted pronoun in the right track is clearly raised from L to H. It should be noted that this tonal shift does not appear if the fronted element is a proper noun (name); see chapter 1 section 1.4.2 and section 7.2 of this chapter for evidence. One tentative explanation may rest on the assumption that pronouns are toneless (Abakah, 2002, 2005).

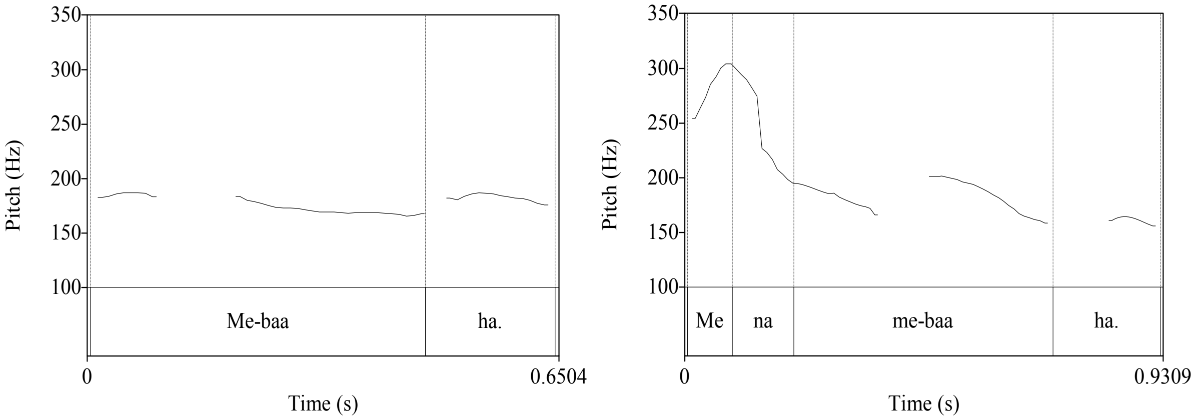


figure 88: Left pitch track: course of F0 for the sentence *Me baa ha*. ‘I came here.’; Right pitch track: course of F0 of the sentence *Me na me baa ha*. ‘It is I who came here’.; female speaker.

Further, we can observe that the verb is raised in the right track compared to the left one. The tense affix is realized with a L tone, contrary to Boadi (1974). Additionaly, the pronoun immediately preceding the verb is clearly realized low, contrary to the expectations. The reasons for the tonal modifications are not clear and need further investigation.

In this section, it has been shown that focused objects are frequently realized in their base position, i.e. without any additional syntactic and morphological marking. It thus can be concluded that the in-situ strategy is the preferred one for focused objects. Furthermore, the results suggest that the ex-situ strategy comes along with a “stronger” interpretation, because it is only used with corrective focus. The frequent use of in-situ focus raises the question of the possibility of prosodic marking of object focus. Prosodic effects of focused objects in Akan are investigated in the following section.

7.2 Prosodic focus marking

Introduction: To gain insights into the prosodic marking of focus in Akan, Kügler & Genzel (2012) conducted a production experiment [53], involving the factors focus type (wide informational focus, narrow informational focus and narrow corrective focus), tone on the object (L and H) and syntactic construction (in-situ and ex-situ). This design enabled us to examine the tonal behavior (scaling) of non-agent non-verbal target words under narrow focus and to see whether there is a prosodic effect of focus on the focused element. The comparison of narrow informational and narrow corrective revealed insides into categorical or gradient uses of grammatical means, already tackled in section 7.1.

Material & Measurements: To investigate the prosodic marking of focus on in-situ and ex-situ objects in Akan, the material presented in table 29 was analyzed. It should be noted that the baseline for the focused object in the ex-situ construction is a homophonous subject in canonical position. TW in table 29 refers to the target word (object). The context questions eliciting narrow informational and narrow corrective focus on the L toned object are displayed in (106)b. and (106)c. in section 7.1. The corresponding answer with the focused object carrying L tones is illustrated in (118). The complete list of materials is introduced in chapter 2, section 2.2, block E.

Condition	in-situ L tone TW	in-situ H tone TW	ex-situ L tone TW	ex-situ H tone TW
wide informational focus (baseline)	(50)a.	(50)b.	(50)c.	(50)d.
narrow focus	(52)a.	(52)b.	(52)c.	(52)d.

table 29: Material used to test the prosodic marking of focus.

(118) Àgyèmàǹ bóà-à [Àdò]F à-nɔ̀pá yí. (52)a.
proper name help-PST proper name NOM-morning this
'Agyeman helped Ado this morning.'

Planned comparison: The procedure applied is illustrated for the L toned target word in (119). The same method is applied for the sentences which contain the target word *àmáǹgò*; only the H tone is considered in the result presentation. The height of both L tones on the target word (object) under narrow informational and narrow corrective focus will be compared to the height of both tones under wide informational focus which serves as a baseline.

[53] The data has already been published as Kügler & Genzel (2012).

(119) baseline (in-situ) | narrow informational focus | narrow corrective focus

baseline (in-situ)	narrow informational focus	narrow corrective focus
L L	L L	L L
| |	| |	| |
…Ado…	…Ado…	…Ado…

An additional relational measurement will be presented. Therefore, the height of the initial H tone (H1) was measured. For the sentence containing the target word *amango*, H1 was extracted at the temporal mid of the vowel /u/ of the subject *Ànúm̀*. For the sentence containing the target word *Ado*, H1 was extracted at the temporal mid of the vowel /o/ of the verb *bóà-à*. To gain insights into the relation of H1 and the tone on the target word in sentence perspective, the difference (Δ) between H1 and the H tone on the target word *àmáǹgò* was calulated. For the sentences containing the L toned target word the difference between H1 and the second L tone of the target word *Àdò* was calculated. Δ of wide informational focus renditions will be compared to Δ in renditions which exhibit narrow corrective focus on the target word.

Results: On the focused in-situ object, a gradual decrease in F0 height can be observed with increasing strength (Féry, 2012) of the focus (wide informational focus < narrow informational focus < narrow corrective focus) on the H tone on the antepenultimate syllable of the target word *amango*; see figure 89. The solid line refers to the wide focus condition. It serves as the baseline for comparison. The target word under narrow informational focus is represented by the dotted line. The dashed line corresponds to the renditions under narrow corrective focus.

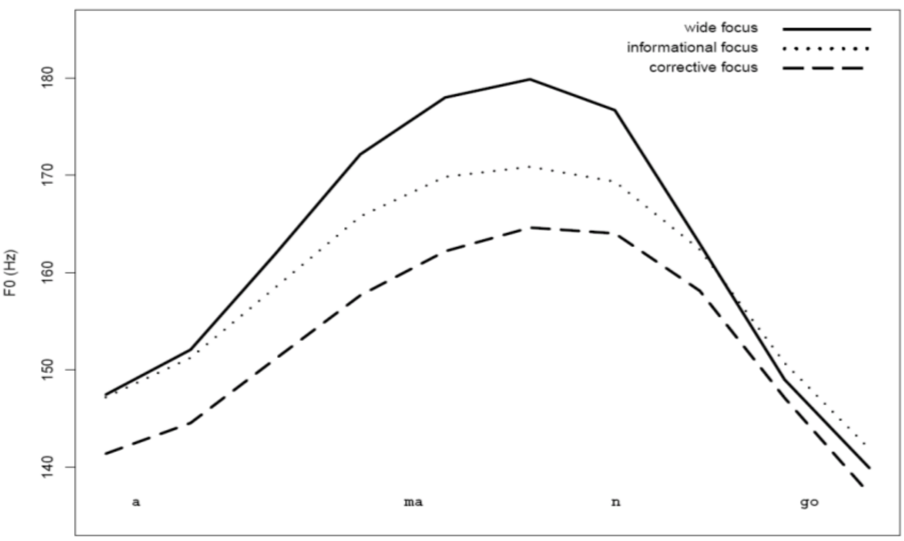


figure 89: Time-normalized course of F0, aggregated over speakers, for *amango* 'mango' uttered under wide informational (solid line), narrow informational (dotted line) and narrow corrective focus (dashed line), taken from the sentence *Anum tɔɔ amango anɔpa yi*. 'Anum bought a mango this morning.'; n=11 (Kügler & Genzel, 2012:344).

The mean F0, aggregated over speakers, for the different focus conditions, amounts to 182 Hz (47) for the wide informational focus, to 172 Hz (41) for the narrow informational focus and to 166 Hz (41) for narrow corrective focus.

The repeated measures ANOVA (including syntactic construction; see below) revealed a significant main effect of focus ($F(2, 10) = 17.4$, $p < 0.001$) for the H tone. Post-hoc *t*-tests showed no significant effect for narrow informational focus compared to the wide focus baseline, but a significant ($t(1, 10) = 3.0$, $p < 0.05$) lowering of narrow corrective focus compared to the wide informational focus baseline. The lowering of the H tone under corrective focus amounts to 1.5 st on average.

The figure 90 shows the course of F0 on the target word *Ado*. Again, the solid line displays the wide informational focus baseline. As for the target word carrying the H tone, the F0, corresponding to the L tone, decreases gradually with the strength of the focus. The second lexical L tone of the target word *Ado* is realized lower than the first one. As pointed out in chapter 1 section 1.4.2, the initial syllable is a toneless nominal prefix which receives its tonal specification via spreading from the predecing syllable. Only the second L tone will be considered.

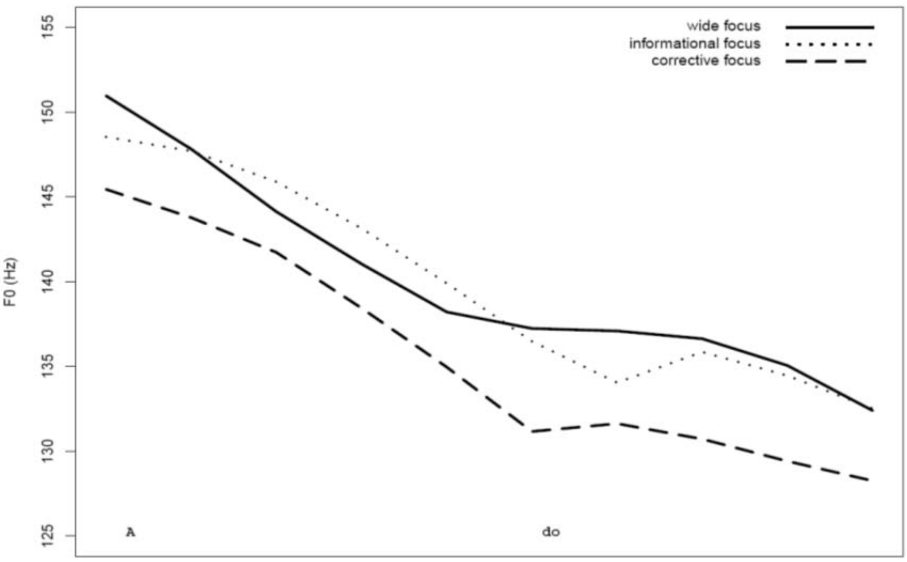


figure 90: Time-normalized course of F0, aggregated over speakers, for the target word *Ado* uttered under wide informational (solid line), narrow informational (dotted line) and narrow corrective focus (dashed line), taken from the sentence *Agyeman boaa Ado anɔpa yi.* 'Agyeman helped Ado this morning.'; n=11 (Kügler & Genzel, 2012:346).

The mean F0, aggregated over speakers, for the different focus conditions, amounts to 137 Hz (36) for the wide informational focus, to 137 Hz (35) for the narrow informational focus and to 132 Hz (34) for narrow corrective focus.

The factor focus approached significance in the repeated measures ANOVA ($F(2, 9) = 2.9$, $p = 0.077$). Post-hoc *t*-tests showed a significant ($t(1, 9) = 4.2$, $p < 0.01$) lowering of F0 for the second L tone under narrow corrective focus compared to the wide informational focus baseline. The lowering of the second L tone amounts to 1.0 st on average.

Turning to the results of focus on the dislocated object, figure 91 shows the time-normalized course of F0, aggregated over speakers, for the target word carrying an H tone. The fronted renditions are represented by the grey lines and the renditions exhibiting the target word in canonical position are repeated (black lines). The solid lines refer to the wide informational focus baseline. Note that the baseline for the ex-situ renditions is a homophoneous subject, see chapter 2, section 2, block E, (50)c. and (50)d. The dotted lines correspond to narrow informational focus and dashed lines to narrow corrective focus conditions. The mean F0 for the H tone of the target word *amango*, aggregated over speakers, amounts to 248 Hz (72) for the wide informational focus, to 248 Hz (76) for the narrow informational focus and to 227 Hz (68) for narrow corrective focus.

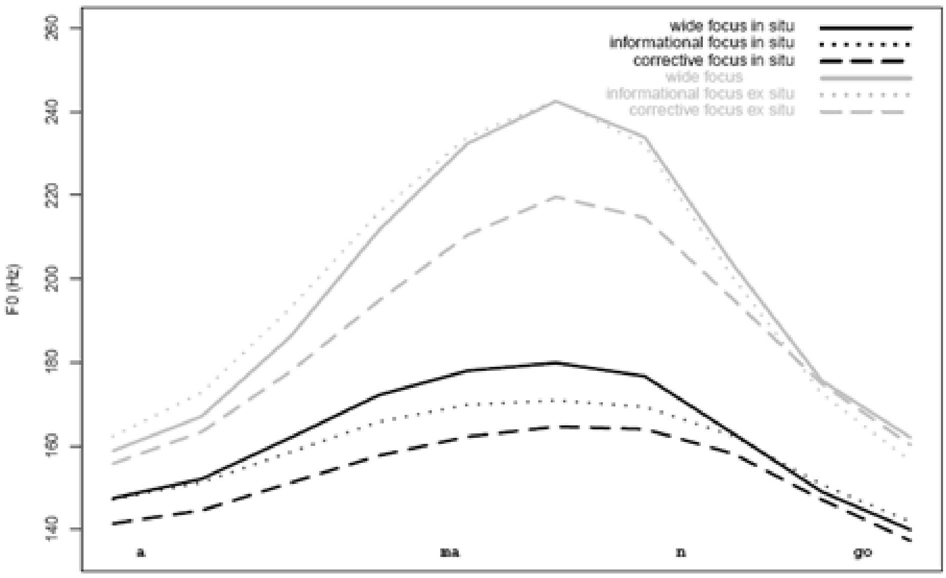


figure 91: Time-normalized course of F0, aggregated over speakers, for *amango* 'mango' uttered in-situ under wide (solid black line), narrow informational (dotted black line) and narrow corrective focus (dashed black line), taken from the sentence *Anum tɔɔ amango anɔpa yi*. 'Anum bought a mango this morning.' and uttered ex-situ under wide (solid grey line), narrow informational (dotted grey line) and narrow corrective focus (dashed grey line), taken from the sentences *Amango ate firi dua no so anɔpa yi*. 'A mango has fallen down the tree this morning.' and *Amango na Anum tɔɔ anɔpa yi*. 'It is a mango that Anum bought this morning.'; n=11 (Kügler & Genzel, 2012:347).

The sentence initial renditions of the target word are generally overall higher in F0 than in the medial (in-situ) position. This is expected since Akan has been shown to exhibit a considerable amount of declination; see chapter 4 section 4.3.1. The repeated measures

ANOVA exposed a significant (F(1, 10) = 41.0, p < 0.001) main effect of Syntactic Construction. The post-hoc *t*-tests showed no significant result for the comparison of the H tones between the baseline and narrow informational focus condition, but for narrow corrective focus ($t(1, 10) = 5.5$, $p < 0.001$). The lowering in the ex-situ construction amounts to 1.8 st, on average. The mean F0 for the second L tone of the target word *Ado*, aggregated over speakers, amounts to 162 Hz (40) for the wide informational focus, to 158 Hz (40) for the narrow informational focus and to 155 Hz (37) for narrow corrective focus. Hence, as for the H tone of the target word *amango*, the lowering is also present, though to a lesser degree (0.8 st). None of the comparisons yielded a significant result.

In the following, it will be investigated whether the observed lowering effect may be due to a global lowering caused by the presence of the negation (*daabi*) in the beginning of the sentences containing a corrective focus on the object; see Greif (2012) for Mandarin Chinese data showing such an effect. During the data elicitation, speakers were instructed to pause after the negation particle. This instruction was given to minimize the influence of it on the following target sentence. All speakers acted according to this instruction. In Kügler & Genzel (2012), we presented two empirical arguments in favor of a local lowering effect, affecting only the element under corrective focus. First, the F0 of the same target word (*amango*) occurring immediately before a narrowly corrected element (pre-focal) will be compared to the F0 of the target word under wide informational focus. The context question is illustrated in (120)a. and the corresponding answer in (120)b. The answer exhibits a sentence initial negation.

(120) a. Ànúm̀ tɔ̀-ɔ̀ à-máǹgò ɛ̀-nórà?
proper name buy.PST NOM-mango NOM-yesterday
'Did Anum buy mango yesterday?'

b. Daabi! Ànúm̀ tɔ̀-ɔ̀ à-máǹgò [à-nɔ̀pá yí]F.
No proper name buy.PST NOM-mango NOM-morning this
'No, Anum bought mango this morning.'

The results are displayed in figure 92; the solid black line refers to the baseline (wide information focus) and the black dotted line to the target word in pre-focal position. The comparison of the F0 on the target word *amango* between the two conditions shows a slightly lower F0 if a negation particle precedes the sentence. However, the statistical comparison between the H tone of the target word in pre-focal position and in the neutral condition was not significant.

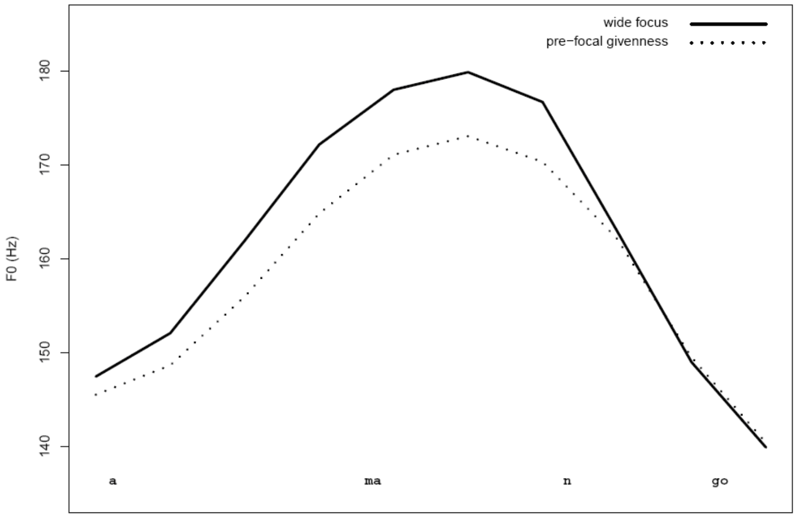


figure 92: Time-normalized course of F0, aggregated over speakers, for the word *amango* 'mango' uttered in wide focus (black solid line) and with the word occurring pre-focally (black dotted line); taken from the sentence *Anum tɔɔ amango anɔpa yi*. 'Anum bought a mango this morning.'; n=11 (Kügler & Genzel, 2012:349).

Second, we presented data from another experiment involving the factors wide informational focus, elicited without context, and narrow corrective focus; see Kügler & Genzel (2012) for details. The context eliciting narrow corrective focus on the object is illustrated in (121)a. and the target sentence with focus marking in (121)b. It is important to note that no negation particle was present in (121)b.

