Tone in Saxwe

Beavon Ham, V.R.

Author: Beavon-Ham, V.R.
Title: Tone in Saxwe
Issue Date: 2019-11-06

7 The phonetic implementation of tone

This chapter examines the specifics of how tone is realized by four Saxwe speakers of similar age and linguistic background. While there are many similarities in the phonetic output of these four speakers, there are also some notable differences, particularly in utterances that push the limits of what might occur in natural speech—such as in lengthy iterative H–M and H–L sequences. Observations are summarized in each section and are discussed more globally in the concluding section of this chapter.

The following is the outline of this chapter. Section 7.1 gives a summary of the methods and instruments used to obtain and analyze the data discussed in this chapter. In section 7.2, baselines for all-H, all-M, and all-L utterances are established. Section 7.3 examines the question of whether there is iterative automatic and non-automatic downstep of H, as well as whether there is iterative automatic downstep of L. In the course of answering this question, we observe the anticipatory raising of H before successive L–H sequences. Section 7.4 looks at the phonetic implementation of L tones in successive L–M sequences and demonstrates the fact that there is no non-automatic lowering of L tone. In the course of examining this question, we also see evidence of anticipatory lowering of L before H.

Turning from multi-word utterances to individual words, section 7.5 looks at the phonetic realization of the most common tone patterns of V.C(C)V nouns. Finally, section 7.6 concludes with a summary of the details of the phonetic implementation of tone in Saxwe.

7.1 Methodology

The data discussed in this chapter were recorded in the Houeyogbe township of the country of Benin in May 2017. These data collection sessions followed a significant amount of analysis that had been done on recordings made previously in 2015 (section 1.3). Based on information I had about the underlying tone of words from data notes and a database in Fieldworks Language Explorer (FLEx) (SIL International, 2011), a preliminary list of the groups of words and sentences to be recorded was prepared before the recording sessions.

In the Houeyogbe township, I worked with Patrice Videgnon and Godefroy Sossou, both of whom had had some basic linguistic training, to finalize the list of words and sentences so that they would reflect the particular sequences of tones desired. In each section of this chapter, the words or sentences tested are described in more detail.

The recordings were done on a Marantz PMD 660 solid state recorder using an external Shure SM10A headworn, unidirectional dynamic microphone. The recordings were done in a cement-walled room with a front door open and a quiet fan running in the background. The recorder was set at a 44,100 Hz sampling rate.

The first recordings were made of André Taïve, a 43-year old male from Adrome. Taïve (hereafter speaker AT) had considerable experience in radio recording, and had worked with me on previous recordings done in 2015. Following this, recordings were done with three other individuals—Nicolas Gbemasse, a 45-year old male from Tohon; Kouessi Sossou, a 51-year old male from Kpovidji; and Béatrice Lokossou, a 45-year old female from Houeyogbe (hereafter speakers NG, KS, and BL). In all, there were three males and one female. These individuals were selected specifically because they were mature adults who were literate in Saxwe, and they had grown up and spent most of their adult years in the Houeyogbe township.

For all of the recordings which were not of himself, speaker AT explained the process and addressed any issues that arose. I operated the recording equipment in all cases.

The process for recording was the following. For each topic of testing, the words or sentences prepared for that topic were printed on individual pieces of paper. These papers were manually shuffled and handed to the speaker. Speakers were instructed to first read to themselves what was on the paper and be sure they understood it, and then to read the word or sentence at a normal speaking rate to be recorded. The speakers did this for each piece of paper in the stack, moving at their own speed from one to the next. If there were questions regarding the meaning of the sentence or word, these were addressed by speaker AT.[109]

In the case of the V.CV nouns, the speakers were instructed to repeat each word twice and there was a single pass through the words. Sentences, however, were read once in each pass, but were reshuffled and re-recorded nine times to make a total of ten tokens per sentence. In all, each speaker produced two repetitions each of 60 different V.CV nouns and ten repetitions each of 50 different sentences. The recording sessions were divided up over the course of two days for each person.

During the recording sessions, if the speaker clearly misread a sentence (as signaled by speaker AT, who was present during all recordings), the recording of that sentence was repeated. Later in the analysis of the sentences, some tokens were not kept because they included reading errors such as the repetition, correction, omission or addition of a word. This meant that for every test sentence spoken by a single speaker there was a maximum of ten tokens, but sometimes there were fewer than ten tokens that were able to be used for analysis. In the examination of the data

[109] I myself answered such questions in the recordings of speaker AT.

in this chapter, I note those cases where the full number of ten tokens was unavailable for analysis.

For every token that was retained, the vowels were manually segmented in Praat (Boersma & Weenink, 2015) based on visual observation of the acoustic waveform and the spectrographic analysis. The end boundary of an utterance-final vowel was placed at the point where the waveform ceased to have a distinctive repeating shape (Baart, 2010).

All vowels were labeled as Praat text files with the underlying tone associated to that TBU. For the V.C(C)V nouns, syllable onsets were labeled as containing either a sonorant, voiced obstruent, or voiceless obstruent onset. The recordings and segmented text files for all the data used in this study can be found at: https://drive.google.com/open?id=1viq0KzW2UEj_uflpW6DYBQ5VJ6PAwysg.

The measurement of pitch F_0 was generated automatically in Praat using two different scripts—one developed for the V.(C)CV nouns and one for the sentences. The scripts used were based on adaptations from other scripts found primarily in Boersma (2014), and secondarily in Kawahara (2014), improved through suggestions found in Styler (2015). Matthew Lee of SIL International and Jos Pacilly of Leiden University helped to adapt and improve these scripts for the purposes of this study.

The effect on the F_0 of a vowel of a preceding consonant in tonal languages is shown to be dissipated after the first 60 ms (Hombert, 1977). However, there can also be a perturbation on the tone of a vowel as it approaches the transition to the following consonant. For example, a voiced obstruent can produce a 'dip' in F_0 that is felt not only on the following vowel but also on the preceding one (Connell & Ladd, 1990). In a visual examination of the data, I found that the point of leveling or 'shoulder' of the F_0 of a level tone occurred most consistently at a position which was not at the midpoint of the vowel, but slightly later in the vowel. For all of these reasons, it was decided that for the measurement of tone within the vowels of the sentences, the Praat script would take a measurement at the time index which was at 66 percent into the duration of the vowel segment. However, in order to avoid too much loss of data in cases where Praat was unable to get a pitch value at that exact time index, the script included a loop which searched for the nearest time index at which a pitch reading could be made, and which recorded the pitch value at this alternative time index, making a note of the difference in time between the 66 percent time index and this new time index. All such alternative time indexes were manually inspected to make sure that the difference in timing was minimal.