(121)	a.	Àfúà	hùnù-ù	ɛ̀-bóɔ̀	à-nɔ̀pá	yí?
		proper name	see-PST	NOM-stone	NOM-morning	this
		'Did Afua see a stone this morning?'				
	b.	Àfúà	hùnù-ù	[wɔ́má]F	á-nɔ̀pá	yí.
		proper name	see-PST	pestle	NOM-morning	this
		'Afua saw a pestle this morning.'				

The figure 93 provides a sentence perspective of the sentence in (121)b. and the wide informational focus baseline. The mean F0 for the H toned target word *wɔma* is slightly lower in the case of narrow corrective focus. Furthermore, the lowering appears only on the target word itself. The previous H tone (H1), associated with the subject of the utterance, is scaled exactly at the same height. In Kügler & Genzel (2012), we did not provide any relational measure between H1 and the H tone on the target word. However, it can be deduced from the figure that the difference (H1 – H tone on the target word) is slightly greater for the corrective focus than for wide informational focus.

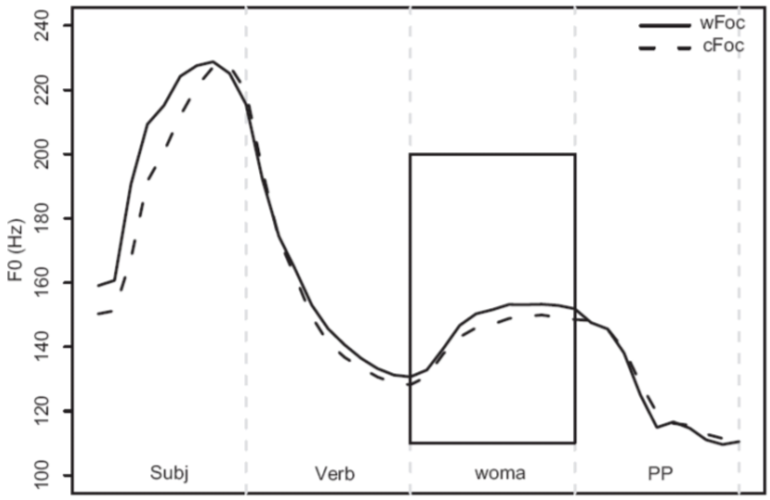


figure 93: Time-normalized course of F0, aggregated over speakers, for the sentence *Afua hunuu wɔma anɔpa yi*. 'Afua saw a pestle this morning.' under wide focus (black solid line) and under corrective focus (black dashed line); n=5 (Kügler & Genzel, 2012:352).

Turning to the relational measurement of the in-situ renditions, uttered under wide informational and narrow corrective focus, containing the target word *amango* and *Ado*, the results are displayed in table 30. In the sentences containing the H toned target word, H1 is realized clearly lower in the narrow corrective focus condition. In this case, H1 is immediately following the negation particle. The difference between H1 and the H tone on the target word is smaller in the narrow corrective focus condition than in the wide information focus baseline, which means that the F0 on the H tone of the target word is raised under corrective focus. In the sentences containing the L toned target word, H1, which was measured on the verb, is only slightly lower in the narrow corrective focus condition. The difference is slightly higher for narrow corrective focus compared to the baseline, which points to the fact that the L tone on the target word is lowered in this condition.

Condition	Tone TW	F0 (Hz) H1	F0 (Hz) TW	Δ (Hz)
wide inf. focus	H	266.59 (71)	181.94 (47)	84.66 (30)
narrow corr. focus	H	222.44 (60)	165.79 (41)	56.65 (22)
wide inf. focus	L	198.96 (55)	137.29 (36)	61.68 (24)
narrow corr. focus	L	197.99 (58)	131.52 (35)	66.47 (30)

table 30: Mean F0 values in Hz for the initial H tone (H1), for the target word (TW) and difference (Δ), aggregated over speakers, split by focus conditions; n=11.

A linear mixed effects model was calculated on the difference. Tone (L/H) and focus condition (wide informational focus/narrow corrective focus) were treated as fixed factors. Speakers were considered as random factor. The factor tone does not show a significant effect. This information is not relevant here. Focus condition approached significance (t = -

2.208, *p*MCMC = 0.0719, 44 observations). However, the effect is not independent. The interaction of tone and condition was significant ($t = 3.120$, *p*MCMC < 0.05); the difference is systematically smaller for narrow corrective focus than for wide informational focus in the sentences containing the H toned target word, whereas there is no difference for the sentences containing the L toned target word.

Summary: The data has shown that the absolute F0 height on an in-situ target word (object) under narrow informational focus does not differ systematically from the absolute F0 height under wide informational focus. The absolute F0 on an in-situ target word under narrow corrective focus, however, differed from its neutral counterpart in being realized systematically lower; see e.g. figure 89. An additional measurement, relating the F0 height on the target word to that of the first H tone (H1) in the same utterance, revealed that H1 on the subject is already systematically lower (~ 40 Hz) when it is preceded by a negation particle, as in the case of narrow corrective focus. The difference between H1 and the H tone on the target word has been found to be generally smaller for narrow corrective focus than for wide informational focus. No systematic effect was observed for the sentences containing the L toned target word; see table 30.

The data at hand did not provide evidence in support of Boadi's (1974) claim that the focus-marked element (proper name) preceding the focus-marker *na* (ex-situ) is raised to H if it was L. In accordance to the results for the in-situ realizations, the comparison of the absolute F0 height of a fronted object under narrow informational focus to a homophone initial subject under wide informational focus showed no systematic difference but a systematic lowering under narrow corrective focus when the target word contains an H tone; see figure 91 for illustration.

Discussion: The search for prosodic effects on the focused element, in the laboratory speech data, has revealed that narrow informational focus does not differ systematically from wide informational focus in terms of absolute tonal height. At first glance, it seems that Akan belongs to the group of non-marking languages (Büring, 2010) together with e.g. the tone languages Hausa (Hartmann & Zimmermann, 2007), Northern Sotho (Zerbian, 2006, 2007), Yucatec Maya (Kügler & Skopeteas, 2006, 2007; Kügler, Skopeteas, & Verhoeven, 2007; Gussenhoven & Teeuw, 2008) and Navajo (McDonough, 2002) and the intonation language Wolof (Rialland & Robert, 2001); see Zerbian, Genzel, Kügler (2010) for an overview of African tone languages. One possible explanation is to assume that in Akan, FocProm must be ranked below other mapping constraints; see e.g. Zerbian (2006) and Downing (2009). The other is to say that focus in answers to wh-questions does not have to be especially marked because it is a prototypical discourse move (Hartmann, 2008). However, Féry (2012) argues that the universal reflex of focus is alignment, which may go together with prominence in some languages, but that alignment and prominence do not necessarily co-occur. In Kügler & Genzel (2012), we also evaluated the duration of the target word and did not find any consistent lengthening effect or pauses due to a possible insertion of a prosodic boundary for narrow informational focus and also not for narrow corrective focus which is assumed to be stronger and "…more prone to be aligned than a weaker one, like an informational focus." (Féry, 2012:1). However, as Féry points out, phrasing related effects of focus may be absent in laboratory speech and are more likely to occur in spontaneous or semi-spontaneous speech. Before evaluating whether we find evidence for phrasing related effects of focus in the semi-spontaneous data from the situation description task presented in section 7.1, I would like to comment on the lowering effect which occurred in the laboratory speech data with narrow corrective focus.

In Kügler & Genzel (2012), we came to the conclusion that narrow corrective focus is accompanied by a lowering which we interpreted as a local effect. We argued that "In the case of corrective focus,…, an additional pragmatic prominence comes into play which speakers may want to express even prosodically. This additional prominence is correlated with a stronger communicative goal to emphasize a certain part of an utterance, and speakers of Akan draw attention to that kind of information by means of pitch register lowering." (Kügler & Genzel, 2012:353). However, closer examination of the scaling relations inside the sentences, presented here, revealed that the lowering effect detected on the narrowly contrasted element is indeed due to a register lowering following the negation. This may also apply to the lowering effect on fronted elements[54]. For elements bearing an H tone, the speakers rather show a raising of F0 on the object under narrow corrective focus. It seems thus that Akan speakers principally have the option to express extra emphasis in the phonetics which may be related to "…the urge to express unexpected discourse moves." (Hartmann, 2008:407). Raising of F0 to signal emphasis is expected. Gussenhoven (2002, 2004) argues that intonational meaning manifests itself in universal and language-specific aspects, which are coded in the intonational grammar of each language, while the universal part is expressed phonetically across all languages. The universal aspects of the interpretation of pitch variation, also known as biological codes (Ohala, 1994), predict "…that the speaker is being forceful because he believes the contents of his message are important." (Gussenhoven, 2002:49). The speaker thus puts more effort in the speech production. More effort is

[54] It should be noted that this explanation may not hold for all the data we presented in the paper.

phonetically detectable as greater articulatory precision and/or pitch register expansion; see also chapter 1 section 1.6.3.

Turning to the evaluation of the semi-spontaneous data from the perspective of prosodic phrasing (Féry, 2012), the assumed default prosodic structure under wide informational focus of the material used, is illustrated in (122) for the sentence containing the L toned target word (*Ado*); the same structures apply to the sentence containing the target word *amango*. In (122)a., the subject is phrased into its own pP. Object and temporal adverb are mapped onto a separate pP, which is embedded in the pP mapping the VP; see chapter 1 section 1.4 for further details on phrasing and phonetic cues in Akan. It is also possible that the temporal adverb is mapped onto its own pP, everything else being equal, as illustrated in (122)b.

(122) a. ((Agyeman)φ (boa-a (Ado a-nɔpa yi)φ)φ)ι
b. ((Agyeman)φ (boa-a (Ado)φ (a-nɔpa yi)φ)φ)ι

In nearly all wide informational focus renditions (lab speech data), the final vowel of the verb is co-articulated with the initial vowel of the object and the final vowel of the object is co-articulated with the initial vowel of the temporal adverb, as illustrated in figure 94.

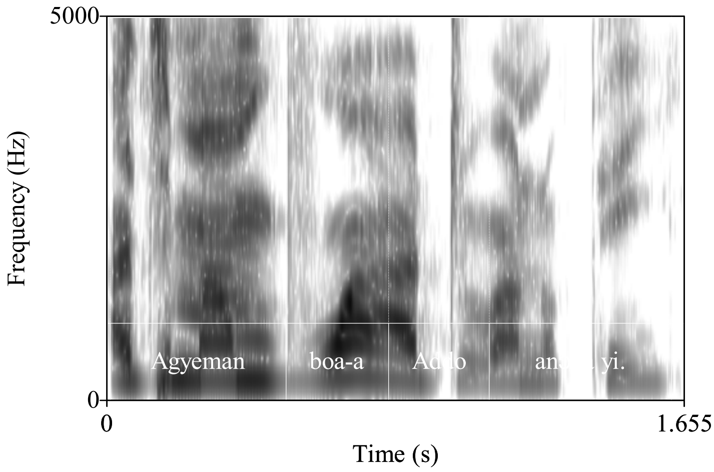


figure 94: Spectrogram of a SVOAdv sentence uttered under wide informational focus; female speaker (11).

However, it seems principally possible to have a pP boundary before the temporal adverb. One speaker produced a pause (92 ms) after the object (*amango*) under wide informational focus. The variant of the prosodic structure is illustrated in (122)b. The F0 pattern of the prevalent realization, reflecting the prosodic structure exemplified in (122)a. under wide informational focus, is illustrated in figure 95 with data from a female speaker. Pitch is continuous and no pauses appear.

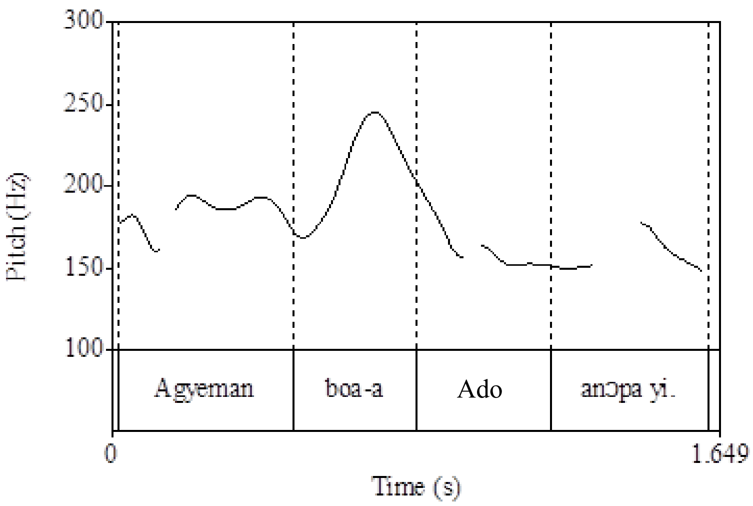


figure 95: SVOAdv sentence uttered under wide informational focus; female speaker (11).

Turning to the prosodic characteristics of focus on objects (semi-spontaneous data), it can be observed that answers to wh-questions contain a pause before (72 ms) and after the focused element (154 ms), as illustrated in figure 96 with data from the female speaker.

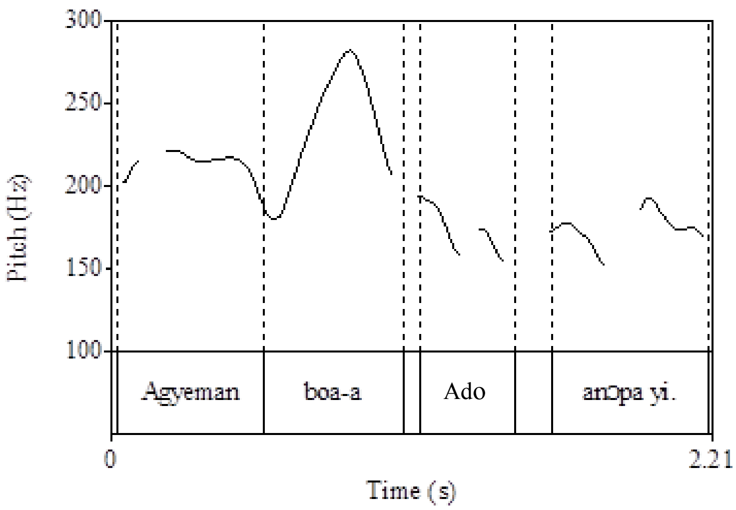


figure 96: SVO$_F$Adv sentence with narrow informational focus on the object; female speaker (11).

After the verb and after the object, the speaker produces a glottal stop; see Dilley, Shattuck-Hufnagel & Ostendorf (1996) for the role of glottalization of word-initial vowels as boundary marker in American English. This is illustrated in figure 97. The waveform of the same sentence is presented with labeled glottal stops.

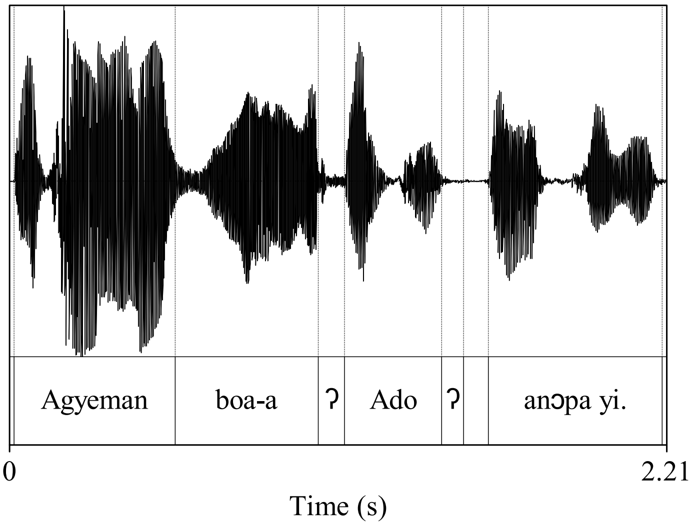


figure 97: Waveform of a SVO$_F$Adv sentence uttered under narrow informational focus; female speaker (11).

The same effect can be observed for narrow corrective focus, as illustrated in figure 98 with data from the female speaker. The pause, preceding the object, is 130 ms long and after the object, 184 ms. Again, the final vowel of the verb is followed by a glottal stop. Furthermore, it is interesting to note that the pitch on the final syllable of the L toned object *Ado* is sharply falling, which goes together with a change in voice quality from modal to glottalized.

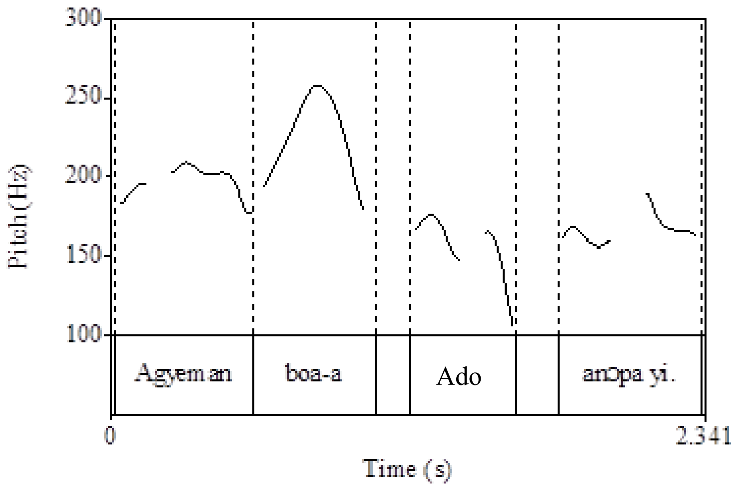


figure 98: SVO$_F$Adv sentence with narrow corrective focus on the object; female speaker (11).

The realization of this particular speaker is not an isolated case. Out of the whole semi-spontaneous data sample, four speakers (1, 3, 5, 8) produced string-identical utterances to the wide informational focus baseline of the laboratory speech data set for both test sentences. From the data of speakers 10 and 11, only the sentences containing the L toned target word, and from speaker 11, the sentence containing the target word *amango* under narrow informational focus, exhibited the target structure. Thus, twenty-one instances were analyzed for occurrences of pauses before and after the object. Altogether, seven pauses occurred before the focused object; four (mean duration 181.5 ms (148)) were produced in answers to a wh-question and three (mean duration 115 ms (52)) in answers to a question eliciting narrow corrective focus. Pauses after the focused element were more frequent; seventeen pauses were found. Nine (mean duration 97.4 ms (41)) were produced in answers to a wh-question and eight (mean duration 149.1 ms (51)) in answers to a question eliciting narrow corrective focus. It is further interesting to note that the length of the final vowel of the object did not differ systematically between the conditions. The mean length for the final vowel of the L toned target under wide informational focus amounts to 94 ms (29), to 102 ms (27) under narrow informational focus and to 90 ms (25) under narrow corrective focus. It thus seems that Akan lacks pre-boundary lengthening. Moreover, a glottal stop frequently occurred after the final vowel of the verb and after the final vowel of the object. The glottal stop has not drawn much attention in the literature on Akan. However, Dolphyne (1988:48f.) notes that the glottal stop is not part of the phoneme inventory, but occurs at the end of a word that is followed by a pause, also after final consonant (/m/, /w/, /r/). "The glottal stop is … a feature of a pause in Akan." (Dolphyne, 1988:49). Furthermore, Dolphyne (1988) observes that it appears at the end of negative sentences like in Dagbani (Hyman, 1989). Moreover, the abrupt lowering of F0 on the final syllable of the object, illustrated in figure 98, is also present in the data of some other speakers but not as pronounced as for speaker 11. We thus have good evidence to conclude that focus in Akan is marked by insertion/enhancement of a pP boundary, adding evidence to Féry's (2012) claim that focus wants to be aligned. If the temporal adverb is not phrased separatly by default, a pP boundary is inserted at the right edge of the focus-marked element. The insertion is phonetically detectable by pauses and/or glottal stop insertion. The default phrasing under wide informational focus is repeated in (123)a. The prosodic structure of an Akan sentence containing a focused object is displayed in (123)b. However, if the focused object is already aligned with a pP boundary at its right edge by default, see (122)a., already existing boundaries are enhanced via pause and/or glottal stop insertion.