For V.C(C)V nouns, the Praat script was adapted to allow a more extensive analysis of F_0 values throughout the duration of the vowel. To this end, pitch values were taken at ten points equally spaced throughout the time duration of each of the vowels.

After F_0 readings were generated using these scripts, the data were verified visually to check for missing or clearly erroneous readings. In several cases where

F_0 readings were proving to be difficult to obtain with consistency, the voicing threshhold in the analysis settings was changed from the 0.45 value used by default to a lower value of 0.25. This was done for a number of the sentences read by speaker KS and NG, who both tended to have some creakiness in their production of certain vowels. Also in several cases, the default low value of the F_0 range was increased from 75 Hz in order to force the pitch analysis calculations in Praat to favor a more reasonable value over one that was clearly erroneous. This was most often necessary for speaker NG. Any measurements from Praat that were still clearly aberrant based on visual and audio inspection were disregarded in statistical calculations and confidence intervals were adjusted accordingly. This is described in the following sections in each case where such adjustments were made.

7.2 Baseline utterances for all-H, all-M, and all-L

7.2.1 Research question and recorded utterances

The first goal was to establish baseline F_0 trends for utterances composed of multiple iterations of a single surface tone. The following is the research question that was proposed: for Saxwe speakers of similar origin and background, how do all-H, all-M, and all-L utterances differ from each other with respect to F_0?

To answer this question, the following six sentences were created and recorded. For each of the three tone options: all-H, all-M, and all-L, there was one sentence of eight syllables and one sentence of ten syllables created. Each sentence was recorded ten times total by each speaker, and the order of sentences was reshuffled between passes.

(422)	/$^{\text{M-}}$ kájí	tɔ́	lá	tʃã́	vǎ	kú	fí/
	Kayi	father	DEF	also	come	die	now
	Kayi's father also finally died recently. (8 surface Hs)						

(423)	/$^{\text{M-}}$ tĩ́kṹsĩ́sjɛ̃́	lá	tʃã́	vǎ	bú	sésé/
	fruit	DEF	also	come	be.lost	completely
	The fruit also was eventually completely lost. (10 surface Hs)					

(424)	/ājānɔ̃̄	mɛ̃̄	nã̄	nyã̄	āwū/
	poor.person	that	FUT	wash	shirt
	That poor person will do laundry. (8 surface Ms)				

(425)	/ɛ̄mɛ̃̄xīxɔ̄	mɛ̃̄	nã̄	nɔ̃̄	ɖ̤ū	ājā/
	servant	that	FUT	HAB	eat	suffering
	That servant will (habitually) suffer. (10 surface Ms)					

(426)

(426)	/$^{M-}$ glàgò	gàgà	gbȉgbɔ̏	mɔ̏	dɾɔ̏/
	Glago	tall	return	REPET	be.remaining

Tall Glago's return is still expected. (8 surface Ls)

(427)	/$^{M-}$ gȁdʒà	glòbòtòtò	mɔ̏	hȕ	gbàjà/
	trap	big.and.round	REPET	open	wide.open

A big round trap again opens wide open. (10 surface Ls)

The ideal for baseline testing would be that there be a total absence of any surface or underlying tone other than the one being focused on. The surface forms of these sentences were indeed all-H, all-M or all-L. However, if we look at underlying forms, we see that they were not strictly all-H, all-M or all-L—partially out of necessity and partially due to error. This is explained below.

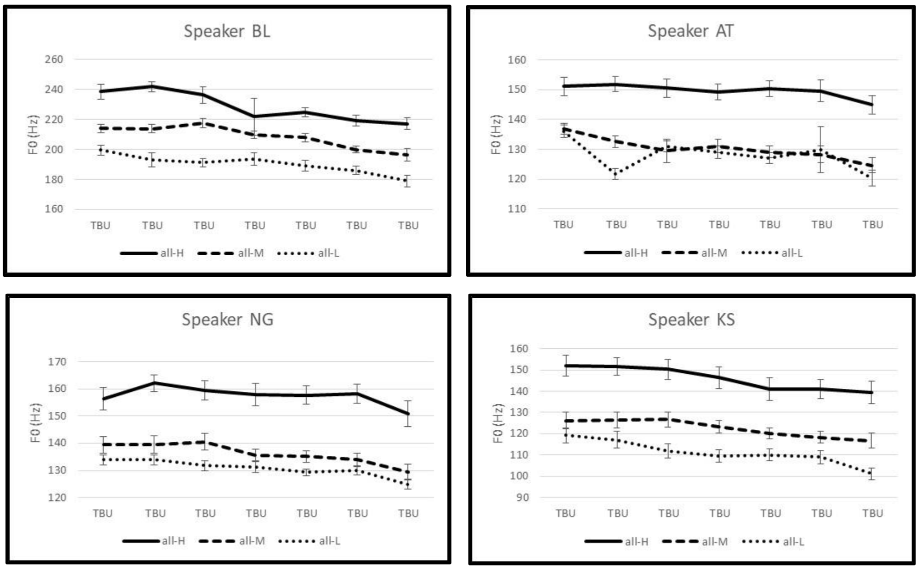
From section 4.3, we recall that PWs which function as the head of a noun phrase have a left M- floating tone if they do not have an initial vowel (all of which together are likely to be vestiges of a historic class-marking system). This presents a challenge in developing sentences of all-H and all-L tones. At the beginning of the all-H and all-L sentences, the left edge M- floating tone is present on the subject of the sentence out of necessity, as the first element in the utterance. However, no further nouns were incorporated in these sentences so as to avoid any further non-automatic downstep triggered by the presence of this left edge M- floating tone (see section 4.3). This deviation from the ideal on the subject noun was unavoidable.

An error was involved in the introduction of the word /vǎ/ 'come', used as a verbal auxiliary in sentences (422) and (423). As described in the rule of Contour simplification A (section 3.6.4), when a TBU has an underlying LH contour, this is simplified following a H by deleting the L. This results in a surface H realization, which is why this discrepancy in underlying tones was not noticed at the time of the finalization of the test sentences. In future tests, it would be preferable not to have the inclusion of this underlying tonal pattern in test sentences used for establishing an all-H baseline.

In Saxwe, there are utterance-final interactions with the right edge $L_{\%}$ IP boundary (section 3.5) that can have a significant effect on the surface realization of the final TBU of an utterance. It has also been documented that H tone can display utterance-final lowering (Herman, 1996). For these reasons, the F_0 measurements of utterance-final TBUs were not retained for any analysis which had to do with establishing baseline levels for tones. For the graphs in section 7.2.2, only the measurements from the first seven TBUs of sentences (422) to (427) were retained for graphing and statistical analysis. These were averaged together, giving a maximum count of 20 tokens per speaker (10 from the sentence of eight TBUs and 10 from the sentence of ten TBUs) for each TBU of an all-H, all-M, or all-L sentence.

7.2.2 Results

The following show results for each of the speakers. Speaker BL is female, while the other three speakers are male. The horizontal graphs show the mean F_0 for each of the first seven TBUs in the all-H, all-M, and all-L utterances. The vertical bars show the upper and lower limits of a 95% confidence interval calculated using a T distribution from the measurements taken at each TBU.[110]



(428) Baseline levels for all-H, all-M, and all-L utterances (avg. 20 tokens)

Rather than normalizing the data from each speaker in an effort to describe a single pattern of implementation for H, M and L in Saxwe, I believe it is more useful to look at the similarities and differences in these speakers' implementation of H, M and L in Saxwe.

---

[110] As noted in the description of the methodology, recordings in which speakers made a reading error were not kept for analysis. Therefore in the following cases, fewer than 20 tokens were available for analysis: (1) speaker BL—the all-L data represent 19 tokens; (2) speaker KS—the all-H data represent 17 tokens and the all-M data represent 19 tokens; (3) speaker NG—the all-H data represent 18 tokens and the all-M data represent 19 tokens; (4) speaker AT—the all-M data represent 16 tokens. These differences are taken into account in the calculation of 95% confidence intervals. In addition, speaker NG consistently omitted the /vǎ/ morpheme in the first all-H sentence with the result that this sentence was seven syllables long rather than eight; this meant that the measurements from the final TBU of the utterance were exceptionally included for his ten tokens of this sentence.

First, I look at similarities. Note first that in all of the utterances, there is overall evidence of declination. In all cases, when a linear regression equation is applied to the data, the slope is negative. There is no generalized pattern, however, of the declination for any one of the three tones being more or less significant than that for the other tones. The following are the slopes of a linear regression line fit to each of the three tone graphs for each speaker: (1) speaker BL—H: -4.3, M: -3.2, L: -2.8; (2) speaker KS—H: -2.4, M: -1.8, L: -2.6; (3) speaker NG—H: -0.9, M: -1.7, L: -1.3; (4) speaker AT—H: -0.8, M: -1.6, L: -1.2.

Another similarity for three of the speakers is that the highest mean F_0 of production of H occurs not on the first TBU of the all-H utterance, but rather on the second. This is true for all speakers except speaker KS. This can be described as peak delay, a known phenomenon of phonetic implementation whereby the F_0 peak may occur on the syllable following the one to which a tone is lexically associated (Xu, 2001). Here, the transition from voicelessness preceding the utterance to the production of a multiple-H sequence results in some speakers 'sliding' up, so to speak, to the target of highest F_0, with the alignment for the peak finally occurring on the second syllable rather than on the first. This peak delay on an utterance-initial sequence of Hs is seen also in section 7.3.4.

The most noteworthy difference between speakers is in the relative distribution of the phonetic heights of H, M, and L within the F_0 range of the individual. Speaker BL is the only speaker who evenly distributes H, M, and L within her F_0 range and who does not have any overlap in the 95% confidence intervals for M and L measurements. We can say that speaker BL is the speaker who shows the least probability of the true value of her M being the same as that of her L at any point in the utterance.

Speakers NG and KS both have a F_0 target for H which is clearly distinct from the targets for M and L. For both speakers, the F_0 targets for M and L are closely spaced within the lower part of their F_0 range. For speaker KS, the 95% confidence intervals for M and L at the first TBU of the utterance are just touching; throughout the rest of the utterances there is more distance between the confidence intervals for M and L. For speaker NG, there is touching or slight overlap of the confidence intervals for M and L at multiple points throughout the utterances—at all but the third and fifth TBUs. Thus for speakers NG and KS, there is still a fairly strong indication that there is a difference between the true F_0 targets for M and L, albeit a relatively small one.

For speaker AT, there is considerable overlap of the 95% confidence intervals for the TBUs all along the lengths of the all-L and all-M utterances. This means that for speaker AT, there is no clear evidence of there being a F_0 target for M different from that for L in single-tone utterances (although we see in section 7.5 that underlying M and L TBUs are phonetically distinct for speaker AT in utterances which are not limited to a single tone). The exception to this is at the second TBU, where there is a single L whose F_0 dips visibly with respect to the general trend for L. This dip on the second L of two utterance-initial Ls seems to be his means of phonetically distinguishing a series of utterance-initial Ls from a series of utterance-initial Ms despite the fact that apart from this he has no clear difference in F_0 targets for M and L.

7.2.3 Discussion of results

The way speakers implement baseline all-H, all-M, and all-L utterances is best understood in light of the two features that lay behind the atomic tones H, M, and L. As seen in section 6.2.3, M and L are both [-upper], in contrast with H, which is [+upper].

(429)		**Underlying**	**After application of default rules**
	H	[+upper]	[+upper, +raised]
	M	[+raised]	[-upper, +raised]
	L	[-raised]	[-upper, -raised]

In the phonetic implementation, the distinction between the higher register and the lower register (between the values [+upper] and [-upper]) seems to be more salient for speakers than the distinction between [+raised] and [-raised] within the register [-upper]. Thus the distinction between H versus M/L is clearly established for all speakers at the level of the phonetic implementation, whereas the distinction between M and L is more variably realized depending on the speaker.

Speaker BL distributes all the three combinations of features quite evenly within her F_0 range in single-tone utterances. Speakers KS and NG appear to divide their F_0 ranges into equal parts for the values [+upper] and [-upper] and then subdivide [-upper] into the F_0 targets for the values [+raised] and [-raised]. And Speaker AT appears to make a single distinction between [+upper] and [-upper] within single-tone utterances, overlapping the values for [+/-raised] within the same F_0 layer in this context.