(123) a. ((Agyeman)φ (boa-a (Ado a-nɔpa yi)φ)φ)ι
b. ((Agyeman)φ (boa-a (Ado$_F$)φ a-nɔpa yi)φ)ι

This chapter revealed that focus in Akan is marked by prosodic means. However, unlike sentence type which is marked by tonal means, the data at hand has shown that a glottal stop is used to mark a pP boundary accompanying the focused element. From a typological perspective, Akan is not the only language employing non-pitch features to convey post-lexical pragmatic meaning. Laryngeal features such as breathy termination have been reported to signal sentence type in some tone languages; see chapter 6. Hyman & Monaka (2008) present further evidence from the literature showing that a glottal stop can take over the function of intonational tones such as in imperatives in Lahu (Matisoff, 1973), questions in Kaingang (Wiesemann, 1972), and negatives in Dagbani (Hyman, 1989). The results from Akan thus add further evidence to the actual discussion to include non-pitch features, e.g. laryngeal features and morphemes, into the definition of intonation (e.g. Hyman & Monaka, 2008; Zerbian, 2010).

A narrow focus only affects the F0 of an Akan sentence marginally if it is contrastive/corrective. It has been argued that the raising effect of H tones under focus is due to emphasis and not due to focus per se. Although narrow focus in semi-spontanous speech can be marked by boundary insertion/enhancement, which leads to an interruption (pause/glottal stop) of the otherwise continuous speech signal, F0 is not affected. The declination pattern is maintained. This is expected from the analysis of declination as a process applying at the level of the IP presented in chapter 4, section 4.3.2. However, superlow realizations, which resemble final lowering, might occur.

Conclusion

The aim of the thesis was to determine which factors, beside lexical tones, contribute to the F0 contour of sentences and thus to gain a deeper understanding of lexical and post-lexical tones in Akan and their phontic implementation. The amount and the sphere of action of the factors were elaborated. The concrete research questions were outlined in chapter 1. In the following, the questions will be repeated and answers will be provided, based on the data analysis and discussions presented in chapters 3-7.

In chapter 1 section 1.1, it was sketched that Dolphyne (1988, 1994) claims !H to be lexicalized i.e. part of the lexical entry of some morphemes in Akan. Abakah (2000), on the other hand, suggests that !H is not lexicalized. He argues that instances of !H are due to non-automatic downstep which is triggered by the presence of L̥ in the derivation.

- Is !H an independent entity in the phonology of Akan?

In chapter 4 section 4.3.1, the observation that sentences with only H tones do not show declination if M or !H is phonologically contrastive (Hyman, 1975) was taken as starting point to answer the research question. The declination rate (st/sec) of sentences with only H tones was measured, with the result that those instances exhibit a considerable amount of declination. This was taken as indirect evidence in support of Abakah's (2000) claim that !H is not part of the lexical entry of Akan words. Hyman's (1975) observation is further interesting from the standpoint of phonetic universals. The Production code (Gussenhoven, 2002, 2004) predicts that declination is automatic since subglottal air pressure is higher and consequently pitch is higher at the beginning of a breath group/utterance than at the end. However, in some tone languages, speakers seem to take strategic control over their respiratory system and suppress declination to prevent perceptual confusion of tonal entities (Hyman, 1975). Grammaticalizations of the Production code are, according to Gussenhoven (2004), limited to the edges and serve discoursal function e.g. H% may signal continuation of a topic. A phonetic model of pitch range control was introduced, simulating discoursal functions associated to the Production code with the use of register tones (Möhler & Mayer, 2001). Since declination has not been instrumentally investigated for Akan, the following questions were raised in chapter 1 section 1.6.1:

- Do declarative sentences show declination?
- Does the degree of declination differ as a function of tone?
- Is declination phonological or phonetic?
- How can declination be modeled?

To answer the questions, controlled material, consisting of sentences with only H and only L tones differing in length, has been analyzed and presented in chapter 4 section 4.3.1. The investigation of the first question supplied evidence for a positive answer. With regard to the second question, we have seen that the rate of declination is higher in sentences with only H tones than in sentences with only L tones. As a side issue, the relation between sentence length and declination rate has been explored. Longer sentences are realized with a slower declination rate than shorter sentences. Evidence has been provided that declination is grammaticalized in Akan. It has been argued to constitute an intonational difference and was modeled phonologically with the use of post-lexical h and l register tones associating to the left and right edge of the IP respectively. The meaning of declination is to signal coherence (Hansson, 2003). Furthermore, declination was modeled as an exponential decay towards a non-zero asymptote (Liberman & Pierrehumbert, 1984; Shih, 2000) using three values: the initial value, the baseline value *r,* and the lowering quotient *s*. It has been argued that *s* is activated by the presence of the register tones in the phonological surface representation. Crucially, declination in sentences with only L and only H tones was successfully predicted with the same lowering quotient, which was interpreted as further evidence to attribute for declination in the phonological component of the grammar. The quotient of determination has been presented as indicator of the goodness of the model.

- Is automatic downstep phonetically similar to non-automatic downstep?

Another open question for Akan, outlined in chapter 1 section 1.3, was whether automatic and non-automatic downstep are phonetically similar. In the original sense of the terms, Stewart (1965) only wanted to express that the lowering in the former case is brought about by an L tone which is phonetically realized, whereas it is phonetically not realized in the latter case (L̥). No difference between the two “types” of downstep was intended, since an L tone is underlyingly present in both cases. However, Dolphyne (1994) presented empirical evidence showing that the amount of lowering in cases of non-automatic downstep is greater than for automatic downstep. Chapter 4 section 4.1 offered controlled data (Genzel & Kügler, 2011) to answer the question. Two associative constructions, one exhibiting automatic downstep and one exhibiting non-automatic downstep, were embedded in an identical sentence frame in order to exclude any influence of sentence length or tonal configuration. Two measures were presented: the absolute pitch level and the pitch drop between H tones. The analysis of the data has shown that there is no phonetic difference between automatic and non-automatic downstep. This finding was interpreted as evidence that an L tone, whose TBU is deleted, remains unassociated in the tonal string and as evidence that downstep is not due to co-articulation.

As worked out in chapter 1 section 1.3, a general property of downstep in terraced level tone languages is that it does not apply to an initial H tone which is preceded by an L tone (Huang, 1985). This has been empirically verified for Yoruba (Connell & Ladd, 1990; Laniran & Clements, 2003) but not for Akan.

- Does downstep apply to initial H tones which are preceded by an L tone?

In chapter 4 section 4.2, controlled phonetic data, consisting of alternating LH and alternating HL tones, has been analyzed to answer the question. Evidence has been provided that downstep is absent in initial position. This observation is essential because it shows that downstep cannot be a co-articulatory effect. Further, the results have been interpreted to expose two central characteristics. Firstly, downstep is relational, i.e. a later H tone is lowered in relation to a preceding one (Huang, 1985). Secondly, initial tones receive a default value (Huang, 1985). In chapter 1 section 1.3, I raised the question whether downstep is different from declination. Downstep has been modeled phonologically with the use of register tones (e.g. Clements, 1983; Huang, 1985; Snider, 1999) and has been claimed to reflect grammaticalization of the Production code (Gussenhoven, 2004).

- Is downstep to be regarded as independent effect i.e. is downstep different from other pitch lowering processes (declination)? If yes, is downstep to be represented in the phonology or in phonetics?
- How can downstep (automatic and non-automatic) be modeled?

Chapter 4 section 4.3.3 was concerned with empirical evidence to answer the questions. The F0 of tones at comparable positions in sentence with only H or only L tones and alternating LH/HL tones has been investigated. Two measures have been presented: the absolute pitch level and the pitch drop between L or H tones. With regard to the first question, the data has shown that declination and downstep are phonetically similar. It was concluded that downstep does not have to be represented as an independent phonological process (e.g. Huang, 1980, 1985; Clements, 1979, 1983, 1990). The surface terracing pattern in sentences with alternating LH/HL tones follows from the phonological post-lexical register tone specification at the left and right edge of the IP (declination). Supplementary to the last question, the following question was raised in chapter 1 section 1.7:

- Which innovations are required to map abstract phonological entities (tones) onto F0 targets?

In chapter 4 section 4.3.4, automatic and non-automatic downstep have been modelled with the same lowering quotient established for the algorithm that predicts declination. Two innovative procedures were introduced to account for the empirical facts. First, the algorithm can differentiate between H and L. Second, if it detects two different tonal entities in a tonal string, it starts to calculate values for both of them and supplies articulatory goals only to the compatible entity. The quotient of determination has been presented as an indicator of the goodness of the model. The following question is intertwined with phonetic implementation and has been said to provide insights into the size of the processing window. It was raised in chapter 1 section 1.7.

- Do Akan speakers employ anticipatory raising?

Stewart (1965) raised the issue of preplanning in relation to the presence of non-automatic downstep and claimed that Akan speakers would realize an initial H tone one step higher if the sentence contained a non-automatic downstep. Schachter (1965) rejected Stewart's claim. In chapter 5 section 5.1, the initial F0 of sentences containing non-automatic downstep was compared to the initial F0 of sentences containing non-automatic downstep. It has been shown that speakers do not raise the initial H tone in anticipation of the presence of non-automatic downstep. In section 5.2, two different data sets have been analyzed to detect anticipatory raising and to determine which information is available to the phonetic implementation component. The investigation of data set 1, consisting of sentences with only H/L tones differing in length (elongation of the first pP) and sentences with alternating LH/HL tones (elongation of the second pP), has revealed that Akan speakers anticipate the length of the whole IP and not only the length of the first pP (subject) in simple (SVO) structures, as in German (Petrone et al., 2011; Fuchs et al., 2013) or Wenzhou Chinese (Scholz, 2012). Anticipatory raising is present on initial L tones and on initial H tones, irrespective of which one is the first. This fact was interpreted as evidence that pitch range is chosen for both tones (L and H) independently and that this may be taken as adaption to the method of operation of the pitch implementation algorithm. The analysis of data set 2, consisting of a matrix clause with constant length followed by a complementizer clause exhibiting only L/only H tones differing in length, provided further details of preplanning in complex (SVAdvCompSVO) sentences. It has been shown that the information that an embedded complementizer clause is coming up is available i.e. information on the level of the IP. The height of the first H tone of the matrix clause was systematically scaled higher in the complex structures compared to the simple SVO sentences. However, the height of the first H tone of the matrix clause was not raised in anticipation of the length of the embedded IP (complementizer clause). The preplanning effect was located only on the initial tone of the complementizer clause. This fact was interpreted as evidence that specific information of the length of an IP is available later. Preplanning (anticipatory raising) is an important process at the level of pitch implementation. It serves to ensure that declination can be maintained throughout the IP and therefore prevents pitch resetting. Furthermore, in chapter 4 section 4.3.3, examples have been shown which indicate that the pitch implementation algorithm calculates articulatory goals more accuratly if anticipatory raising was applied.

One of the major topics of the thesis was intonation. In chapter 1 section 1.5.2, it was outlined that the intonational marking of sentence type (Yes – No questions) in Akan is typologically interesting because it combines H-pitched and non H-pitched features (Rialland, 2007). Furthermore, Akan has been classified as "lax" prosody language (Rialland, 2009). The following questions were raised.

- What is the intonational morpheme marking Yes – No questions in Akan and which effects are a by-product of the phonetic implementation of it?
- How can Akan be classified along the lines of Rialland's (2007, 2009) typology?

The empirical investigation of material consisting of string-identical statements and Yes – No questions was undertaken in chapter 6. Concerning the first question, it has been argued that the intonational question morpheme in Akan is L%. Complementry to L% as question marker, it has been proposed in chapter 4, section 4.3.1 that declarative sentences are marked by a toneless boundary tone, 0% (Grabe, 1998), associating to the right edge of the IP. L% has been represented as post-lexical tonal entity on the right periphery of the tonal tier. It has been shown to cause several phonetic side effects such as extra final lengthening, higher intensity and leads generally to an increase in articulatory effort which is also reflected in the formant structure. Furthermore, it has been observed that the pitch register is raised in Yes – No questions which, however, did not affect the declination/downstep pattern. The use of L% as question morpheme deviates from the universal form-function relation summerized under the notion Frequency code (Ohala, 1994; Gussenhoven, 2002, 2004). In chapter 1 section 1.6.2 the following question was raised.

- Does Akan show phonetic compensation for the ‘unnatural’ intonational question morpheme?

In chapter 6, it was argued that the higher register is employed at the phonetic level to compensate for the ‘unnatural’ form of the intonational question morpheme (e.g. Haan, 2002). Concerning the classification of Akan along the lines of Rialland’s typology, it was concluded that it does not meet the description of a “lax” question prosody language because L% is phonetically implemented by means of increasing articulatory effort, and breathy termination is not employed. Thus, it has been suggested to add the new category called “low tense” question prosody languages to Rialland’s (2007, 2009) typology. Apart from intonational marking of sentence type, syntactic and prosodic marking of focus have been an area of interest. In chapter 1 section 1.5.3, it was pointed out that a constituent under narrow informational focus has been observed to remain in-situ, whereas a constituent under narrow corrective focus appears in sentence initial position (ex-situ) and is followed by the particle *na* (e.g. Ermisch, 2006). The following questions, concerning syntactic marking of focus, have been raised.

- How frequently is the ex-situ construction used in general and specifically with corrective focus?
- Do our distributional findings speak in favor of a focus type or emphasis analysis?

The data, obtained from a situation decription task (Genzel & Kügler, 2010), was analyzed in chapter 7 section 7.1. It has been shown that the ex-situ construction is not used in connection with narrow informational focus at all and that uses of ex-situ constructions are rare with narrow corrective focus. These findings have been interpreted with the use of paralinguistic notions such as hearer expectation, discourse expectability (Zimmermann, 2007) and emphasis (Hartmann, 2008). Further, it was suggested that the results speak against the proposal of different types of foci (e.g. Drubig, 2003). Concerning the prosodic marking of focus, the following questions were raised in chapter 1 section 1.5.3.

- Is focus prosodically marked?
- Does the prosodic marking of narrow informational focus differ from that of narrow corrective focus?

Two data sets (laboratory speech & semi-spontenous speech), in which focus was elicited with the help of context questions that put target words either in narrow informational or corrective focus (wide informational focus was elicited without context and served as baseline), have been analyzed, in chapter 7 section 7.2, to answer the questions. The absolute F0, corresponding to the tones on the target words, has been measured for the laboratory speech data (Kügler & Genzel, 2012). The results obtained from the laboratory data set did not provide evidence for a prosodic stategy to signal the focus of a sentence. However, a significantly lowered F0 was observed on target words under narrow corrective focus. In chapter 1 section 1.6.3 the following question was raised supplementary.

- Is emphasis expressed in terms of higher F0?

A relational measure, relating the height of the tones on the target words under wide informational focus and narrow corrective focus to the height of the first H tone in the same sentence, has been presented. The results revealed that the lowering, detected on the target word under narrow corrective focus, is the consequence of a global register lowering effect triggered by the negative particle *daabi,* that preceds those renditions (Greif, 2012). Furthermore, it has been shown that target words that exhibit an H tone are systematically raised under narrow corrective focus, which was interpreted as manifestation of the Effort code; emphasis is signaled in terms of higher F0 (Gussenhoven, 2002, 2004). The analysis of the semi-spontanous data set relied on the following variables: pause occurance and duration, pre-boundary lengthening and occurrence of glottal stops. The data has shown that focus is marked by insertion or enhancement of a prosodic boundary (φ), presumably to the right of the focused constituent. The boundary is accompanied or enhanced by a silent pause before and after the focused element. Furthermore, the use of the non-tonal intonational feature glottal stop (Dolphyne, 1988) has been observed in connection with focus. The glottal stop appeared before and after the focused element. Pre-boundary lengthening was not present. The results have been taken to support the claim that focus marking and prominence should be disentangled and that the universal reflex of focus is alignment (Féry, 2012). Furthermore, it was argued that non-tonal features such as the glottal stop should be included into the definition of intonation (Hyman & Monaka, 2008).

Finally, co-articulatory tonal processes that affect the height of tones in connected speech have been of interest. In chapter 1 section 1.8, the following questions were raised.

- Are L tones subject to local carry-over raising from a preceding H tone?
- Are H tones subject to local raising when they are followed by an L tone?

The first questions were investigated in chapter 3 section 3.1. Two different measures were presented: the height of the first L tone in a sentence with alternating LH tones was compared to the height of the second L tone and the height of an L tone in a sentence with only L tones was compared to the height of an L tone in a sentence with alternating LH/HL tone at a comparable position. The results have shown that L raising is present in Akan. L raising has been analyzed as local carry-over effect (co-articulation), following Gandour et al., (1994) and Laniran & Clements (2003). The second question has been tackled in chaper 3 section 3.2. Three different measures have been presented: the height of the first H tone in a sentence with only H tones was compared to the height of the first H tone in a sentence with alternating LH/HL tones, the height of a later H tone in a sentence with only H tones was compared to the height of a later H tone in a sentence with alternating LH/HL tones and the height of an H tone in which the L tone is one, two or three syllables away to the H tones at the same position in a sentence with only H tones. The results revealed that H raising is present in Akan. It is most pronounced on an initial H tone which is immediately followed by an L tone. The effect is not detectable on later H tones. It has been argued that H raising is not a co-articulatory effect, since it is also triggered by L̥ (Gussenhoven, 2004), but a local anticipatory planning effect which functions to make tones of the opposite identity more distinct in the tonal space (Chen, 2012).

The thesis offered controlled experimental data on Akan. Special attention was given to the investigation of declination since it can be regarded as instrumentally understudied in African tone languages, albeit the fact that interesting interactions with the number and nature of lexical tones in a language have been observed (Hombert, 1974; Hyman, 1975). The substantial presence of declination in Akan has lead to the conclusion that it exhibits only two lexical tones L and H. The investigation of the relationship of declination and downstep presented here may be regarded as pilot study whose result will hopefully lead to further research on the matter in other (tone) languages. What has been commonly described as downstep in many typologically different languages may be caused by fundamentally different mechanisms. It can be purely co-articulatory as in Mandarin Chinese (e.g. Wang & Xu, 2011), it can be lexicalized as in Dschang (Hombert, 1974), it can be phrasal as in Tswana (Zerbian & Kügler, 2012) and it can be the result of phonologized declination as proposed here for Akan. A phonological account of declination, involving post-lexical h and l register tones (Möhler & Mayer, 2001), has been offered. Further research is needed to shed light on the question how these post-lexical tones interact with discourse related pitch range/register effects. The proposed analysis may, hopefully, motivate further research on the nature of downstep in other terraced level tone languages.

The investigation of the use of intonational post-lexical tones, as an indicator of sentence type, has revealed that a complex interplay of phonetic surface effects can be broken down into a phonological trigger, L% as question morpheme marking Yes – No questions in Akan, and by-products of the phonetic implementation of it (extra final lengthening, intensity increase, lack of final lowering). The proposed methodology may hopefully inspire further research in the field. Special attention was given to the co-occurrence of register raising and L%. This combination has been observed in a few other tone languages e.g. Konni, Ga, Izon and Turkana (Rialland, 2007). Furthermore, a higher register has been observed in connection with breathy termination e.g. Ikaan (Salffner, 2010) and in connection with cancelation of penultimate lengthening e.g. Zulu and Southern Sotho (Rialland, 2007). All these languages show, besides register raising, phonetic characteristics contradicting the view of the Frequency code that predicts a correlation of questions with some kind of high pitch (Ohala, 1994; Gussenhoven, 2002, 2004). It has been suggested here that the higher register in Akan reflects a compensation of L% at the phonetic level to satisfy the Frequency Code. However, in other languages a higher register may appear as a by-product of the phonetic implementation of H% (van Heuven & Haan, 2002; Myers, 2004). The proposed analysis may hopefully lead to further investigations of the interplay between phonology and phonetics.

The investigation of the frequency of occurrence of the ex-situ construction as a grammatical device signaling narrow corrective focus as a focus type (Drubig & Schaffar, 2001), with the help of a situation description task (Genzel & Kügler, 2010), has revealed that it is not obligatorily used. Comparable results have been obtained for Hausa, with the help of a corpus study (Hartmann & Zimmermann, 2007). The findings point to the need to conduct more empirical investigations in other languages which have been observed to use ‘focus’ constructions. The results may add to the ongoing debate on focus types, their status as grammatical categories and the relationship of syntax and information structure (Hartmann, 2008; Fanselow & Lenertová, 2011). It has been shown here, and in Genzel & Kügler (2010), that focus frequently remains in-situ. The use of prosodic strategies of focus marking in Akan has received attention only recently. The data (Kügler & Genzel, 2012) obtained from controlled laboratory settings is less prone to reveal insights into the prosodic marking of focus (Féry, 2012). The analysis of semi-spontaneous data has disclosed that Akan marks focus, on objects in canonical position, by insertion of a phonological phrase boundary and insertion of a glottal stop. The findings point to the need to conduct experiments that elicit semi-spontaneous/spontaneous speech material, to gain a better understanding of the grammatical expression of focus in a language; e.g. Questionnaire for Information Structure (Skopeteas et al., 2006).

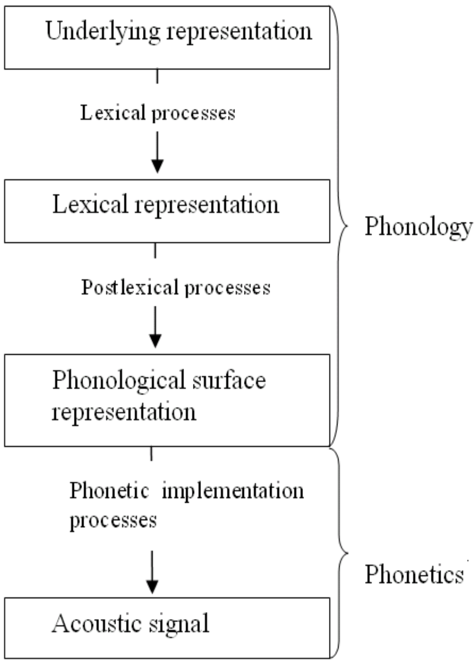
The investigation of processes applying at the level of phonetic implementation has shown that, besides lexical tones (L/H), post-lexical tones (0%, L%, l & h), the glottal stop as non-tonal post-lexical feature and pitch variation related to the biological codes, local and global anticipatory processes play a role in determining the surface characteristics of the acoustic signal in Akan. This finding is not new. Laniran & Clements (2003:243) conclude their seminal paper on interacting factors in the tone production of Yoruba with the following words: “This study has motivated a compositional approach to tone production in which the overall shape of an f_0 contour is viewed as resulting not from any single factor (such as downstep) but from the interaction of a number of factors, each of which can vary independently of the others to at least some degree.”. This thesis has provided insights into factors that contribute to the overall shape of the F0 contour in Akan. Quantitative data on the impact of each factor has been presented. Furthermore, it has been discussed at which stage of the speech production process the specific factors come into play, which functions they serve and how they interact.

The overall picture emerging from the presented phonetic investigations is illustrated in figure 99, which revisits figure 1. Akan is a tone language. This is exemplified with the proper name *Anane*, the verb *bisa* - ‘to ask’ and the noun *sika* - ‘money’ at the top of the righthand side of figure 99. *Anane* exhibits an H tone on the antepenultimate, an L tone on the penultimate and a H tone tone on the ultimate syllable. The verb and the noun *sika* carry an L on the penultimate syllable and an H tone on the ultimate syllable. The tones are part of the lexical entry of these Akan words. Let us assume that an Akan speaker wants to form a statement out of these three words with *Anane* as subject, *sika* as object and the verb *bisa* in the habitual aspect. The habitual aspect is regarded as underlying form. No lexical tonal processes apply. Thus, the underlying representation is also the lexical representation in this case. The lexical representations are then linearized into the sentence *Anane bisa sika*. ‘Anane asks for money.’, which exhibits a syntactic SVO structure. The assumed syntactic structure of this simple SVO sentences is represented in (124)a. Post-lexically the syntactic structure is mapped onto the prosodic structure. The prosodic structure of the sentence is illustrated in (124)b.

(124) a. Syntactic structure: [$_{\text{InflP}}$ [$_{\text{NP}}$ Ánànè́] [$_{\text{VP}}$ bìsá [$_{\text{NP}}$sìká]]

b. Prosodic structure: ((Ánànè́)φ (bìsá (sìká)φ)φ)ι

Furthermore, post-lexical tones are inserted. These are: post-lexical tones mediating sentence type meaning (0% or L%) and post-lexical register tones (h & l) signaling information about the prosodic structure at the IP level.



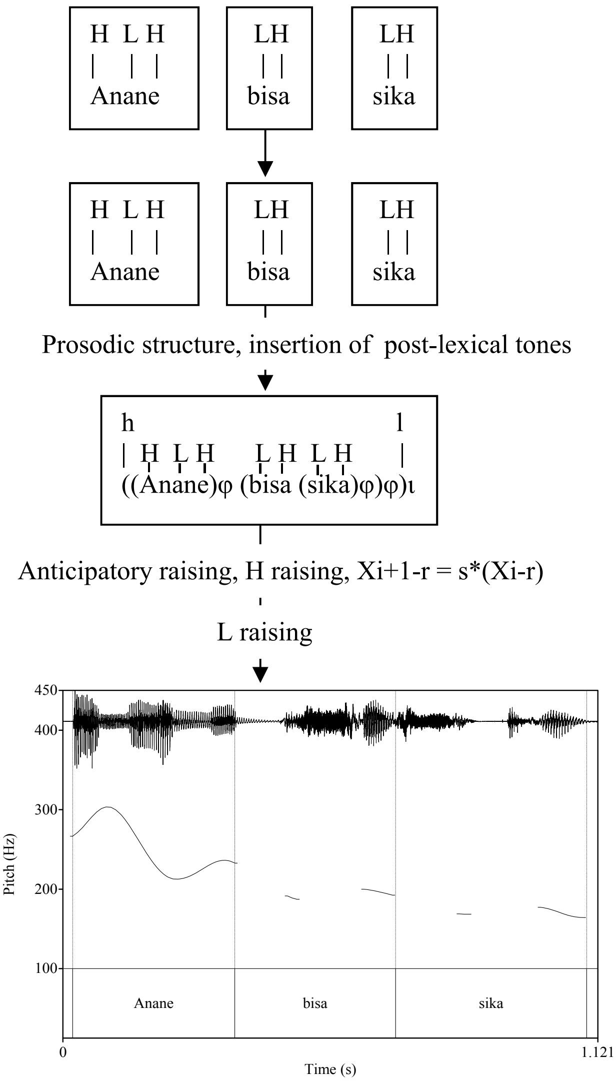


figure 99: Model of Lexical Phonology with information about lexical tones, post-lexical processes and phonetic processes of Akan.

The phonological surface representation of the sentences is displayed at the right hand side of figure 99. Since the example sentence is a statement, no tonally specified boundary tone occurs. The post-lexical register tones, h and l, associate with the left and right edge of the IP. As soon as the phonological surface representation is available, the phonetic implementation component starts to plan globally and locally. Higher level information of the length/complexity of the sentence is taken into account to set an adequate initial reference value X_i via anticipatory raising. Furthermore, local information of the immediately neighboring tones is used to dissimilate H to L via H raising. Once the planning of the acoustic goals has taken place, the phonetic implementation algorithm starts to compute acoustic goals (tonal targets), see right hand side of figure 99 for details. The presence of the register tones in the phonological presentation leads to an activation of the lowering quotient *s* in the pitch implementation algorithm. The lowering quotient, which stabilizes at 0.75, ensures that the F0 of the sentence decreases gradually. The figure 100 provides insights into the operation mode of the pitch implementation algorithm in Akan. The initial H tone and L tone values serve as reference and input values for the calculation of the tonal targets and set the pitch register of the utterance within the speakers' pitch range. To calcuate the target of the seocnd tone, the argorithm lowers the intial tone H_i, which is a H tone, by *s* in relation to the baseline quotient *r*. When it encounters the second tone and it is an H tone the value for H_{i+1} is generated on the output. In our case it is a L tone, hence, the value H_{i+1}, is stored in the memory and a second algortihm, operating with the same quotients as the H tone algorithm, starts to calculate the value for L_{i+1} with L_i as input. Since the tone following L_i is an H tone, the value L_{i+1}, is stored in the memory. The H tone algorithm generates the value for H_{i+2}, which was calculated with H_{i+1} as input value, on the output. This procedure applies iteratively from left to right with one tone look-ahead following Pierrehumbert (1980).

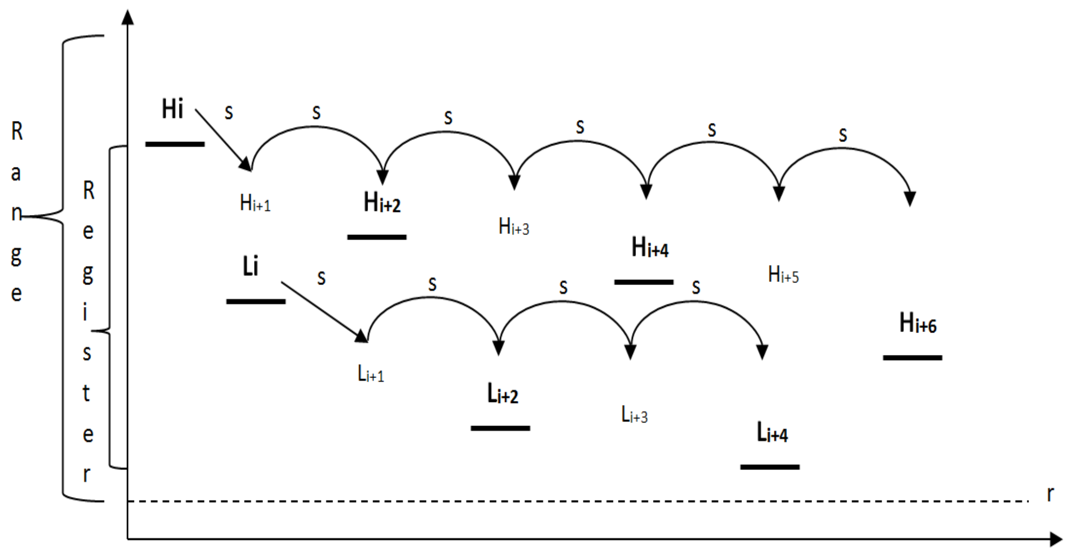


figure 100: Operation mode of the pitch implementation algorithm in Akan. H_i & L_i are initial reference values, *s* is the lowering quotient and *r* is the baseline quotient.

Once, the abstract phonological representations have been mapped onto concrete articulatory/acoustic targets, articulation takes place. During articulation, L tones are raised above their intended values. The end-product of the derivation is the acoustic signal. Its F0 and waveform are displayed on bottom of figure 99.

It has been demonstrated that the charachteristics of the F0 contour of the acoustic signal in Akan are attributed to the interplay of a number of factors. The F0 contour is a result of the composition of lexical tones, post-lexical tones, planning processes, pitch implementation and co-articulatory processes.

Bibliography

Abakah, E. N. (2000). A closer look at downstep in Akan. *Afrika und Übersee, 83*, 1-23.

Abakah, E. N. (2002). The low tone in Akan. In *Aktuelle Forschungen zu afrikanischen Sprachen,* Schumann, T., Reh, M., Kießling, R. & Gerhardt, L. (eds), Cologne: Koeppe. 193-210.

Abakah, E. N. (2005). Tone rules in Akan. *Journal of West African Languages, XXXII,* 1-2, 109-134.

Abakah, E. N. & Koranteng, L. (2007). The Interaction of Tone, Syntax and Semantics in Akan. *Studies in the Languages of the Volta Basin, 4, 2,* 63-86.

Abramson, A. S. (1979). The Coarticulation of Tones: An Acoustic Study of Thai. In *Studies in Thai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J.A. Henderson*, Thongkum T. L., Kullavanijaya, P., Panupong V. & Tingsabadh, K. (eds), 1-9.

Adomako, K. (2008). *Vowel epenthesis and vowel deletion in loanwords: A study of Akan*. MA Thesis, Department of Language and Linguistics, Tromsø.

Adriaens, L. M. H. (1991): *Ein Modell deutscher Intonation.* Phd Thesis, TU Eindhoven, IPO.

Ahoua, F. (1996). *Prosodic Aspects of Baule, with special reference to the German of Baule speakers*. Cologne: Koeppe.

Ameka, F. K. (2010). Information packaging in Kwa: micro-variation and typology. In *Studies in Kwa syntax* (Studies in Natural Language and Linguistic Theory, 78), Aboh. E. & Essegbey, J. (eds), Dordrecht: Kluwer Publishers, 141-176.

Amfo, N. A. A. (2007). Akan Demonstratives. In *Selected Proceedings of the 37th Annual Conference on African Linguistics.* Payne, D. & Peña, J. (eds), Sommerville: Cascadilla Press. 134-148.

Amfo, N. A. A. (2010a). Lexical signaling of information structure in Akan. *Linguistics*, *48*, 95-225.

Amfo, N. A. A. (2010b). Noun Phrase Conjunction in Akan: the Grammaticalization Path. *Pragmatics, 20, 1*, 27-41.

Anderson, J. C. (2009). Preliminary to preliminary: Studies of speech rhythm in Akan/Twi. In *African linguistics across the discipline: Indiana University Linguistics Club Working Papers, 10*, Anderson, J. C., Green, C. R. & Obeng, S. G. (eds), Bloomington, IN: IULC Press, 133-143.

Anderson, J. C. (2011). A Preliminary Study of Speech Rhythm in Akan. Paper presented at the *Acoustical Society of America*, San Diego.

Anderson, S. (1978). Tone features. In *Tone: A linguistic survey.* Fromkin, V. (eds), *Tone: A linguistic survey.* New York: Academic Press.133-176.

Arkoh, R. B. (2011). *Semantics of Akan bi and nY.* MA Thesis, University of Ghana & University of British Columbia.

Armstrong, R. (1968). Yala (Ikom): A terraced-level language with three tones. *Journal of West African Languages, 5*, 49-58.

Arvaniti, A. & Godjevac, S. (2003). The origins and scope of final lowering in English and Greek. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Universidad Autónoma de Barcelona, 1077-1080.

Arvaniti, A. (2007). On the presence of final lowering in British and American English. In *Tone and Tunes, vol:2: Experimental Studies in Word and Sentence Prosody*, Gussenhoven, C. & Riad, T. (eds), Berlin and New York: Mouton de Gruyter, 317-347.

Asu, E. L., Schötz, S. & Kügler, F. (2009). The acoustics of Estonian Swedish long close vowels as compared to Central Swedish and Finland Swedish. In *Proceedings of Fonetik.* Stockholm.

Baayen, R. H. (2008) *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.

Baković, E. (2003). Vowel harmony and stem identity. *San Diego Linguistics Papers, 1*, *2,* 1-35.

Ballard, L. W. (2010). *Akan vowel harmony and optimality theory*. MS Thesis. University of Florida.

Baumann, S. (2006). *The intonation of givenness.* Tübingen: Niemeyer.

Baumann, S., Grice, M. & Steindamm, S. (2006). Prosodic Marking of Focus Domains - Categorical or Gradient?. In *Proceedings of Speech Prosody 2006,* Hoffmann, R. & Mixdorff, H. J. (eds). Dresden: TUDpress, 301-304.

Baumann, S., Becker, J., Grice, M. & Mücke, D. (2007). Tonal and Articulatory Marking of Focus in German. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, 1029-1032.

Bearth, T. (2002). Fokus im Akan - Outline. Handout of a presentation at the *Linguistic Colloquium of the Seminar of African Studies* at the Humboldt-University, Berlin.

Becker, L. A. (1979). Once more on the nature of downdrift. *Studies in African Linguistics*, *10*, *3*, 233-245.

Beckman, M. E. & Pierrehumbert, J. (1992). Comments on chapters 14 and 15. In *Papers in laboratory phonology II: gesture, segment, prosody*, Docherty, G. J. & Ladd, D. R. (eds), Cambridge: Cambridge University Press, 387-397.

Beckman, M. E. & Venditti, J. J. (2010). Tone and Intonation. In *The Handbook of Phonetic Sciences*, Wiley-Blackwell, 603-652.

Berry, J. & Aidoo, A. A. (1975). *An introduction to Akan*. Evanston: Northwestern University.

Bickel, B., Comrie, B. & Haspelmath, M. (2004). *The Leipzig Glossing Rules: Conventions for Interlinear Morpheme-by-Morpheme Glosses*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://www.eva.mpg.de/.

Bickel, B., Hildebrandt, K. & Schiering, R. (2009). The distribution of phonological word domains: a probabilistic typology. In *Phonological domains: universals and deviations*, Grijzenhout, J. & Kabak, B. (eds), Berlin: Mouton de Gruyter, 47-75.

Bickmore, L. S. (2000). Downstep and fusion in Namwanga. *Phonlogy, 17*, 297-331.

Bird, S. (1994). Automated Tone Transcription. In *Proceedings of the First Meeting of the ACL Special Interest Group in Computational Phonology*, Las Cruces, 1-12.

Boadi, L. A. (1974). Focus-marking in Akan. *Linguistics: An International Review*, *140*, 5-57.

Boadi, L. A. (1990). Questions in Akan. *Frankfurter afrikanistische Blätter, 2,* 70-92.

Boadi, L. A. (2005). *Three major syntactic structures in Akan Interrogatives, Complementation and relativisation*. Black Mask Limited.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* (Version 5.3.21) [Computer program]. Retrieved from http://www.praat.org/.

Bolinger, D. L. (1962). Intonation as a universal. In *Proceedings of the 9th International Congress of Linguistics,* Lunt, H.G. (ed), The Hague: Mouton, 833-844.

Bolinger, D. L. (1978). Intonation across languages, In *Universals of human language*, Greenberg, J. (eds), Stanford: Stanford University Press. 471-524.

Breckenridge, J. (1977). The declination effect. *Journal of the Acoustical Society of America, 61*, 1-90.

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

Bruce, G., Granström, B., Gustafson, K. & House, D. (1991). Prosodic phrasing in Swedish, In *Lund Working Papers in Linguistics, 38*, 5-17.

Brunelle, M. (2009). Northern and Southern Vietnamese Tone Coarticulation: A Comparative Case Study, *Journal of the Southeast Asian Linguistics Society, 1*, 49-62.

Büring, D. (1997). *The meaning of topic and focus: The 59th Street Bridge accent*. London/New York: Routledge.