In all cases, what is *not* observed is the hypothetical situation where the [+/-raised] alternation would be clearly distinguished in the phonetic implementation at the expense of clearly making a difference between the [+/-upper] values.

7.3 Iterative automatic and non-automatic downstep of H

7.3.1 Definition of downstep in terms of pitch observations

A crucial issue for tests attempting to demonstrate the phenomenon of downstep in a language is the question of what defines downstep as distinct from other pitch lowering phenomena. Connell (2011) gives the following definition.

> An important, indeed defining, feature of downstep, in addition to its lowering of a H relative to a preceding H (or lowering of other tones relative to preceding tones of like phonological value) via a L (either surface or floating) that conditions the lowering is that, within specifiable bounds, the downstepped H sets a new ceiling for all subsequent Hs within a specifiable domain; *i.e.* these Hs do not rise above the height of the downstepped one, hence the descriptive label terracing… A further characteristic of downstep, it will be remembered, is its cumulative nature: successive downsteps result in successively lower pitch levels. (pp. 838-839)

It is a tricky issue to define downstep of H in a language where declination is always present and anticipatory raising of H is a reality—both of which are true in Saxwe. When we speak of lowering, we must specify what the lowering is in reference to. There are three possibilities: lowering relative to a preceding H, lowering relative to a baseline measurement from an all-H utterance, and lowering at a rate that exceeds the lowering of declination. All three types of lowering are observed among the four Saxwe speakers tested here, although not necessarily to the same degree by individual speakers. In addition, when we speak of successively lower levels, it is helpful to specify how many additional levels beyond the first must be observed in order to qualify as downstep.

Having considered all the challenges in assigning labels to lowering phenomena, I recognize here that my own labels will be subjective. I try, however, to be clear about the criteria I use so that comparisons can be made with pitch lowering observations from other languages. The following are the criteria I use in assigning or not assigning the label of "downstep".

First, there must be a minimum of two steps down, each representing a decline from the preceding H that exceeds the rate of declination. Because there is anticipatory raising of H, the lowering need not descend below the baseline for an all-H utterance. However, if it does not descend below this baseline, I label it as "delimited downstep", and if it does not continue within the prosodic domain further than the minimum of two steps down, I label it as "arrested downstep", borrowing from Rialland (2001). In a case of arrested downstep, there can still be, due to declination, progressive lowering of H. However, the rate of lowering will no longer exceed the rate of declination. Alternatively, there may be an upward reset of the level of H after the two steps of lowering. A speaker may implement downstep which is both delimited and arrested, but these two do not necessarily co-occur.

7.3.2 Research questions and recorded utterances

Automatic downstep of H in Saxwe is the lowering of the level of H triggered by a surface L, whereas non-automatic downstep of H is the lowering of the level of H triggered by a floating M. This floating M can be present because of the synchronic elision of a vowel (section 4.2), because it is part of an underlying tonal pattern (section 3.7.4), because of vestigial effects from loss of noun class marking (section 4.3), or because of the way the word has been incorporated into the language through borrowing (section 4.5). In the most common cases, however, a floating M is present because a preceding H or L tone has spread onto an underlying M TBU, causing it to be delinked (section 3.2). This latter situation is what we see in the test utterances of this section.

The following four research questions describe the information related to automatic and non-automatic downstep which is sought after: (1) Is the lowering of the F_0 of H which is attributed to a surface L between Hs greater than that which could be attributed to declination? (2) Is the lowering of the F_0 of L when it alternates with H greater than that which could be attributed to declination? (3) Is the lowering of the F_0 of H which is attributed to a floating M between Hs greater than that which could be attributed to declination? (4) Is the lowering of the F_0 of H which is attributed to a surface L between Hs equal to that which is attributed to a floating M between Hs?

In order to answer these questions, the following set of eight sentences was created and recorded. Each sentence was recorded ten times by each speaker, and sentences were reshuffled between each pass.

(430) /ᴹ⁻ télà sɔ̃́/
The tailor left. (3 TBUs - /HLH/)

(431) /ᴹ⁻ télà xé mɔ̃̀ sɔ̃́/
This tailor left again. (5 TBUs - /HLHLH/)

(432) /ᴹ⁻ télà xé mɔ̃̀ tú vɔ̀ ké/
This tailor again finished paying. (7 TBUs - /HLHLHLH/)

(433) /ᴹ⁻ télà ᴹ⁻ sɛ̃́gbàtɔ́ mɔ̃̀ kpɔ̃́ ᴹ⁻ mĩ̀ ké/ [111]
The lawbreaking tailor saw you (PL) again. (9 TBUs - /HLHLHLHLH/)

[111] It would have been preferable not to have the initial floating M- tone on /ᴹ⁻ sɛ̃́gbàtɔ́/ and /ᴹ⁻ mĩ̀/, but the reality is that any noun which is used in a test sentence will have this initial floating M- tone if there is no initial vowel. All observations and tests indicate that a floating M word-initially or between a L and a H tone has no effect on the surface output.

(434) /ᴹ⁻ tʃītʃā sɔ̋/
The teacher left. (3 TBUs - /HMH/)

(435) /ᴹ⁻ tʃītʃā lá nā̋ sɔ̋/
That teacher will leave. (5 TBUs - /HMHMH/)

(436) /ᴹ⁻ tʃītʃā lá nā̋ kpɔ̋ ōtī́/
That teacher will see a tree. (7 TBUs - /HMHMHMH/)

(437) /ᴹ⁻ tʃītʃā lá nā̋ kpɔ̋ ōtī́ ātū̋/
That teacher will see five trees. (9 TBUs - /HMHMHMHMH/)

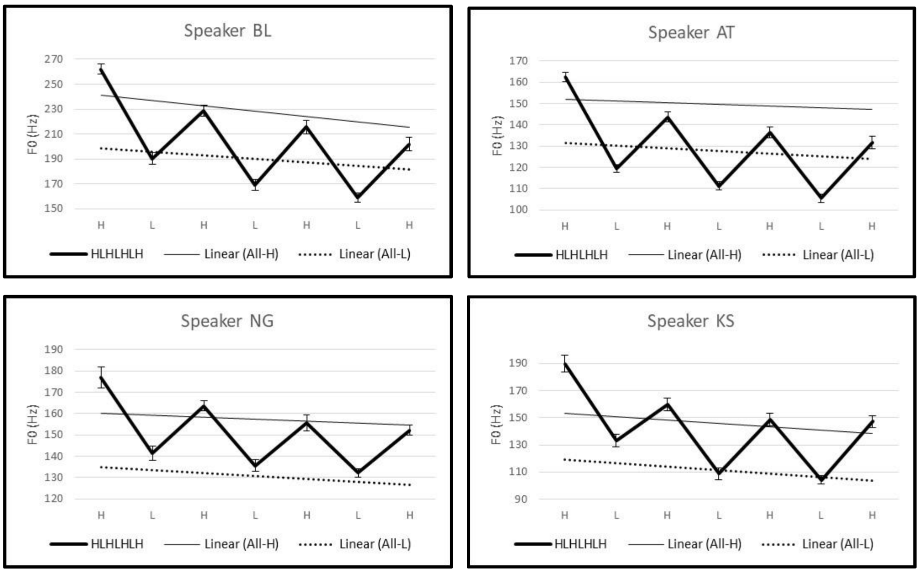
The sections below describes how the measurement obtained from these recordings were used.

7.3.3 Results: automatic downstep of H in alternating H–L sequences

This section deals specifically with the results pertaining to the lowering of H and L in alternating H-L sequences, and answers the first two research questions proposed in section 7.3: (1) Is the lowering of the F_0 of H which is attributed to a surface L between Hs greater than that which could be attributed to declination? (2) Is the lowering of the F_0 of L when it alternates with H greater than that which could be attributed to declination?

The data used to demonstrate the phonetic realization of alternating surface H and L tones come from the 10 repetitions of both the 7-TBU sentence (432) and the 9-TBU sentence (433). The F_0 measurements at each TBU (20 measurements in all) are averaged together to give a mean.[112] A 95% confidence interval is calculated using a T distribution, shown in these graphs by the vertical bars. In addition, a linear regression line is generated from the baseline data for all-H and all-L utterances seen in section 7.2.2. This is added to the graphs of alternating H and L surface tones.

[112] Occasionally a F_0 generated by Praat had to be excluded because it showed clear discrepancies (by being far outside of the realm of all other readings or of the pitch range of a given speaker). The 95% confidence intervals reflect this difference in number of readings.



(438) Iterative H–L sequences over 7 TBUs (20 tokens each)

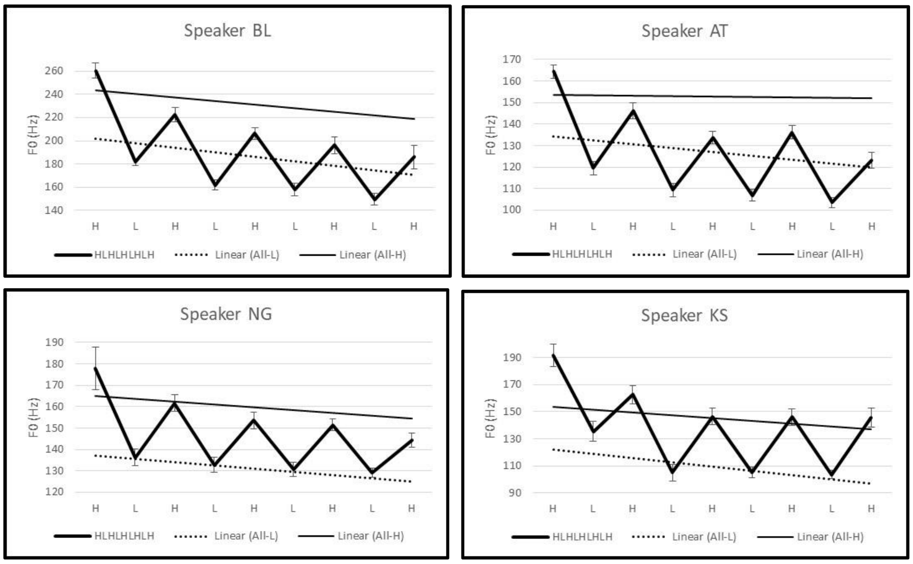
First, we consider whether the lowering of the F_0 of H attributed to a surface L between Hs is greater than that which could be attributed to declination. Clearly for all four speakers the lowering of F_0 from the first H (at the first TBU) to the second H (at the third TBU) is a greater decrease in terms of Hz than the corresponding declination seen for the all-H sentences between the first and third TBUs. What is also clear for all speakers is that this difference is achieved in large part because of the anticipatory raising of the initial H above the baseline level of H seen in all-H utterances. This strategy has been described for other languages (Rialland, 2001) and is in Saxwe a significant reason for the general auditory perception that there is automatic downstep of H.

For speaker AT, there is a smaller but recognizable second instance of lowering between the second and the third H (corresponding to the third and fifth TBUs). For this speaker, the confidence interval for the second H of the alternating utterance is slightly below the baseline for H, whereas the confidence interval for the third H of the alternating utterance is well below the baseline for H—in fact, midway between the baselines for all-H and all-L utterances.

For speakers NG and KS, there also seems to be a second instance of lowering of H beyond that which could be attributable to declination since the confidence interval bars for the second H are above the all-H baseline and the confidence interval bars for the third H are overlapping with the all-H baseline. For speaker BL, it is unclear whether there is a statistically significant second instance of lowering of H beyond that which could be attributable to declination.

The strongest conclusion that can be made from the graphs in (438) is that all of these Saxwe speakers show clear evidence of there being statistically significant lowering of H between the first and second H in an utterance composed of alternating H and L surface tones, achieved in large part through the anticipatory raising of H utterance-initially. After this, there seems to be for some speakers a second instance of lowering, smaller in Hz than the first. Following this, speakers use what Rialland (2001) terms a cancellation strategy; any subsequent lowering of the F_0 of H no longer exceeds that which can be attributed to declination. Recall from section 7.3.1 that according to the definition of downstep used here, there must be two steps down which both exceed the rate of declination in order for lowering to receive the label of downstep. If, after the two steps, the lowering does not exceed the rate of declination, the downstep is said to be arrested.

We can get another view on the matter by looking solely at the data coming only from the utterances of alternating H and L tones over 9 TBUs. For the graphs below, the F_0 measurements at each TBU are calculated from 10 repetitions of the 9-TBU sentence (433) only. A 95% confidence interval and a linear regression line for all-H and all-L utterances is again included. Here, the linear regression lines are calculated only from the all-H and all-L utterances that are 10 TBUs in length ((423) and (427)). This helps to ensure that we are comparing declination and downstep in sentences of roughly equivalent length.