Büring, D. (2010). Towards a typology of focus realization. In *Information Structure. Theoretical, Typological, and Experimental Perspectives,* Zimmermann, M. & Féry, C. Oxford: Oxford University Press, 177-205.
Büring, D. (2012). Focus and Intonation. In *Routledge Companion to the Philosophy of Language,* Russell G. & Graff Fara, D. (eds), London: Routlegde, 103-115.
Byrd, D. (1992). Pitch and duration of yes-no questions in Nchufie. *Journal of the International Phonetic Association*, *22*, *1 & 2*, 12-26.
Cahill, M. (1985). *An autosegmental analysis of nasality and tone*. MA Thesis, Texas: Arlington.
Callier, P. (2011). Sociophonetics on the (right) edge: Boundary tones and final lengthening in Mandarin Chinese. *eVox: Georgetown Working Papers in Language, Discourse, and Society, 5, 1*, 16-36.
Campione, E. & Véronis, J. (2002). A large-scale multilingual study of silent pause duration. In *Proceedings of the Speech Prosody 2002*. Bel, B. & Marlien, I. (eds). Aix-en-Provence: Laboratoire Parole et Langage. 199-202.
Chen, Y. (2012). Tonal Variation. In *The Oxford handbook of laboratory phonology*, Cohn, C. A., Fougeron, C. & Huffman, M. K. (eds), Oxford University Press, 103-114.
Childs, J. T. (2003). *An Introduction to African Languages*. Amsterdam/Philadelphia: John Benjamins.
Christaller, J. G. (1875). *A grammar of the Asante and Fante language*. Basel: Evangelical Mission Society.
Christaller, J. G. (1933). *Dictionary of the Asante and Fante language called Tshi (Twi)*. Basel: Evangelical Missionary Society.
Clark, J. & Yallop, C. (1990). *An introduction to phonetics and phonology.* Oxford: Blackwell.
Clark, M. (1993). Representation of downstep in Dschang Bamileke. In *The phonology of tone: the representation of tonal register*, Snider, K. & van der Hulst, H. (eds), Berlin: Mouton de Gruyter, 29-73.
Cleland, E., Gyang, K. O., Imbeah, N. K. & Imbeah, P. K. (2009). *Akan: a concise introduction to the Akuapem, Fanti and Twi language*. 2nd Ed., kasahorow Language Guides.
Clements, G. N. & Ford, K. C. (1979). Kikuyu Tone Shift and its Synchronic Consequences. *Linguistic Inquiry, 10*, 179-210.
Clements, G. N. (1979). The description of terraced-level tone languages. *Language, 55, 3*, 536-558.
Clements, G. N. (1981). Akan vowel harmony: a nonlinear analysis. *Harvard Studies in Phonology, 2,* 108-77.
Clements, G. N. (1983). The hierarchical representation of tone features. In *Current Approaches to African Linguistics, 1*, Dihoff, I. (ed), Dordrecht: Foris. 145-176.
Clements, G. N. (1990). The Status of Register in Intonation Theory: Comments on the Papers by Ladd and by Inkelas and Leben. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech,* Kingston, J. & Beckman, M. E. (eds), Cambridge: Cambridge University Press, 58-72.
Clements, G. N. & Rialland A. (2008). Africa as a phonological area. In *A Linguistic Geography of Africa,* Heine B. & Nurse D. (eds), Cambridge: Cambridge University Press, 36-85.
Clements, G. N, Michaud, A. & Patin, C. (2011). Do We Need Tone Features? In *Tones and Features*, Hume, Goldsmith, J. E. & Wetzels, W. L., Berlin: De Gruyter Mouton, 3-24.
Cohen, A. & 't Hart, J. (1967). On the anatomy of intonation. *Lingua, 19*, 177-192.
Collier, R. (1975). Physiological correlates of intonation pattern. *Journal of the Acoustical society of America, 58*, 249-255.

Cohn, A. C. (2007). Phonetics in Phonology and Phonology in Phonetics. *Working Papers of the Cornell Phonetics Laboratory*. 1-31.
Collier, R. & Gelfer, C. E. (1983). Physiological explanation of F0 declination. In *Proceedings of the 10th international congress of phonetic sciences.* Utrecht, Netherlands. 354-360.
Collier, R. (1985). F0 Declination: The Setting and Resetting of the Baseline, *Ann. Bull. RILP, 19*, 111-132.
Connell, B. & Ladd, D.R. (1990). Aspects of pitch realization in Yoruba. *Phonology, 7, 1*, 1-29.
Connell, B. (1999). Four tones and downtrend: a preliminary report on pitch realization in Mambila. In *New Dimensions in African Linguistics and Languages. Trends in African Linguistics, 3*. Kotey P. (eds), Trenton, N.J.: Africa World Press, 75-88.
Connell, B. (2002a). Downdrift, downstep and declination. In *Typology of African Prosodic Systems*, Gut, U. & Gibbon, D. (eds), Bielefeld: University of Bielefeld, 3-12.
Connell, B. (2002b). Tone languages and the universality of intrinsic F0: evidence from Africa. *Journal of Phonetics, 30*, 101-129.
Connell, B. (2003). Pitch realization and the four tones of Mambila. In *Cross-Linguistic Studies of Tonal Phenomena: Historical development, phonetics of tone, and descriptive studies*. Shigeki, K.(eds), Tokyo, 181-197.
Connell, B. (2004). Tone, utterance length and F0 Scaling. In *Proceedings of the International Symposium on Tonal Aspects of Languages: Emphasis on tone languages.* Bel, B. & Marlien, I. (eds), Bejing, China, 41-43.
Connell, B. (2011). Downstep. In *The Blackwell Companion to Phonology,* van Oostendorp, M, Ewen, C. J., Hume E. & Rice, K. (eds), Wiley-Blackwell, 824-847.
Cooper, W. E. & Sorensen, J. M. (1981). *Fundamental Frequency in Sentence Production.* New York: Springer.
Courtenay, K. (1971). Yoruba: A 'terraced-level' language with three tonemes. *Studies in African Linguistics, 2, 3,* 239-255.
Cruttenden, A. (1997). *Intonation (2nd edn.)*. Cambridge: Cambridge University Press.
Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
de Lacy, P. (1999). *Tone and prominence*. Ms, University of Massachusetts, Amherst. Available as ROA-333 from the Rutgers Optimality Archive.
Delattre, P. (1962). Comparing the prosodic features in English, German, Spanish, and French. *International Review of Applied Linguistics,1*, 193-210.
Dilley, L. C., Shattuck-Hufnagel, S. & Ostendorf, M. (1996). Glottalization of vowel-initial syllables as a function of prosodic structure. *Journal of Phonetics*, *24*, 423-444.
Dilley, L. C. (1997). *The Phonetics and Phonology of Tonal Systems*. Phd Thesis. S.B., Brain and Cognitive Sciences MIT.
Dolphyne, F. A. (1988). *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Ghana: University Press.
Dolphyne, F. A. & Kropp Dakubu, M. E. (1988). The Volta-Comoe Languages. In *The Languages of Ghana*, Kropp Dakubu, M.E. (ed), 50-90.
Dolphyne, F. A. (1994). A phonetic and phonological study of downdrift and downstep in Akan. Paper presented at the *25th annual conference of African Linguistics*, Rutgers University.
Downing, L. J., Mtenje, A. & Pompino-Marschall, B. (2004). Prosody and information structure in Chichewa. In *ZAS Papers in Linguistics*, *37,* Fuchs, S. & Hamann, S. (eds), Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung, 167-186.

Downing, L. J. (2008). Focus and prominence in Chichewa, Chitumbuka and Durban Zulu. In *ZAS Papers in Linguistics*, *49,* Zygis, M. & Fuchs, S. (eds), Berlin: Zentrum Für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung, 47-65.

Downing, L. J. (2009). Stress-Focus Mismatches in Chichewa, Talk held at the *35th conference of the Generativen Grammatik des Südens* (GGS), Leipzig.

Downing, L. J. & Rialland, A. (2012). Local and global effects of intonational register on lexical tone: a comparative case study of Chichewa and Mbosi, Talk held at the *5th European Conference on Tone and Intonation*, Oxford.

Downing, L. J. & Pompino-Marschall, B. (to appear). The focus prosody of Chichewa and the Stress-Focus constraint: A response to Samek-Lodovici (2005). *Natural Language and Linguistic Theory*.

Drubig, H. B. & Schaffar, W. (2001). Focus constructions. In *Language Typology and Language Universals.* Haspelmath, M., König, E., Oesterreicher, W. & Raible, W. (eds), Berlin: de Gruyter, 1079-1104.

Drubig, H. B. (2003). Toward a Typology of Focus and Focus Constructions. *Linguistics, 41*, 1-50.

Eady, S. J. & Cooper, W. E. (1986). Speech Intonation and Focus Location in Matched Statements and Questions. *Journal of the Acoustical Society of America*, *80*, 402-415.

Eady, S. J., Cooper, W. E., Kloouda, G. V., Mueller, P. R. & Lotts, D.W. (1986). Acoustical Characteristics of Sentential Focus: Narrow vs. Broad and Single vs. Dual Focus Environments. *Language and Speech*, *29*, 233-251.

Elordieta, G. (2008). An overview of theories of the syntax-phonology interface. *International Journal of Basque Linguistics and Philology, 42,* 209-286.

Erickson, D. M. (1976). *A physiological analysis of the tones of Thai*. Ph.D. thesis, University of Connecticut.

Erickson, D. M., Honda, K., Hirai, H. & Beckman, M. E. (1995). The production of low tones in English intonation, *Journal of Phonetics, 23*, 179-188.

Ermisch, S. (2006). Focus and topic Akan constructions in Akan. In *Frankfurter Afrikanistische Blätter, 18,* Ermisch, S. (eds), Köln: Köppe, 51-68.

Fanselow, G. & Lenertová, D. (2011). Left peripheral focus: Mismatches between syntax and information structure. *Natural Language and Linguistic Theory, 29*, 169-209.

Farnetani, E. & Recasens, D. (2010). Coarticulation and connected speech processes. In *Handbook of phonetic sciences*, 2nd Ed., Hardcastle, W. J., Laver, J. & Gibbon, F. E. (eds), Wiley - Blackwell, 316-352.

Faytak, M. & Yu, A. C. L. (2011). A typological study of the interaction beween level tones and duration. In *Proceedings of the 17th international congress of phonetic sciences.* Lee, W. S. & Zee, E (eds), Hong Kong, 659-662.

Fedden, S. (2007). *A grammar of Mian, a Papuan language of New Guinea*. Phd Thesis, University of Melbourne.

Ferreira, F. (1993). Creation of Prosody During Sentence Production. *Psychological Review, 100, 2*, 233-253.

Ferreira, F. & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*, 57-84.

Féry, C. & Truckenbrodt, H. (2005). Sisterhood and tonal scaling. *Studia Linguistica, 59, 3*, 223-243.

Féry, C. & Krifka, M. (2008). Information structure. Notional distinctions, ways of expression. In *Unity and diversity of languages*, Piet van Sterkenburg (ed), Amsterdam: John Benjamins, 123-136.

Féry, C. & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German, *Journal of Phonetics, 36,* 680-703.

Féry, C. (2010). Syntax, information structure, embedded prosodic phrasing, and the relational scaling of pitch accents. *The Sound Patterns of Syntax*, Erteschik-Shir, N. & Rochman, L. (eds), Oxford: Oxford University Press. 271-290.

Féry, C., Hörnig, R. & Pahaut, S. (2010). Phrasing in French and German: an experiment with semi-spontaneous speech In *Intonational Phrasing at the Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic*, Gabriel, C. & Lleó, C. (eds), Amsterdam: John Benjamins Publishing Company, 11-41.

Féry, C. & Schubö, F. (2010). Hierarchical prosodic structures in the intonation of center-embedded relative clauses. *The Linguistic Review 27, 3*, 293-317.

Féry, C. (2012). Focus as prosodic alignment. To appear in *Natural Language and Linguistic Theory*, 1-52.

Fiedler, I. & Schwarz, A. (2005). Out-of-focus Encoding in Gur and Kwa. In *Interdisciplinary Studies on Information Structure (ISIS), 3*, Ishihara, S., Schmitz, M. & Schwarz, A. (eds), Potsdam: Potsdam University. 111-142.

Fiedler, I. & Schwarz, A. (2007). Narrative focus strategies in Gur and Kwa. In *Focus strategies in Niger-Congo and Afroasiatic – On the interaction of focus and grammar in some African languages,* Enoch, A. Hartmann, K. & Zimmermann, M. (eds), Berlin: Mouton de Gruyter, 267-286.

Fiedler, I., Hartmann, K., Reineke, B., Schwarz, A. & Zimmermann, M. (2010). Subject Focus in West African Languages. In *Information Structure from Different Perspectives*, Zimmermann, M. & Féry, C. (eds), Oxford:Oxford University Press, 234-257.

Fletcher, J. (2010). The Prosody of Speech: Timing and Rhythm. In *The Handbook of Phonetic Sciences*, Wiley-Blackwell, 523-602.

Fougeron, C. & Jun, S. A. (1998). Rate Effects on French Intonation: Phonetic Realization and Prosodic Organization, *Journal of Phonetics, 26,* 45-70.

Frajzyngier, Z. (1989). *A grammar of Pero*. Berlin: Dietrich Reimer Verlag.

Fuchs, S., Petrone, C., Krivokapić, J. & Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics, 41*. 29-47.

Gandour, J., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., Boongird, P., & Potisuk, S. (1993). Tonal coarticulation in Thai disyllabic utterances after unilateral brain damage: A preliminary study. *Nopparat Rajathanee General Hospital Medical Journal, 4,* 24-47.

Gandour, J., Potisuk, S. & Dechongkit, S. (1994) Tonal coarticulation in Thai , *Journal of Phonetics, 22* ,477-492.

Gee, J. P. & Grosjean, F. (1983). Performance structures. A psycholinguistic and linguistic appraisal. *Cognitive Psychology, 15*, 411-458.

Genzel, S., & Kügler, F. (2010). How to elicit semi-spontaneous focus realizations with specific tonal patterns. *Working Papers of the SFB 632*, *Interdisciplinary Studies on Information Structure (ISIS), 13*, Potsdam: University Press Potsdam, 77-102.

Genzel, S. & Kügler, F. (2011). Phonetic realization of automatic (downdrift) and non-automatic downstep in Akan. In *Proceedings of the 17th international congress of phonetic sciences.* Lee, W. S. & Zee, E (eds), Hong Kong, 735-738.

Greif, M. (2012). *Corrective Focus in Mandarin Chinese: "A Question of Belief?"*, Lincom Europa.

Goldsmith, J. (1976). *Autosegmental Phonology*. Phd Thesis, MIT.

Goldsmith, J. (1979). *Autosegmental phonology*. New York: Garland.

Grabe, E. (1998). *Comparative Intonational Phonology: English and German*. (MPI Series in Psycholinguistics 7). Wageningen: Ponsen and Looijen.

Grice, M. (2006). Intonation. In *Encyclopedia of Language and Linguistics, vol. 5*, Brown, K. (ed), Elsevier: Oxford, 778-788.

Gussenhoven, C. & Rietveld, A. C. M. (1988). Fundamental frequency declination in Dutch: Testing three hypotheses. *Journal of Phonetics, 16*, 355-69.

Gussenhoven, C. & Jacobs, H. (1998). *Understanding Phonology*. London: Hodder Arnold. New York: Oxford University Press.
Gussenhoven, C. (1999). Discreteness and gradience in intonational contrasts. *Language and Speech, 42*, 283-305.
Gussenhoven, C. & Chen, A.-J. (2000). Universal and language-specific effects in the perception of question intonation**,** In *Proceedings of the International Conference on the Processing of Spoken Language (ICSLP), 6, vol. 2,* Beijing, 91-94.
Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and Phonology. In *Proceedings Speech Prosody*, Aix-en-Provence, 47-57.
Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.
Gussenhoven, C. & Teeuw, R. (2008). A moraic and a syllabic H-tone in Yucatec Maya. In *Fonología instrumental: Patrones fónicos y variación*, Herrera, E. Z. & Butrageño, P. M. (eds), Mexico City: El Colegio de México, 49-71.
Haan, J., van Heuven, V. J., Pacilly, J. J. A. & van Bezooijen, R. (1997). An Anatomy of Dutch Question Intonation, *Linguistics in the Netherlands, 14*, 97-108.
Haan, J. (2002). *Speaking of Questions: An Exploration of Dutch Question Intonation*. Phd Thesis, Utrecht: Netherlands Graduate School of Linguistics (LOT).
Halliday, M. A .K. (1985). *An Introduction to Functional Grammar*. London: Arnold.
Hansson, P. A. (2003). *Prosodic phrasing in spontaneous Swedish*. Phd Thesis, Lunds Universitet, Sweden.
Harley, M. (2005). *A descriptive grammar of Tuwuli, a Kwa language of Ghana*. Phd Thesis, School of Oriental and African Studies, London.
Hartmann, K. (2007). Focus and tone. In *Interdisciplinary Studies on Information Structure, The Notions of Information Structure (ISIS), 6,* Féry, C., Fanselow, G. & Krifka, M. (eds), 221-235.
Hartmann, K. & Zimmermann, M. (2007). In place – out of place? Focus in Hausa. In *On information structure, meaning and form: Generalizing across languages,* Schwabe, K. & Winkler, S. (eds), Amsterdam: Benjamins, 365-403.
Hartmann, K. (2008). Focus and Emphasis in Tone and Intonation Languages. In *The discourse potential of underspecified structures*. Steube, A. (ed), Berlin: Walter de Gruyter. 389-412.
Heeren, W. F. L. & van Heuven, V. J. (2011). Acoustics of whispered boundary tones: effects of vowel type and tonal crowding. In *Proceedings of the 17th International Congress of Phonetic Sciences,* Lee, W. S. & Zee, E (eds), Hong Kong, 851-854.
Herman, R. (1996). Final lowering in Kipare. *Phonology, 13*, 171-196.
Herman, R., Beckman, M. E. & Honda, K. (1996). Subglottal pressure and final lowering in English. In *Proceedings of the 4th International Conference on Spoken Language Processing*, Bunnell, H. T. & Idsardi, W. (eds), Philadelphia, 145-148.
Herman, R. (2000). Phonetic markers of global discourse structures in English. *Journal of Phonetics, 28,* 466-493.
Hirschberg, J. (2002). The Pragmatics of Intonational Meaning. In *Proceedings of the Speech Prosody 2002 conference,* Bel, B. & Marlien, I. (eds), Aix-en-Provence: Laboratoire Parole et Langage, 65-68.
Hockett, C. F. (1963). The problem of universals in language. In *Universals of Language,* Greenberg, J. H. (ed), Cambridge: MIT Press, 1-29.
Hombert, J. M. (1974). Universals of downdrift: their phonetic basis and significance for a theory of tone. *Studies in African Linguistics, 5,* 169-183.
Horne, M., Strangert, E. & Heldner, M. (1995). Prosodic boundary strength in Swedish: Final lengthening and silent interval duration. In *Proceedings of the 13th International Congress of Phonetic Sciences,* Elenius, K. (ed), Stockholm, Sweden, 170-173.