(439) Iterative H–L sequences over 9 TBUs (10 tokens each)

Here we see again that from the initial raised level of H, there is clearly a first, statistically more significant lowering of H, followed by a second, less significant lowering. The second instance of lowering is clearly apparent for speakers AT and KS, and less clearly discernible for speakers BL and NG. After this, the lowering of Hs in the alternating H and L utterance does not significantly exceed the lowering that can be attributed to declination; the rate at which the Hs lower after this point produces a slope of no steeper incline than the rate at which Hs lower in an all-H utterance. Thus automatic downstep of H, if it is present for two steps, is arrested after these two first steps. In fact, for speakers AT and KS, there are some upward tendencies for H (a slight reset of H) after the second downstep.

Referring to the definition established for downstep in section 7.3.1, we can say that the automatic lowering of H of speakers AT and KS can be labeled as arrested automatic downstep and that for speakers BL and NG, it is not entirely clear whether there are two steps of lowering of which both decline at a rate that exceeds the rate of declination. If the pattern seen for speakers BL and NG is not labeled as downstep, we could describe it as an initial localized raising of H followed by a single instance of subsequent lowering of H, proceeded afterwards by lowering that is consistent with declination.

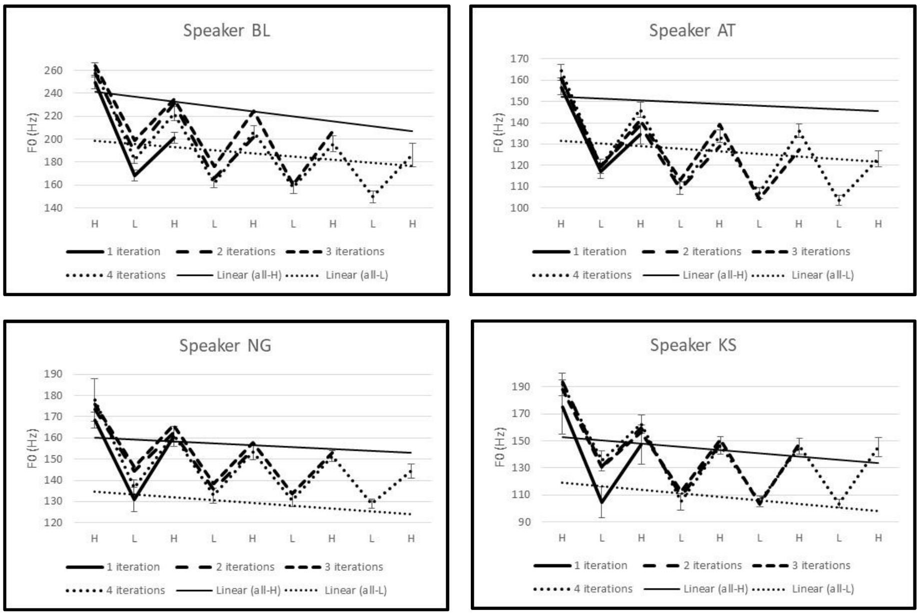
Briefly before moving on to answer the second research question, I highlight here a distinction that one can see in comparing the graphs of speakers BL and AT to the graphs of speakers NG and KS. For speakers NG and KS, the baseline for each tone correlates roughly with the lower limit of the F_0 range of that tone. A H tone can be produced above the baseline for H (in anticipatory raising) but not significantly below it, and a L can be produced above the baseline for L but not significantly below it. For speakers BL and AT, the baselines for H and L do not correlate with the lower limits of the F_0 range of these tones. A H can be realized both above and below the H baseline. Similarly, a L can be realized below the all-L baseline. (The data for speakers BL and AT do not answer the question whether a L can be realized above the L baseline.) This presence or absence of correlation between the baseline and the lower limit of F_0 range of a tone is seen again in the discussion of non-automatic downstep of H in section 7.3.4.

As stated in section 7.3.1, downstep which does not bring the level of H below the baseline level of H in all-H utterances is given the label "delimited" downstep. Therefore we can say that the downstep produced by speakers AT and KS is arrested downstep, and that additionally, the downstep produced by speaker KS is delimited downstep.

The graphs in (438) and (439) also address the second research question, which is whether the lowering of the F_0 of L when it alternates with H (automatic downstep of L) is greater than that which could be attributed to declination. For speakers BL, AT, and KS (the latter most clearly), there is one instance of lowering between the first and second L in the alternating utterance which exceeds the lowering of Ls in the all-L baseline. Following this initial downward trend for L, any further lowering of L does not exceed the lowering that can be attributed to declination, as indicated by the slope of the baseline all-L regression line. For speaker NG, there is a steady lowering of Ls in the alternating utterance which parallels the lowering attributable to declination in the all-L regression line.

For all four speakers, there is at most a single instance of lowering of L that would exceed the lowering that is attributable to declination. If we apply the same criteria to downstep of L as we do to downstep of H, we can conclude that none of these speakers produce automatic downstep of L. No speaker produces two steps of lowering of L which exceed the rate of decline attributable to declination. There is instead, for some speakers, a single localized instance of lowering of L from the first level of L found at the beginning of the utterance.

The question arises whether this initial single instance of lowering (for speakers BL, AT, and KS) is made possible by an anticipatory raising of the initial L in an utterance, similar to the anticipatory raising of an utterance-initial H. The graphs below show increasing numbers of iterations of Hs separated by L, with the data coming from recorded sentences (430) to (433). Bars indicating 95% confidence intervals are indicated only for the HLH utterances and the HLHLHLHLH utterances.



(440) Utterances of increasing iterations of Hs separated by L (10 tokens each)

In these graphs, we see that for three speakers (all except AT), the L in a short HLH utterance is lower in F_0 than the initial L in longer utterances of alternating H and L tones. This suggests that there could be a relationship between the anticipatory raising of the the initial L in longer utterances and the single instance of lowering of L that is observed for some speakers. It would be useful to explore this relationship further in studies involving more speakers and a larger data set.

To summarize the results in this section, we see that in utterances of alternating H and L surface tones, Saxwe speakers will implement one step of automatic lowering of H from a previously raised H, sometimes followed by a second, smaller step of lowering of H. Thus for some speakers, there is automatic downstep of H. Following this, any further lowering of H can generally be attributed to declination.

Another parameter related to downstep is how Hs and Ls in an utterance of alternating tones are realized in relation to all-H and all-L baselines. In longer utterances of alternating Hs and Ls, some speakers will permit Hs found late in the utterance to drop well below the baseline of an all-H utterance, and some speakers will not. Those who do not are labeled as having delimited downstep. The same speakers who permit Hs to drop well below the baseline of an all-H utterance also permit Ls to drop below the baseline of an all-L utterance.

According to the stated criteria for assigning the label of downstep (along with its sub-categories), the following statements can be made about these speakers' production of downstep: (1) speaker BL—weak evidence of arrested automatic downstep of H; (2) speaker AT—stronger evidence of arrested automatic downstep of H; (3) speaker NG—no clear evidence of downstep of H after the first step of lowering; and (4) speaker KS—some evidence of delimited, arrested automatic downstep of H. These initial conclusions are based on visual comparison of the descriptive statistics from multiple sets of data, taking into account confidence intervals. Further testing with larger data sets would allow more conclusive statements to be made.

We can also draw some conclusions about the lowering of L in alternating H and L surface tones. For speakers who do demonstrate any lowering of L in these utterances of alternating H and L, only one step of lowering of L is ever observed. Therefore, generalizing from these four speakers, I draw the conclusion that in Saxwe, there is no automatic downstep of L.

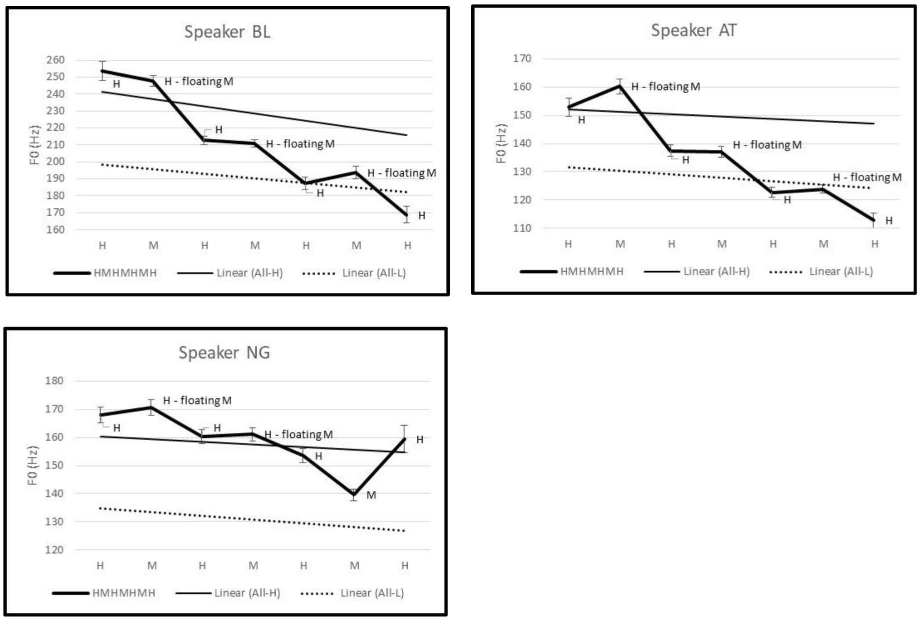
7.3.4 Results: the lowering of H triggered by a floating M

I turn now to the topic of the lowering of H triggered by a floating M, answering the question (3) from section 7.5: Is the lowering of the F_0 of H which is attributed to a floating M between Hs greater than that which could be attributed to declination?

In order to demonstrate the phonetic realization of utterances of alternating underlying H and M tones, F_0 data were obtained from 10 repetitions each of the 7-TBU sentence (436) and the 9-TBU sentence (437). This gave 20 measurements for each TBU, which were averaged to obtain a mean.[113] A 95% confidence interval was calculated using a T distribution, shown in these graphs by the vertical bars. In addition, a linear regression line was generated from the baseline data for all-H and all-L utterances seen in section 7.2.2.

In section 3.2, the claim is made that in Saxwe, H tone spreads to an underlying M TBU, delinking the M. This floating M between surface Hs triggers a lowering of the level of H. Thus an underlying /H–H–H/ sequence will be realized [H–H–H], whereas a /H–M–H/ sequence will be realized [H–H–ꜜH]. This lowering is phonologically contrastive as it is the means by which the underlying distinction between a M and a H can be recovered when these two tones occur between Hs.

The following graphs show the phonetic implementation of alternating underlying H and M tones for three of the four speakers. (The fourth is discussed separately.) Note that in these graphs, the underlying tones are marked on the x-axis and the output from the phonology is marked on the graphs.



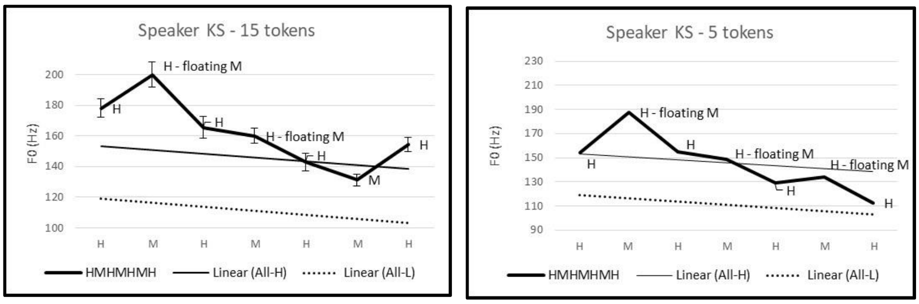
(441) Iterative underlying H–M sequences over 7 TBUs (20 tokens each)

[113] Occasionally, Praat generated a clearly erroneous measurement which was far out of the range of other measurements in the utterance. When this happened, the measurement was not included in the mean and the calculation of the 95% confidence interval took into account this difference in number of tokens.

In these graphs, we see that all speakers begin with a surface height for H which has been raised in anticipation of the lowering that will occur later in the utterance. As discussed in section 7.2, due to peak delay, the highest F_0 measurement of H occurs for some speakers on the second surface H TBU in the utterance rather than on the utterance-initial H. Here we see that happening for speakers AT and NG.