Huang, C. T. J. (1980). The metrical structure of terraced-level tones. *NELS, 10*, 257-70.
Huang, C. T. J. (1985). The Autosegmental and Metrical Nature of Tone Terracing. In *African Linguistics: Essays in Memory of M. W. K. Semikenke,* Goyvaerts, D. L. (ed), Amsterdam: John Benjamins, 209-238.
Hyman, L. M. (1975). *Phonology: theory and analysis*. New York: Holt, Rinehart & Winston.
Hyman, L. M. (1985). Word Domains and Downstep in Bamileke-Dschang, *Phonology, 2*, 47-83.
Hyman, L. M. (1989). The phonology of final glottal stops. In *Proceedings of WECOL 1988*, 113-130.
Hyman, L. M. (1993). Register tones and tonal geometry. In *The phonology of tone: the representation of tonal register*, Snider, K. & van der Hulst, H. (eds), Berlin: Mouton de Gruyter, 75-100.
Hyman, L. M. (2001). Tone systems. In *Language Typology and Language Universals: An International Handbook*, Haspelmath, M., König, E., Oesterreicher, W. & Reible, W. (eds), Berlin: Mouton de Gruyter, 1367-1380.
Hyman, L. M. (2003). African languages and phonological theory. *GLOT International,* 7, *6,* 153-163.
Hyman, L. M. (2007). Universals of tone rules: 30 years later. In *Tones and Tunes, vol. 1: Studies in Word and Sentence Prosody*. Gussenhoven, C. & Riad, T. (eds), Berlin: Mouton de Gruyter, 1-34.
Hyman, L. M. & Monaka, K. C. (2008). Tonal and non-tonal Intonation in Schekgalagari, *UC Berkeley phonology Lab Annual Report*, 269-288.
Hyman, L. M. (2009). Penultimate lengthening in Bantu: Analysis and Spread. In *UC Berkeley Phonology Lab Annual Report*, 195-209.
Hyman, L. M. (2011). Tone: is it different? In *The Handbook of Phonological Theory*, 2nd Ed., Goldsmith, J., Riggle, J. & Yu, A. (eds), Blackwell, 197-240.
Ikekeonwu, C. (1993). Intonation and Focus: A Reanalysis of Downdrift and Downstep in Igbo. In *Lund Working papers in Linguistics, 40*, 95-112.
Inkelas, S., Leben, W. R. & Cobler, M. (1987). The Phonology of Intonation in Hausa. In *Proceedings of NELS, 17,* McDonough, J. & Plunkett, B. (eds), Amherst, MA: GLSA, 327-341.
Inkelas, S. & Leben, W. R. (1990). Where phonology and phonetics intersect: The case of Hausa intonation. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Kingston J. & Beckman, M. E. (eds), Cambridge, UK: Cambridge University Press. 17-34.
Itô, J. & Mester, A. (2007). Prosodic adjunction in Japanese compounds. In Miyamoto Y. & Ochi M. (eds.), *Formal approaches to Japanese linguistics (FAJL) 4*, Cambridge, MA: MITWPL. 97–111.
Jun, S. A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology, 15,* 189-226.
Jun, S. A. (2003). Prosodic Phrasing and Attachment Preferences. *Journal of Psycholinguistic Research*, *32, 2,* 219-249.
Karlsson, A., House, D., Svantesson, J.-O., & Tayanin, D. (2007). Prosodic phrasing in tonal and ton-tonal dialects of Kammu. In *Proceedings of the 16th International Congress of Phonetic Sciences*, Trouvain, J. & Barry, W. J. (eds), Saarbrücken, 1309-1312.
Karneva, J. M. (1990). Focusing on phonological phrases in Chichewa. In *The phonology-syntax connection,* Inkelas, S. & Zec, D. (eds), Chicago: Chicago University Press, 145-161.
Karvonen, D. (2008). Explaining Nonfinality: Evidence from Finnish. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, Chang, C. B. & Haynie, H. J. (eds), Somerville, MA: Cascadilla Proceedings Project, 306-314.

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In *Phonetic linguistics, Essays in honor of Peter Ladefoged*, Fromkin , V. A. (ed), London: Academic Press, 115-132.

Kidda, M. (1993). *Tangale phonology*. Berlin: Dietrich Reimer Verlag.

Kobele, G. & Torrence, H. (2006). Intervention and focus in Asante Twi. In *ZAS Papers in Linguistics, 46,* Fiedler, I. & Schwarz, A. (eds), Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung, 161–184.

Kotei, N. A. (1969). A description of modern spoken Ga with particular reference to Tone and Intonation. Everston: Northwestern University.

Kotey, A. P. (2009). *Twi-English/English-Twi dictionary*. New York: Hippocrene books.

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

Kropp Dakubu, M. E. (1986). Downglide, Floating Tones, and Non-Wh Questions in Ga and Dangme. In *The Phonological Representation of Suprasegmentals*, Bogers, K., van der Hulst, H. & Mous, M., Dordrecht: Foris Publications, 153-74.

Kropp Dakubu, M. E. (2005). The syntax of focus in Ga and Akan and the significance of related constructions. Paper presented at the *International Conference on "Focus in African languages"* Humboldt University & ZAS Berlin.

Krifka, M. (2007). Basic notions of information structure. In *The notions of information structure, Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS), 6*, Féry, C., Fanselow, G., & Krifka, M. (eds), Universitätsverlag Potsdam, Potsdam, 13-56.

Kügler, F. & Skopeteas, S. (2006). Interaction of lexical tone and information structure in Yucatec Maya. In *Proceedings of the Second International Symposium on Tonal Aspects of Languages,* Belotel-Grenié, A. & Grenié, M. (eds), La Rochelle, 83-88.

Kügler, F. & Skopeteas, S. (2007). On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. In *Proceedings of the 16th International Congress of Phonetic Sciences*, Trouvin, J. & Barry, W. (eds), Saarbrücken, 1025-1028.

Kügler, F., Skopeteas, S. & Verhoeven, E. (2007). Encoding information structure in Yucatec Maya: On the interplay of prosody and syntax. In *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS), 8,* Ishihara, S., Jannedy, S. & Schwarz, A. (eds), Potsdam: University Press Potsdam, 187-208.

Kügler, F. (2012). Segmental and prosodic cues for phrasing in Akan. Talk held at the *Workshop of the Syntax-Phonology Interface from a Cross-linguistic perspective*, ZAS, Berlin.

Kügler, F. & Genzel, S. (2012). On the prosodic expression of pragmatic prominence: the case of pitch register lowering in Akan. *Language & Speech, 55, 3,* 331-359.

Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

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

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

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lai, C., Evanini, K. & Zechner, K. (2011). Comparative Analysis of Prosodic Features of Native and Non-native Spontaneous Speech. Poster presented at *New Tools and Methods for Very-Large-Scale Phonetics*, Philadelphia.

Laniran, Y. O. (1992). Phonetic aspects of tone realization in Igbo. *Progress Reports from Oxford Phonetics, 5*, 35-51.

Laniran, Y. O. & Gerfen, C. (1997). High raising, downstep and downdrift in Igbo. *Paper presented at the 71st annual meeting of the Linguistic Society of America*, Chicago.

Laniran, Y. O. & Clements, G. N. (2003). Downstep and High Raising: Interacting Factors in Yoruba Tone Production. *Journal of Phonetics, 31, 2*, 203-50.
Leben, W. R., Inkelas, S. & Cobler, M. (1989). Phrases and Phrase Tones in Hausa. In *Current Approaches to African Linguistics*. Newman, P. & Botne, R. (eds), Dordrecht: Foris. 45-61.
Leben, W. & Ahoua, F. (2006). Phonological reflexes of emphasis in Kwa languages of Cote d'Ivoire. In *West African Linguistics: Papers in Honor of Russell G. Schuh, Studies in African Linguistics, Supplement 11,* Newman, P. & Hyman, L. M. (eds), Columbus: Ohio State University, 145-158.
Lee, S. J. (2009). H tone, depressors and downstep in Tsonga. In *Selected Papers from the 38th Annual Conference on African Linguistics*. Matondo, M., McLaughlin, F. & Potsdam, E. (eds), Somerville, MA, 26-37.
Lefebvre, C. (1998) Multifunctionality and Variation Among Grammars: The Case of the Determiner in Haitian and in Fongbe. *Journal of Pidgin and Creole Languages*, *13*, 93-150.
Levelt, W. J. M. (2002). Picture naming and word frequency. *Language and Cognitive Processes*, *17*, 663-671.
Lewis, M. P. (ed) (2009). *Ethnologue: Languages of the world*, *Sixteenth edition*. Dallas, TX: SIL.
Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8,* 249-336.
Liberman, M. & Pierrehumbert, J. (1984). Intonational Invariance under Changes in Pitch Range and Length. In *Language Sound Structure*, Aronoff, M. & Oehrl, R.T. (eds), Cambrige, MA: MIT Press.157-234.
Liberman, M., Schultz, J. M., Hong S. & Okeke V. (1992). The phonetic interpretation of tone in Igbo. In *Proceedings of the Institute of Research in Cognitive Science Workshop on Prosody in Natural Speech.* Philadelphia: University of Pennsylvania, 83-92.
Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.
Lindau, M. (1986). Testing a model of intonation in a tone language. *Journal of the Acoustical Society of America*, 80, 757-764.
Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling*, Hardcastle, W. J. & Marchal, A. (eds), Amsterdam: Kluwer, 403-439.
Maddieson, I. (1978). Universals of Tone. In *Universals of Human Language, Volume 2: Phonology*, Greenberg, J. H. (eds), Stanford: Stanford University Press. 335-366.
Maeda, S. (1976). A characterization of American English intonation. Phd Thesis, MIT, Cambridge.
Manfredi, V. B. (1993). Spreading and downstep: prosodic government in tone languages. In *The phonology of tone. The representation of tonal register,* van der Hulst, H. & Snider, K. (eds), Berlin: Mouton de Gruyter, 133-184.
Manyah, K. A. (2006). Relation between tone and vowel quality in Twi. In *Proceedings of the Second International Symposium on Tonal Aspects of Languages,* Belotel-Grenié, A. & Grenié, M. (eds), La Rochelle: Université de La Rochelle. 37-40.
Marfo, C. O. (2003). The prosody-syntax interface in Akan focus and topic constructions. In *Proceedings of the 4th postgraduate research forum on Linguistics*, Hong Kong. 125-133.
Marfo, C. O. (2004). On tone and segmental processes in Akan phrasal words: A prosodic account. *Linguistik online, 18,* 93-110.
Marfo, C. O. (2005). *Aspects of Akan Grammar and the Phonology-syntax Interface*. Phd Thesis, University of Hong Kong.

Marfo, C. & Bodomo, A. (2005). Information structuring in Akan question-word fronting and focus constructions. *Studies in African Linguistics*, *34, 2*, 179-208.
Marfo, C. & Yankson, S. (2008). The structure of the CCV syllabus of the Akan Concentric. *Studies in Linguistics, 34*, *2,* 85-100.
Matisoff, J. A. (1973). *The Grammar of Lahu.* University of California Publications in Linguistics, 75, Berkeley / Los Angeles: University of California Press.
McDonough, J. (2002). The intonation of Navajo focus and interrogative constructions. In *Formal approaches to functional phenomena,* Carnie, A., Harley, H. & Willie, M. A. (eds), Amsterdam: Benjamins, 191-206.
Michaud, A. (2008). Tones and intonation: some current challenges. In *Proceedings of 8th International Seminar on Speech Production*, Strassbourg, 13-18.
Möhler, G. & Mayer, J. (2001). A Discourse Model for Pitch-Range Control. In *Proceedings of the 4th workshop on Speech Synthesis of the international Speech Communication Association*, Perthshire.
Mücke, D. & Grice, M. (2006). Articulatory anchors for tonal targets in rising pitch accents. Poster presented at the *10th Conference on Laboratory Phonology*, Paris.
Myers, S. P. (1996). Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African linguistics, 25, 1*, 29-60.
Myers, S. P. (2004). The effects of boundary tones on the f0 scaling of lexical tones. In *TAL 2004: International Symposium on Tonal Aspects of Language*. Bel, B. & Marlien, I. (eds), Beijing. 147-150.
Nakai, S., Kunnari, S., Turk, A., Suomi, K. & Ylitalo, R. (2009). Utterance-final lengthening and quantity in Northern Finnish. *Journal of Phonetics, 37,* 29-45.
Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
Nolan, F. (2006), Intonation. In *Handbook of English Linguistics*, Aarts, B. & McMahon, A. (eds), Oxford: Blackwell, 433-459.
Nordhoff, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. Phd Thesis. University of Amsterdam.
Obeng, S. G. (1989). *Advanced level Akan phonology: essays in memory of P.C. Sgt. D. M. Gyasi*. Legon: Department of Linguistics. Mimeograph.
Obeng, S. G. (2000). Vowel harmony and tone in Akan toponyms. *Studies in the linguistic sciences, 30, 2,* 173-184.
Odden, D. (1995). Tone: African Languages. In *The Handbook of phonology*, Goldsmith, J. (ed), Blackwell, Oxford, 444-475.
Ofori, S. A. (2011). On the Basic Focus Marker, and the Basic Focus Sentence, in Akan (Twi). *Nordic Journal of African Studies 20, 3,* 241-262.
Ohala, J. J. (1972). The physiology of tone. In *Consonant types and tone. Southern California Occasional Papers in Linguistics, 1,* Hyman, L. M. (ed.), 1-14.
Ohala, J. J. (1978). The production of tone. In *Tone: a linguistic survey*, Fromkin, V. A. (ed.), New York: Academic Press. 5-39.
Ohala, J. J. (1984). An ethological perspective on common cross - language utilization of F0 of voice*. Phonetica, 41*, 1-16.
Ohala, J. J. (1994). The frequency codes underlies the sound symbolic use of voice pitch. In *Sound symbolism*. Hinton, L., Nichols, J. & Ohala, J. J. (eds), Cambridge: Cambridge University Press, 325-347.
Ohala, J. J., Dunn, A., & Sprouse, R. (2004). Prosody and phonology. In *Proceedings of Speech Prosody 2004*, Bel, B. & Marlien, I. (eds), Nara, 161-163.
O'Keefe, M. (2003). *Akan vowel harmony*. BA Thesis, Department of Linguistics, Swarthmore College.

Pan, H. H. (2007). Focus and Taiwanese unchecked tones. In *Topic and Focus. Crosslinguistic perspectives on meaning and intonation,* (Studies in Linguistics and Philosophy), Chungming, L., Gordon, M. & Büring, D. (eds), Dordrecht: Springer, 197-216.

Parker, S. (2002). *Quantifying the Sonority Hierarchy*. Phd Thesis. UMass, Amherst.

Paster, M. (2010). The verbal morphology and phonology of Asante Twi. *Studies in African Linguistics, 39, 1*, 77-120.

Paster, M. & Kim, Y. (2011). Downstep in Tiriki. *Linguistic Discovery, 9*, 71-104.

Petrone, C., Fuchs, S. & Krivokapić, J. (2011). Consequences of working memory differences and phrasal length on pause duration and fundamental frequency. In *Proceedings of the International Seminar on Speech Production,* Laprie, Y. (ed), 393-400.

Peng, S.-H. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics*, *25*, 371-400.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Phd Thesis, Bloomington: Indiana University Linguistics Club.

Pierrehumbert, J. & Beckman, M. E. (1988). *Japanese tone structure*. Linguistic Inquiry Monograph, 15, Cambridge, MA: MIT Press.

Pierrehumbert, J. & Hirschberg, J. (1990). The Meaning of Intonational contours in the Interpretation of Discourse, In *Intentions in Communication*, Cohen, P., Morgan, J. & Pollack, M. (eds), MIT Press, Cambridge MA, 271-311.

Pierrehumbert, J. (2000). Tonal elements and their alignment. In *Prosody: Theory and experiment,* Horne, M. (ed), Dordrecht: Kluwer Academic Publishers, 11-36.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the Emergence of Linguistic Structure*, Bybee, J. L. & Hopper, P. (eds), John Benjamins, Amsterdam, 137-157.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

Pike, K. L. (1948). *Tone languages*. Ann Arbor: University of Michigan Press.

Poser, W. (1984). *The Phonetics and Phonology of Tone and Intonation in Japanese*. Phd Thesis, MIT.

Potisuk, S., Gandour, J., & Harper, M. (1997). Contextual variations in trisyllabic sequences of Thai tones. *Phonetica*, *54,* 22-42.

Prieto, P., Shih, C. & Nibert, H. (1996). Pitch Downtrend in Spanish, *Journal of Phonetics*, *24*, 445-473.

Prieto, P. (2004). The Search for Phonological Targets in the Tonal Space: Evidence from Five Sentence-Types in Peninsular Spanish. In *Laboratory Approaches to Spanish Phonology*. Face, T. (eds). Mouton de Gruyter: The Hague, 29-59.

Prieto, P., D'Imperio, M., Elordieta, G., Frota, S. & Vigário, M. (2006). Evidence for soft preplanning in tonal production: Initial scaling in Romance. In *Proceedings of Speech Prosody 2006*. Hoffmann, R. & Mixdorff, H. J. (eds). Dresden: TUDpress, 803-806.

Pulleyblank, D. G. (1986). *Tone in Lexical Phonology*, Reidel, Dordrecht.

Purvis, T. M. (2009). Speech rhythm in Akan oral praise poetry. *Text & Talk*, *29*, *2,* 201-218.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing* (Version 2.13.2) [Computer program]. Retrieved from http://www.r-project.org/.

Remijsen, B. & van Heuven, V. J. (2005). Stress, tone, and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology*, *22, 2*, 205-235.

Rialland, A. (1983). Le système tonal du moba comparé à celui du gurma. In *Current approaches to African linguistics, 2*, Kaye, J., Koopman, H., Sportiche, D. & Dugas, A. (eds), Dordrecht: Foris, 217-234.

Rialland, A. & Somé, P. A. (2000). Dagara downstep: How speakers get started. In *Advances in African linguistics. Trends in African Linguistics, 4,* Carstens, V. & Parkinson, F. (eds), Trenton, NJ: Africa World Press, 251-263.
Rialland, A. (2001). Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*, Kaji, S. (ed), Tokyo: Institute for Languages and Cultures of Asia and Africa / Tokyo University of Foreign Studies, 301-321.
Rialland, A. & Robert, S. (2001). The intonational system of Wolof. *Linguistics, 39,* 893-939.
Rialland, A. (2007). Question prosody: an African perspective. In *Tones and tunes, vol.1: Typological studies in word and sentence prosody*, Riad, T. & Gussenhoven, C. (eds), Berlin, New York: Mouton de Gruyter, 35-64.
Rialland, A. (2009). The African lax question prosody: Its Realisation and geographical distribution. *Lingua*, *119*, 928-49.
Rialland A. & Aborobongui, M. E. (2010). The Intonational system and prosodic hierarchy of Embosi. Talk held at the *International Conference on Prosodic Interfaces*, New-Delhi.
Rietveld, T., & van Heuven, V. J. (1997). *Algemene fonetiek*. Bussum: Uitgeverij Coutinho.
Rizzi, L. (1997). The fine structure of the left periphery. In *Elements of Grammar*, Haegeman, L. (ed), Dordrecht: Kluwer, 281-337.
Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics, 1, 1*, 75-116.
Saah, K. K. (1988). Wh-questions in Akan. *Journal of West African Languages*, *18*, 17–28.
Saah K. K. (1994). *Studies in Akan Syntax*. Phd Thesis. University of Ottawa, Canada.
Salffner, S. (2010). *Tone in the phonology, lexicon and grammar of Ikaan*. Phd Thesis. School of Oriental and African Studies, London.
Schachter, P. (1965). Some comments on J. M. Stewart's ‘‘The typology of the Twi tone system.", Preprint from *The Bulletin of the Institute of African Studies*, *1*, Legon, Ghana.
Schachter, P. & Fromkin, V. (1968). *A phonology of Akan: Akuapem, Asante and Fante*. Los Angeles: University of California.
Schachter, P. (1973). Focus and relativization. *Language, 49,* 19-46.
Scholz, F. (2012). Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese, Phd Thesis, Utrecht: Netherlands Graduate School of Linguistics (LOT).
Schwiertz, G. (2009). *Intonation and prosodic structure in Beaver (Athabaskan): Explorations on the language of the Danezaa*, Phd Thesis, University of Cologne.
Selkirk, E. O. (1984). *Phonology and Syntax. The Relation between Sound and Structure. Cambridge* (Mass.): MIT Press.
Selkirk, E. O. (1995). Sentence prosody: intonation, stress, and phrasing. In *The Handbook of Phonological Theory,* Goldsmith, J. A. (ed), Cambridge, MA, and Oxford, UK: Blackwell, 550-569.
Selkirk, E. O. (2007). Contrastive Focus, Givenness and the Unmarked Status of "Discourse-New". In *The notions of information structure, Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS), 6,* 125-145.
Selkirk, E. O. (2009). On Clause and Intonational Phrase in Japanese: The Syntactic Grounding of Prosodic Constituent Structure, *Gengo Kenkyu, 136,* 35-73.
Shih, C. (2000). A declination model of Mandarin Chinese. In *Intonation: Analysis, Modelling and Technology*, Botinis, A. (ed), 243-268.
Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., Féry, C. & Krifka, M. (2006). *Questionnaire on Information Structure,* (ISIS Vol. 4). Potsdam: Universitätsverlag Potsdam.
Smith, C. L. (2002). Prosodic Finality and Sentence Type in French. *Language and Speech, 45, 2,* 141-178.

Snider, K. L. (1990) Tonal upstep in Krachi: Evidence for a register tier. *Language, 66, 3*, 453-474.
Snider, K. L. & van der Hulst, H. (1993). Issues in the representation of tonal register, In *The Phonology of Tone – the representation of tonal register*, van der Hulst, H. & Snider, K. L. (eds), Berlin: Mouton de Gruyter, 1-27.
Snider, K. L. (1998). Phonetic realisation of downstep in Bimoba. *Phonology, 15, 1*, 77-101.
Snider, K. L. (1999). *The Geometry and Features of Tone*. Dallas: The Summer Institute of Linguistics & The University of Texas at Arlington.
Snider, K. L. (2007). Automatic and nonautomatic downstep in Chumburung: an instrumental comparison. *Journal of West African languages, 34*, 105-114.
Sosa, J. M. (1999). *La entonación del español*, Cátedra, Madrid.
Stewart, J. M. (1965). The typology of the Twi tone system. Preprint from *The Bulletin of the Institute of African Studies*, *1*, Legon, Ghana, 1-27.
Stewart, J. M. (1967). Tongue root position in Akan vowel harmony. *Phonetica, 16*, 185-204.
Stewart, J. M. (1970). A theory of the origin of Akan vowel harmony. In *Proceedings of the 6th International congress of phonetic sciences.* Prague, 863-865.
Stewart, J. M. (1971). Niger-Congo, Kwa. In *Linguistics in Sub-Saharan Africa, Current trends in linguistics, 7,* Sebeok, T. A., Berry, J. & Greenberg, J. (eds), Mouton, The Hague, 179-212.
Stewart, J. M. (1981). Key lowering (downstep/downglide) in Dschang. *Journal of African Languages and Linguistics*, *3*, 113–38.
Strik, H. & Boves, L. (1995). Downtrend in F0 and Psb. *Journal of Phonetics, 23,* 203-220.
Swerts, M., Strangert, E. & Heldner, M. (1996). F0 declination in spontaneous and read aloud speech. In *Proceedings of the International Conference on Spoken Language Processing, 3*, Philadelphia. 1501-1504.
Szendröi, K. (2003). A stress-based approach to the syntax of Hungarian focus. *Linguistic Review, 20*, 37-78.
Terken, J. (1993). Baseline revisited: reply to Ladd. *Language and Speech, 36,* 453-459.
Thorson, J. (2007). The scaling of utterance-initial pitch peaks in Puerto Rican Spanish: Evidence for tonal preplanning. In *University of Rochester Working Papers in the Language Sciences, 3, 1*, Wolter, L. & Thorson, J. (eds), 91-97.
Truckenbrodt, H. (1995). *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*, Phd Thesis, MIT.
Truckenbrodt, H. (2002). Upstep and embedded register levels. *Phonology, 17, 1,* 77-120.
Truckenbrodt, H. & Féry, C. (2003). More on hierarchical organization and tonal scaling. Ms, Tübingen/Potsdam, 1-20.
Truckenbrodt, H. (2004). Final lowering in non-final position. *Journal of Phonetics, 32,* 313-348.
Truckenbrodt, H. (2007). Upstep of edge tones and of nuclear accents. *Tones and Tunes. Vol. 2: Experimental Studies in Word and Sentence Prosody*, Gussenhoven C. & Riad, T. (eds), Berlin: Mouton, 349-386.
Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In *Methods in empirical prosody research,* Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N. & Schließer, J. (eds), Berlin: Mouton de Gruyter, 1-27.
Urua, E. A. (2002). The tone system of Ibibio. In *Typology of African Prosodic Systems*, Gut, U. & Gibbon, D. (eds), Bielefeld: University of Bielefeld, 65-85.
van den Berg, R., Gussenhoven, C. & Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In *Papers in Laboratory Phonology II: gesture, segment, prosody.* Docherty, G. & Ladd, D. R. (eds), Cambridge: Cambridge University Press. 335-367.

van Heuven, V. J., & Haan, J. (2000). When and how do we hear whether a Dutch speech utterance of a statement or a declarative question? In *Linguistics in the Netherlands 2000*, *17*, Hoop, H., de & van der Wouden, T. (eds), 93-104.
van Heuven, V. J. (2004). Planning in speech melody: Production and perception of downstep in Dutch. In *On Speech and Languages: Studies for Sieb G. Nooteboom*, Quené, H. & van Heuven, V. J. (eds), Utrecht: Netherlands Graduate School of Linguistics (LOT), 83-93.
Vaissière, J. (1983). Language-independent prosodic features. In *Prosody: Models and Measurements*, Cutler, A. & Ladd, D. R. (eds), Berlin. Springer Verlag, 53-66.
Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 323-340.
Wang, B. & Xu, Y. (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics, 39*, 595-611.
Weber, A., Grice, M. & Crocker, M. W. (2005). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eyes movements. *Cognition, 99, 2,* 1-10.
Welmers, W. E. (1959). Tonemics, morphotonemics, and tonal morphemes. *General Linguistics*, *4*, 1-9.
Welmers, W. E. (1973). *African language structures*. Berkeley: University of California Press.
Westermann, D. & Ward, I. C. (1990). *Practical Phonetics for Students of African Languages.* London: Kegan Paul International.
Wiesemann, U. (1972). *Die phonologische und grammatische struktur der Kaingáng-Sprache.* The Hague: Mouton.
Xu, Y. (1993). *Contextual variation in Mandarin Chinese*, Phd Thesis, University of Connecticut.
Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*, 61-83.
Xu, Y. & Wang, Q. E. (1997). What can tone studies tell us about intonation? In *Proceedings of the European Speech Communication Association Workshop Intonation: Theory, Models and Applications,* Botinis A., Kouroupetroglou, G., Carayannis, G. (eds), Athens, Greece. 337-340.
Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, *27, 1*, 55-105.
Xu, Y. & Sun X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America, 111*, 1399-1413.
Yip, M. (1980). The tonal phonology of Chinese. Phd Thesis, MIT.
Yip, M. (2002). *Tone*. Cambridge University Press.
Yu, K. M. (2009). Contextual tonal variation in level tone languages. Poster presented at the *Acoustical Society of America*, Portland, OR.
Yuan, J. & Liberman, M. (2010). F0 declination in English and Mandarin broadcast news speech. In *Proceedings of Interspeech 2010*, 134-137.
Zee, E. (1978). Duration and intensity as correlates of F0. *Journal of Phonetics, 6*, 213-220.
Zerbian, S. (2004). Phonological Phrases in Xhosa (Southern Bantu), *ZAS Papers in Linguistics, 37,* 71-99.
Zerbian, S. (2006). Expression of Information Structure in Northern Sotho. Phd Thesis, Humboldt University, Berlin.
Zerbian, S. (2007). The Subject/Object-Asymmetry in Northern Sotho. In *Information Structure and the Architecture of Grammar: A Typological Perspective*. Schwabe, K. & Winkler, S. (eds), Amsterdam: John Benjamins, 323-345.
Zerbian, S. (2010). Recent developments in the typology of intonation, *Linguistics and Language Compass, 4, 9,* 874-889.

Zerbian, S., Genzel, S., & Kügler, F. (2010). Experimental work on prosodically-marked information structure in selected African languages (Afroasiatic and Niger-Congo). *Workshop on Experimental Approaches to Focus*, Speech Prosody 2010. Chicago.

Zerbian, S. & Kügler, F. (2012). Sequences of high tones across word boundaries: downstep and phrasing in Tswana, Talk held at the *Workshop on the Phonology-Syntax Interface*, ZAS Berlin.

Zhenglai, G., Hiroki, M. & Hideki. K. (2003). Analysis of vowel formant frequency variations between focus and neutral speech in Mandarin Chinese, *Acoustical Science and Technology, 4, 4,* 192–193.

Zimmermann, M. (2007). Contrastive Focus and Emphasis. In *The Notions of Information Structure Interdisciplinary Studies on Information Structure (ISIS)*, *6*, Féry, C., Fanselow, G. & Krifka, M., Potsdam: Potsdam University. 147-159.

List of figures

figure 1: Model of Lexical Phonology; reproduced from Gussenhoven (2004:58) with slight modifications (replacement of rules with processes). 7
figure 2: Left side: L toned proper name (*Yaw*) taken from the initial position of the sentence *Yaw fi Akyemfo.* 'Yaw comes from Akyemfo.'; Right side: H toned proper name (*Esi*) taken from the initial position of the sentence *Esi yi lɛta*. 'Esi takes a letter away.'; female speaker. 11
figure 3: Sentence with alternating HL tones; uttered by the female speaker. 17
figure 4: Sentence involving non-automatic downstep; female speaker. 19
figure 5: Mandarin Chinese sentence containing only H tones (solid line), Mandarin Chinese sentence with alternating H and L tones (dotted line); reproduced from Wang & Xu (2011:604). 20
figure 6: Sentence illustrating the 'ceiling effect'; female speaker. 21
figure 7: Left pitch track Ikaan sentences with only L tones Dʒɛ̀jȁnìhjɛ̀hjɛ̀g 'I bought sponges.'; Right pitch track sentences with alternating LH tones Dʒɛ̀r̥ȃn ɔ̀wóg àrákpà. 'I made bean soup.'; reproduced from Salffner (2010:137, 143); male speaker. 23
figure 8: Embosi sentence with superimposed high boundary tone (H%); reproduced from Rialland & Aborobongui (2010:45). 24
figure 9: Left hand side, pitch track of a sentence with LH toned subject NP and perfective aspect uttered at normal speech rate; Right hand side, pitch track of a sentence with LH toned subject NP and perfective aspect uttered at low speech rate; male speaker. 31
figure 10: Left hand side, pitch track of a sentence with L toned subject NP and perfective aspect uttered at normal speech rate; Right hand side, pitch track of a sentence with L toned subject NP and perfective aspect uttered at slow speech rate; male speaker. 32
figure 11: Left hand side, pitch track of a sentence with LH toned subject and mono-syllabic verb in progressive aspect, uttered at moderate speech rate; Right hand side, pitch track of a sentence with H toned subject and di-syllabic verb in progressive aspect, uttered at moderate speech rate; male speaker. 33
figure 12: Pitch track of a sentence with progressive aspect and fronted subject NP, uttered at moderate speech rate; male speaker. 37
figure 13: Pitch track of a sentence with perfective aspect and fronted object NP, Esi na Yaw abo no. 'It is Esi who Yaw has beaten.' uttered at moderate speech rate; male speaker. 38
figure 14: Pitch track of a sentence with past tense and fronted object NP, *Mango na Anum tɔɔ no anɔpa yi.* 'It is mango that Anum bought this morning.', spontaneously uttered; female speaker. 39
figure 15: Surface representation for the sentence *Ane-no akai se'etaa-wa do'ko desu ka?* 'Where is big sister`s sweater?', reproduced from Pierrehumbert & Beckman (1988:21). 43
figure 16: Left hand side, pitch track of a sentence with only L tones, *Yaw fi Akyemfo.* 'Yaw is from Akyemfo.'; Right hand side, pitch track of a sentence with only H tones, *Aba twa serɛ.* 'Aba cuts grass.'; female speaker. 52
figure 17: L raising and downstep in Yoruba in sentences consisting of alternating HL and alternating LH tones; reproduced from Laniran & Clements (2003:213). 73
figure 18: Mean F0, aggregated over repetitions, of sentences consisting of alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba.* 'Papa Kofi talks to his child.'; Mean F0, aggregated over repetitions, of sentences containing only L tones (solid line), *Wɔfa Asare fi Akyemfo.* 'Asare come from Akyemfo.'; female speaker. 77
figure 19: Mean F0, aggregated over repetitions, in sentences consisting of alternating LH (solid line), *Papa Kofi kasa kyerɛ ne ba.* 'Papa Kofi talks to his child.'; Mean F0,

aggregated over repetitions, in sentences consisting of alternating HL tones (dashed line), *Anane bisa sika ɛnɛ.* ‘Anane asks for money.’, female speaker. 78
figure 20: Left boxplot: F0 of L2 in sentences with only L tones and L1 in sentences with alternating HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. Right boxplot: F0 of L3 in sentences with only L tones and L2 in sentences with alternating LH tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. 79
figure 21: Boxplot F0 of L1 and L2, in sentences with alternating LH tones, aggregated over repetitions and lengths, split by speakers and positions. 80
figure 22: Yoruba sentence containing only H tones (white circles), Yoruba sentence containing only L tones (white triangles), Yoruba sentences with alternating L and H tones (black circles), female speaker; reproduced from Laniran & Clements (2003:217). 81
figure 23: F0 of sentences with HLHM (white triangles) and HLHL (black circles) tones, male speaker; reproduced from Laniran & Clements (2003:232). 82
figure 24: H raising and downstep in Akan. Mean F0, aggregated over repetitions, of sentences with alternating HL tones (dashed line), *Anane bisa.* ‘Anane asks.’; Mean F0, aggregated over repetitions, of sentences containing only H tones (solid line), *Kukuoba no.* ‘The small pot.’; female speaker. 85
figure 25: Boxplot F0 of H1, in sentences with only H tones and in sentences with alternating LH/HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. 86
figure 26: Boxplot F0 of H3 & H4, in sentences with only H tones and of H2 in sentences with alternating LH/HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. 87
figure 27: Boxplot F0 of H1, in sentences with only H tones, alternating LH tones, initial HHHL tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration. 88
figure 28: Boxplot F0 of H3, in sentences with only H tones, initial HHHL tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration. 89
figure 29: Boxplot F0 of H4, in sentences with only H tones and initial HHHHL tones, aggregated over repetitions and split by speakers and tonal configuration. 90
figure 30: Pitch track of the word *Anane*, uttered by the female speaker. 91
figure 31: Automatic downstep (black solid line) and non-automatic downstep (grey solid line), aggregated over speakers (5); based on Dolphyne (1994:23). 96
figure 32: Mean F0 of sentences, aggregated over speakers and repetitions, exhibiting automatic downstep (solid line), *Afua hunu Kofi papa anɔpa yi.* ‘Afua saw Kofi’s father this morning.’; Mean F0 of sentences, aggregated over speakers and repetitions, exhibiting non-automatic downstep (dashed line), *Afua hunu Kofi dan anɔpa yi.* ‘Afua saw Kofi’s house this morning.’; n=6. 97
figure 33: Left boxplot: F0 of H3 in sentences containing downstep vs. F0 of !H3 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. Right boxplot: F0 of H3 in sentences containing downstep vs. F0 of H4 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. 98
figure 34: Left boxplot: F0 drop between H2-H3 in sentences containing downstep vs. F0 drop between H2-!H3 in sentences containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. Right boxplot: F0 drop between H2-H3 in sentences containing downstep vs. F0 drop between H2-H4 in sentences

containing non-automatic downstep, aggregated over repetitions and split by downstep type and speakers. 99
figure 35: Mean F0, aggregated over repetitions, of sentences consisting of alternating LH (solid line), *Papa Kofi kasa.* 'Father Kofi talks.'; Mean F0, aggregated over repetitions, of sentences consisting of alternating HL tones (dashed line), *Anane bisa.* 'Anane asks.'; female speaker. 105
figure 36: Boxplot F0 of H1, in sentences with alternating LH tones and HL tones, aggregated over repetitions and lengths, split by speakers and tonal configuration. 105
figure 37: F0 of a Hausa sentence containing only H tones, *Muudii yaa zoo gidaa.* 'Muudi came home.' (originally from Lindau, 1986:761); reproduced from Connell (2011:838). 107
figure 38: Pitch range of the primary registers h and l for the male speaker; reproduced from Möhler & Mayer, (2001). 110
figure 39: Mean F0 of sentences containing only H tones differing in length, aggregated over repetitions; short sentences (solid line), *Kukuoba no.* 'The small pot.', medium sentences (dotted line), *Kukuoba papa no.* 'The good small pot.', long sentences (dashed & dotted line), *Kukuoba papa no bɔ daa.* 'The good small pot breaks everyday.', longest sentences (dashed line), *Kukuoba papa paa no bɔ daa.* 'The very good small pot breaks everyday.'; female speaker. 113
figure 40: Mean F0 of sentences containing only L tones differing in length, aggregated over repetitions; short sentences (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akyemfo.', medium sentences (dotted line), *Asare fi Akyemfo.* 'Asare comes from Akyemfo.', long sentences (dashed & dotted line) *Wɔfa Asare fi Akyemfo.* 'Uncle Asare comes from Akyemfo.', longest sentences (dashed line), *Wɔfa Ado Asare fi Akyemfo.* 'Uncle Ado Asare comes from Akyemfo.'= 12 syllables; female speaker. 114
figure 41: Left boxplot: declination rate for sentences with only H tones, aggregated over repetitions, split by speakers and utterance lengths (short = 5σ's, medium = 7σ's, long = 10σ's, longest = 12σ's). Right boxplot: declination rate for sentences with only L tones, aggregated over repetitions and split by speakers and utterance lengths (short = 7σ's, medium = 8σ's, long = 10σ's, longest = 12σ's). 115
figure 42: Interaction of tone, length (short – dotted and dashed line, medium - dotted line, long – dashed line, longest – solid line) and declination rate. 116
figure 43: Left pitch track: statement consisting of only H tones, *Aba twa serɛ.* 'Aba cuts grass.'; Right pitch track: Yes – No question consisting of only H tones, *Aba twa serɛ?* 'Aba cuts grass?'; female speaker. 117
figure 44: Obtained mean F0, aggregated over repetitions, of short sentences containing only L tones (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' and predicted F0 values (dashed line); female speaker. 121
figure 45: Obtained mean F0, aggregated over repetitions, of short sentences containing only L tones (solid line), *Yaw fi Akyemfo.* 'Yaw comes from Akymenfo.' and predicted F0 values (dashed line); male speaker 3. 122
figure 46: Obtained mean F0, aggregated over repetitions, of short sentences containing only H tones (solid line), *Kukuoba no.* 'The small pot.' and predicted F0 values (dashed line); male speaker 5. 123
figure 47: Left hand side, F0 of a Hausa sentence containing only H tones, *Muudii yaa zoo gidaa.* 'Muudii came home.'; (originally from Lindau, 1986:761); reproduced from Connell (2011:838); Right hand side, F0 of a Hausa sentence containing alternating HL tones, *Maalam yaa auni leemoo.* 'The teacher weighed the oranges.'; (originally from Lindau, 1986:758); reproduced from Connell (2002a:7). 127

figure 48: Mean F0 of L2, L4 and L6, aggregated over speakers and repetitions, of the medium sentences containing only L tones (solid line), *Asare fi Akyemfo.* 'Asare comes from Akyemfo.'; Mean F0 of L1, L2 and L3, aggregated over speakers and repetitions, of the medium sentences with alternating HL tones (dashed line), *Anane bisa sika*. 'Anane asks for money.'; n=6. 131
figure 49: Mean F0 of L1, L3, L5, L7, L9 and L12, aggregated over speakers and repetitions, of the longest sentences containing only L tones (solid line), *Wɔfa Ado Asare fi Akyemfo.* 'Unlce Ado Asare comes from Akyemfo.'; Mean F0 of L1, L2, L3, L4, L5 and L6, aggregated over speakers and repetitions, of sentences with alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba bio.* 'Father Kofi asks for his child again.'; n=6. 131
figure 50: Mean F0 of H1, H3, H5 and H7, aggregated over speakers and repetitions, of the medium sentences containing only H tones (solid line), *Kukuoba papa no.* 'The good small pot.'; Mean F0 of H1, H2, H3 and H4, aggregated over speakers and repetitions, of the medium sentences with alternating HL tones (dashed line), *Anane bisa sika.* 'Anane asks for money.'; n=5. 132
figure 51: Mean F0 of H2, H4, H6, H8, H10 and H12 aggregated over speakers and repetitions, of the longest sentences containing only H tones (solid line), *Kukuoba papa paa no bɔ daa.* 'The very good small pot breaks everyday.'; Mean F0 of H1, H2, H3, H4, H5 and H6 aggregated over speakers and repetitions, of the longest sentences with alternating LH tones (dashed line), *Papa Kofi kasa kyerɛ ne ba bio.* 'Father Kofi talks to his child again.'; n=5. 133
figure 52: Boxplot of pitch level of L tones, aggregated over repetitions, lengths and tonal configurations, split by positions (1 = early)/(2 = late), and speaker. 134
figure 53: Boxplot pitch level of L tones in long sentences, aggregated over repetitions and positions, split by tonal configuration (L = only L)/(LH= alternating) and speakers. ... 135
figure 54: Boxplot of pitch drop of L tones, aggregated over repetitions and lengths, split by tonal configuration (L = only L)/(LH/HL = alternating tones) and speakers. 135
figure 55: Boxplot of pitch drop of L tones, aggregated over repetitions and tonal configuration, split by utterance lengths (medium = 7σ's, long = 10σ's, longest = 12σ's) and speaker. 136
figure 56: Boxplot of pitch level of H tones, aggregated over repetitions, lengths and tonal configurations, split by positions (1 = early)/(2 = late) and speakers. 137
figure 57: Boxplot of pitch level of H tones in long sentences, aggregated over repetitions and positions, split by tonal configuration (H = only H)/(LH= alternating) and speakers. .. 138
figure 58: Left boxplot: pitch drop of H tones, aggregated over repetitions and lengths (medium & long), split by tonal configuration (H = only H)/(LH/HL = alternating tones) and speakers; Right boxplot: pitch drop of H tones for the longest sentences. 139
figure 59: Left boxplot: pitch drop of H tones in sentences with only H tones, aggregated over repetitions, split by length (medium = 7σ's, long = 10σ's, longest = 12σ's) and speakers. Right boxplot: pitch drop of H tones in sentences with alternating LH/HL tones. 140
figure 60: Obtained F0 values (solid line) of the medium sentences containing alternating HL tones, *Anane bisa sika*. 'Anane asks for money.', aggregated over repetitions, and predicted F0 values (dashed line); female speaker. 146
figure 61: Obtained F0 values of the longest sentence with alternating HL tones (solid line), *Anane bisa sika bɔne ɛnɛ*. 'Anane asks for money today.', aggregated over repetitions and predicted F0 values (dashed line); male speaker. 147
figure 62: Obtained F0 values (solid line), aggregated over repetition and predicted F0 values, for the medium sentence with alternating LH tones, *Papa Kofi kasa*. 'Papa Kofi talks.' and predicted F0 values (dashed line); male speaker. 148

figure 63: Left plot obtained F0 values (solid line), aggregated over repetitions, *Afua hunu Kofi papa anɔpa yi.* 'Afua sees Kofi's father this morning.' and predicted F0 values (dashed line) for the sentences with automatic downstep; Right plot obtained F0 values (solid line), aggregated over repetitions, *Afua hunu Kofi dan anɔpa yi.* 'Afua sees Kofi's house this morning.' and predicted F0 values (dashed line) for the sentences with non-automatic downstep, female speaker. 149
figure 64: F0 values of the initial H tone, aggregated over repetitions and split by speakers and downstep type, DD refers to automatic downstep and DS refers to non-automatic downstep. 154
figure 65: F0 maxima (st) on the first subject peak in Wenzhou Chinese, broken down by constituent length, averaged across speakers (19). T-bars = ±2 SE; reproduced from Scholz (2012:134). 156
figure 66: Left plot: Mean F0, aggregated over repetitions, for short (solid line), *Kukuoba no*. 'The small pot.' and long sentences, *Kukuoba papa no bɔ daa.* 'The good small pot breaks everyday.' (dashed line) consisting of only H tones; Right plot: Mean F0, aggregated over repetitions, for short, *Yaw fi Akyemfo*. 'Yaw comes from Akyemfo.' and long sentences, *Wɔfa Asare fi Akyemfo*. 'Uncle Asare comes from Akyemfo.' consisting of only L tones; female speaker. 160
figure 67: Left plot: Mean F0, aggregated over repetitions, for short (solid line), *Anane bisa*. 'Anane asks.' and long sentences (dashed line), *Anane bisa sika ɛnɛ*. 'Anane asks for money today.', consisting of alternating HL tones; Right plot: mean F0, aggregated over repetitions, for short, *Papa Kofi kasa*. 'Father Kofi talks.' and long sentences, *Papa Kofi kasa kyerɛ ne ba*. 'Father Kofi talks to his child.', consisting of alternating LH tones; female speaker. 160
figure 68: Mean F0 values of the initial H tone in sentences with only H tones, aggregated over repetitions and split by speakers and length group. 161
figure 69: Mean F0 values of the initial L tone in sentences with only L tones, aggregated over repetitions and split by speakers and length group. 161
figure 70: Left boxplot: F0 values of the initial H tone in sentences with alternating HL tones, aggregated over repetitions and split by speakers and length group; Right boxplot: F0 values of the initial L tone in sentences with alternating HL tones, aggregated over repetitions and split by speakers and length group. 162
figure 71: Left boxplot: F0 of the initial L tone in sentences with alternating LH tones, aggregated over repetitions and split by speakers and length group; Right boxplot: F0 of the initial H tone in sentences with alternating LH tones, aggregated over repetitions and split by speakers and length group. 163
figure 72: Mean F0 values, aggregated over speakers and repetitions, of complex sentences with complementizer clauses containing only L tones; short complementizer clause (solid line), *Nana kaa ɛnora sɛ Yaw fi Akyemfo.* 'Nana said yesterday that Yaw comes from Akyemfo.', medium complementizer clause (dashed line), *Nana kaa ɛnora sɛ Asare fi Akyemfo.* 'Nana said yesterday that Asare comes from Akyemfo.' and long complementizer clause (gazed line), *Nana kaa ɛnora sɛ wɔfa Asare fi Akyemfo.* 'Nana said yesterday that uncle Asare comes from Akyemfo.',; n=6. 165
figure 73: Mean F0 values, aggregated over speakers and repetitions, of complex sentences complementizer clauses containing only H tones; short complementizer clause (solid line), *Nana kaa ɛnora sɛ kukuoba bɔ.* 'Nana said yesterday that a small pot breaks.', medium complementizer clause (dashed line), *Nana kaa ɛnora sɛ kukuoba papa bɔ.* 'Nana said yesterday that a good small pot breaks.' and long complementizer clause (gazed line), *Nana kaa ɛnora sɛ kukuoba papa no bɔ daa.* 'Nana said yesterday that the good small pot breaks everyday.'; n=6. 165

figure 74: Left boxplot: F0 of the initial L tone of the matrix clause, aggregated over repetitions and split by speakers and lengths of the complementizer clause; Right boxplot: F0 of the initial H tone of the matrix clause, aggregated over repetitions and split by speakers and lengths of the complementizer clause. 166
figure 75: Left boxplot: F0 of the initial L tone of the complementizer clause containing only L tones, aggregated over repetitions and split by speakers and lengths; Right boxplot: F0 of the initial H tone of the complementizer clause containing only H tones, aggregated over repetitions and split by speakers and lengths. 167
figure 76: Left plot: mean F0 (10 points of measurement per vowel), aggregated over speakers and repetitions, for the complementizer (COMP) and first H tone (H1) of the short (solid line), medium (dashed line) and long (gazed line) complementizer clauses containing only H tones; Right plot: mean F0, aggregated over speakers and repetitions, for the complementizer (COMP) and first L tone (L1) of the short (solid line), medium (dashed line) and long (gazed line) complementizer clauses containing only L tones; n=6. 171
figure 77: Global (b-c) and local (d) variations in pitch range: a. neutral realization (baseline) of an HLH sequence in a hypothetical tone language; b. higher register (H and L raising); c. bottomline raising (L raising); d. span expansion (H raising and L lowering), d'. bottomline lowering (L lowering), d''. topline raising (H raising), e. raising of the last H, e'. raising of the Hs in the final phrase. 175
figure 78: Time normalized course of F0 for the sentences: *Mwamuna amalamula amayi*./? 'The man bosses around the woman./?'; utterance final; one phrase; reproduced from Myers (1996:33f.) 176
figure 79: Mean F0, aggregated over speakers and repetitions, of the sentence *Papa Kofi kasa*./? 'Father Kofi talks./?', statement solid line, Yes – No question dashed line*, T* terminal F0; n=6. 183
figure 80: Mean F0, aggregated over speakers and repetitions, taken from the last vowel of the sentence *Papa Kofi kasa*./? 'Father Kofi talks./?' 7 points of measurement: F0 extracted at the first 30 ms, mid of vowel and last 30 ms, statement solid line, Yes – No question dashed line*,* n=6 185
figure 81: Mean F0, aggregated over speakers and repetitions, taken from the last vowel of the sentence Papa Kofi kasa kyerɛ ne ba bio./? 'Father Kofi talks to his child again./?' 7 points of measurement: F0 extracted at the first 30 ms, mid of vowel and last 30 ms, statement solid line, Yes – No question dashed line*,* n=6. 186
figure 82: Mean duration (V-F2), aggregated over items, speakers and repetitions, of the final vowel measured from the beginning of phonation (V) till the end of F2 and from the end of F2 till the end of phonation (BT), BT stands for breathy termination; the Yes – No questions are represented by the white bar and the statements by the black bar; n=17. 188
figure 83: Mean vowel formants (F1, F2, F3) of the last vowel, aggregated over speakers, items and repetitions, split by sentence type and vowel quality; /a/ left panel; /o/, /ɔ/ right panel; F1 black quadrate, F2 grey quadrate, F3 light grey quadrate; n=17.................... 190
figure 84: Mean duration, aggregated over speakers and repetitions, measurement for /a/ taken from the final vowel of the statement/Yes- No question *Papa Kofi kasa*./? 'Papa Kofi talks./?'; measurement for /aa/ taken from the final vowel of the statement *Kukuoba papa no bɔ daa*. 'The good small pot breaks everyday.'; n=6. 194
figure 85: Time normalized course of F0, aggregated over speakers and repetitions, of the sentence *Papa Kofi kasa*./? 'Father Kofi talks./?', Yes – No question grey dotted line; statement black dotted line and *Papa Kofi kasa kyerɛ ne ba*./? 'Papa Kofi talks to his child./?'; Yes – No question grey solid line; statement black solid line; n=6. 197
figure 86: Scatterplot displaying F0 (st) and intensity (db) values, measured at the mid of the initial vowel of the Yes – No questions *Papa Kofi kasa*? 'Father Kofi talks?', *Papa Kofi*

kasa kyerɛ ne ba? ‘Father Kofi talks to his child?’ and *Papa Kofi kasa kyerɛ ne ba bio*? ‘Father Kofi talks to his child again?’; n = 41 ... 199
figure 87: Pictures for the description of a single situation task. Target sentence left picture: *Agyeman boaa Ado a-nɔpa yi*. ‘Agyeman helped Ado this morning.’. Target sentence right picture: *Anum tɔɔ amango anɔpa yi*. ‘Anum bought mango(s) this morning.’. ... 202
figure 88: Left pitch track: course of F0 for the sentence *Me baa ha*. ‘I came here.’; Right pitch track: course of F0 of the sentence *Me na me baa ha*. ‘It is I who came here’.; female speaker. ... 208
figure 89: Time-normalized course of F0, aggregated over speakers, for *amango* ‘mango’ uttered under wide informational (solid line), narrow informational (dotted line) and narrow corrective focus (dashed line), taken from the sentence *Anum tɔɔ amango anɔpa yi*. ‘Anum bought a mango this morning.’; n=11 (Kügler & Genzel, 2012:344) ... 210
figure 90: Time-normalized course of F0, aggregated over speakers, for the target word *Ado* uttered under wide informational (solid line), narrow informational (dotted line) and narrow corrective focus (dashed line), taken from the sentence *Agyeman boaa Ado anɔpa yi.* ‘Agyeman helped Ado this morning.’; n=11 (Kügler & Genzel, 2012:346). ... 211
figure 91: Time-normalized course of F0, aggregated over speakers, for *amango* ‘mango’ uttered in-situ under wide (solid black line), narrow informational (dotted black line) and narrow corrective focus (dashed black line), taken from the sentence *Anum tɔɔ amango anɔpa yi*. ‘Anum bought a mango this morning.’ and uttered ex-situ under wide (solid grey line), narrow informational (dotted grey line) and narrow corrective focus (dashed grey line), taken from the sentences *Amango ate firi dua no so anɔpa yi*. ‘A mango has fallen down the tree this morning.’ and *Amango na Anum tɔɔ anɔpa yi*. ‘It is a mango that Anum bought this morning.’; n=11 (Kügler & Genzel, 2012:347). ... 212
figure 92: Time-normalized course of F0, aggregated over speakers, for the word *amango* ‘mango’ uttered in wide focus (black solid line) and with the word occurring pre-focally (black dotted line); taken from the sentence *Anum tɔɔ amango anɔpa yi*. ‘Anum bought a mango this morning.’; n=11 (Kügler & Genzel, 2012:349) ... 214
figure 93: Time-normalized course of F0, aggregated over speakers, for the sentence *Afua hunuu wɔma anɔpa yi*. ‘Afua saw a pestle this morning.’ under wide focus (black solid line) and under corrective focus (black dashed line); n=5 (Kügler & Genzel, 2012:352). ... 215
figure 94: Spectrogram of a SVOAdv sentence uttered under wide informational focus; female speaker (11). ... 218
figure 95: SVOAdv sentence uttered under wide informational focus; female speaker (11). 219
figure 96: SVO$_F$Adv sentence with narrow informational focus on the object; female speaker (11). ... 219
figure 97: Waveform of a SVO$_F$Adv sentence uttered under narrow informational focus; female speaker (11). ... 220
figure 98: SVO$_F$Adv sentence with narrow corrective focus on the object; female speaker (11). ... 220
figure 99: Model of Lexical Phonology with information about lexical tones, post-lexical processes and phonetic processes of Akan. ... 233
figure 100: Operation mode of the pitch implementation algorithm in Akan. H_i & L_i are initial reference values, *s* is the lowering quotient and *r* is the baseline quotient. ... 234

List of tables

table 1: Vowels of Akan and their phonetic description. 28
table 2: Summary of the main research entities; ordered by the presumed domain of the entity and their presumed place in the grammar. 60
table 3: F0 values for L tones in Hz and differences in pitch drop (Δ), L1-L2, L2-L3, L3-L4, aggregated over speakers (5), for sentences exhibiting downstep; reproduced from Dolphyne (1994:8). 74
table 4: Material used to test L raising; method 1. 75
table 5: Material used to test L raising; method 2. 76
table 6: Material used to test H raising. 83
table 7: F0 values for H tones in Hz and differences in pitch drop (Δ), H1-H2, H2-H3, H3-H4, aggregated over speakers (5), for sentences exhibiting automatic downstep and non-automatic downstep; based on Dolphyne (1994:5f.). 95
table 8: Material used to test initial downstep. 104
table 9: Interaction of number/type of tonal contrasts and declination. 108
table 10: Material used to test declination. 112
table 11: Reference values *r* and *s*; mean initial F0 (T1), aggregated over repetitions, for short, medium, long and longest sentences containing only L tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6. 120
table 12: Reference values *r* and *s*; mean initial F0 (T1), aggregated over repetitions, for short, medium, long and longest sentences containing only H tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6. 122
table 13: Material used to test the difference between declination & downstep. 129
table 14: Mean F0 values of the selected L tones plus standard deviation and mean pitch drop between the selected L tones, aggregated over speakers and repetitions for sentences with only L tones and sentences with alternating (LH/HL) tones; n=6. 133
table 15: Mean F0 values for the selected H tones plus standard deviation and mean pitch drop between the selected H tones, aggregated over speakers and repetitions for sentences with only H tones and sentences with alternating (LH/HL) tones; n=5. 136
table 16: Reference values *r* and *s*; mean initial F0 (H1 & L1), aggregated over repetitions, for short, medium, long and longest sentences with alternating HL tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6. 145
table 17: Reference values *r* and *s*; mean initial F0 (L1 & H1), aggregated over repetitions, for short, long and longest sentences with alternating LH tones and Pearson's R^2 for obtained vs. predicted F0 value; n=6. 147
table 18: Reference values *r* and *s*; mean initial F0 (L1 & H1), aggregated over repetitions, for sentences containing automatic and non-automatic downstep and Pearson's R^2 for obtained vs. predicted F0 value; n=6. 149
table 19: Material for testing anticipatory raising, data set 1. 157
table 20: Material for testing anticipatory raising, data set 2. 158
table 21: Summary of prosodic cues, found in the literature, marking Yes – No questions in Akan. 180
table 22: Initial F0 values, aggregated over speakers and repetitions and split by sexes, tones and sentence type. 184
table 23: Pitch drop for items 4, 5 & 7, aggregated over speakers and repetitions and split by positions, sentence type and sexes; PoD refers to position of the drop (1 = H1-H2; 2 = H2-H3; 3 = H3-H4; 4 = H4-H5). 185
table 24: Final mean F0 values, aggregated over speakers and repetitions, split by items, sexes, tones and sentence type. 186

table 25: Mean ΔL values, aggregated over speakers and repetitions, split by positions, sexes and sentence type. Position 1 refers to the antepenultimate value, position 2 to the penultimate value and position 3 to the ultimate value. 187
table 26: Mean duration of the final vowel, aggregated over speakers and repetitions, split by items, tones and sentence types. 189
table 27: Mean intensity, aggregated over speakers and repetitions, split by items, tones and sentence type; Maximal intensity of the utterance (MaxIntU), maximal intensity of the final vowel (MaxIntfinV) and the difference between the two (Int_{rel}). 189
table 28: Mean formant values (F1, F2, F3), aggregated over speakers and repetitions, split by items, vowels, tones and sentence type. 191
table 29: Material used to test the prosodic marking of focus. 209
table 30: Mean F0 values in Hz for the initial H tone (H1), for the target word (TW) and difference (Δ), aggregated over speakers, split by focus conditions; n=11. 215

Statement of Authorship/Selbstständigkeitserklärung

I hereby solemnly affirm that this thesis was written by myself and describes my own work, unless otherwise acknowledged in the text. The thesis has not been submitted for the award of any other degree in any other tertiary institution.

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ferner erkläre ich, dass die Arbeit noch nicht in einem anderen Studiengang als Prüfungsleistung verwendet wurde.

Susanne Genzel