Speakers BL and AT have iterative lowering of the level of H which continues well below the baseline level for H and even below the baseline level for L in all-L utterances. In these graphs, we see three downsteps produced by these speakers. Speaker NG, however, limits the number of downsteps to two. Following this, downstep is arrested in a cancellation strategy (Rialland, 2001). For this speaker, Tonal spread no longer operates after two downsteps and he produces instead a surface M followed by a surface H which is reset above the level of the previous H. This allows speaker NG to avoid lowering the level of H significantly below the baseline level for H established in all-H utterances.

Speaker KS is discussed separately because in the 20 total tokens analyzed, he shows two different patterns—one observed in 15 tokens, and one observed in 5 tokens. These are shown below.



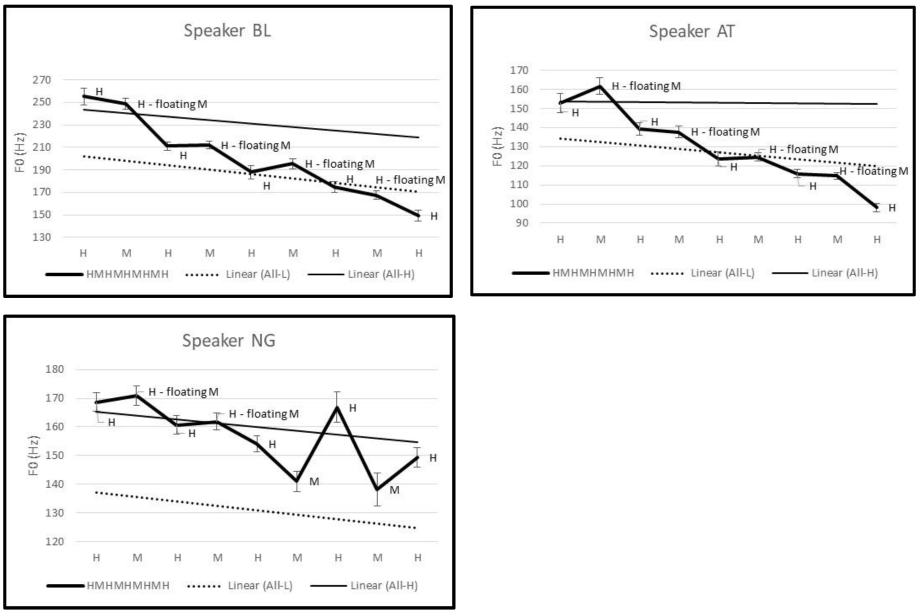
(442) Iterative underlying H–M sequences over 7 TBUs for speaker KS
(20 tokens total divided into 2 patterns)

Note first that for speaker KS, there is again anticipatory raising of H above the baseline for H; the initial raised level of H reaches its peak on the second TBU of the utterance rather than on the first.

In fifteen of the twenty tokens, speaker KS employs the cancellation strategy. Just as with speaker NG, this cancellation occurs after two downsteps and allows speaker KS to avoid realizing H below the baseline level for H seen in all-H utterances. When Tonal spread fails to operate, H is reset above the level of the previous H.

In five of the twenty tokens, however, speaker KS employs the strategy seen above for speakers AT and BL and allows for continuous non-automatic downstepping of the level of H below the baseline for H (but not below the baseline for L). Thus speaker KS shows some variation in his realization of these longer utterances with some of his tokens following the pattern seen for speaker NG and others of his tokens following the pattern seen for speakers BL and AT.

We can focus on what happens in longer utterances by looking at the data coming solely from the utterance of 9 TBUs (437). First we see the patterns displayed by three of the four speakers. Speaker KS is again examined separately.

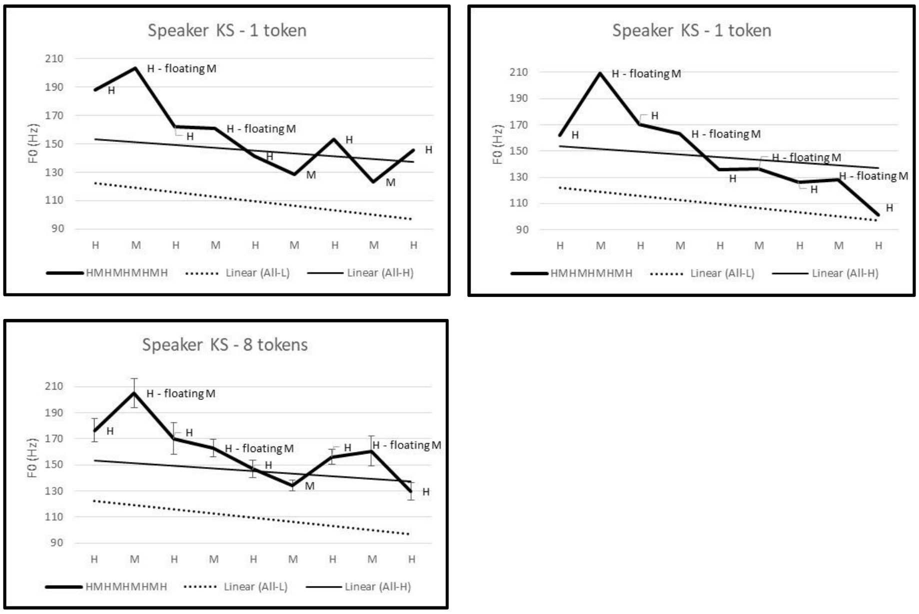


(443) Iterative underlying H–M sequences over 9 TBUs (10 tokens each)

Here we see that in an utterance of nine TBUs, speakers BL and AT continue to lower the F_0 of H until the final H ends well below the baseline level of L in all-L utterances of equivalent length. We also see that as the utterance progresses, there is a compression of the steps, so that the third step is smaller than the second, which is smaller than the first. (This relationship does not necessarily hold for the last downstep, however.) This compression of steps is likely a concession to accommodate the physical limitations of the pitch range.

In the graph of the data from speaker NG, non-automatic downstep is again canceled after two steps. Tonal spread no longer operates after this point and the following H is reset above the level of the previous H. Speaker NG does not spread H tone either in the third or final iteration of underlying /H–M/ sequences, so the underlying M tones in these sequences are realized as surface Ms with a F_0 produced roughly between the all-H and all-L baselines.

Speaker KS is not consistent in his production of the 9-TBU utterance. He employs three different patterns in his ten repetitions of the utterance. One pattern is employed for eight tokens, and the other two are each employed for a single token. This is demonstrated below.



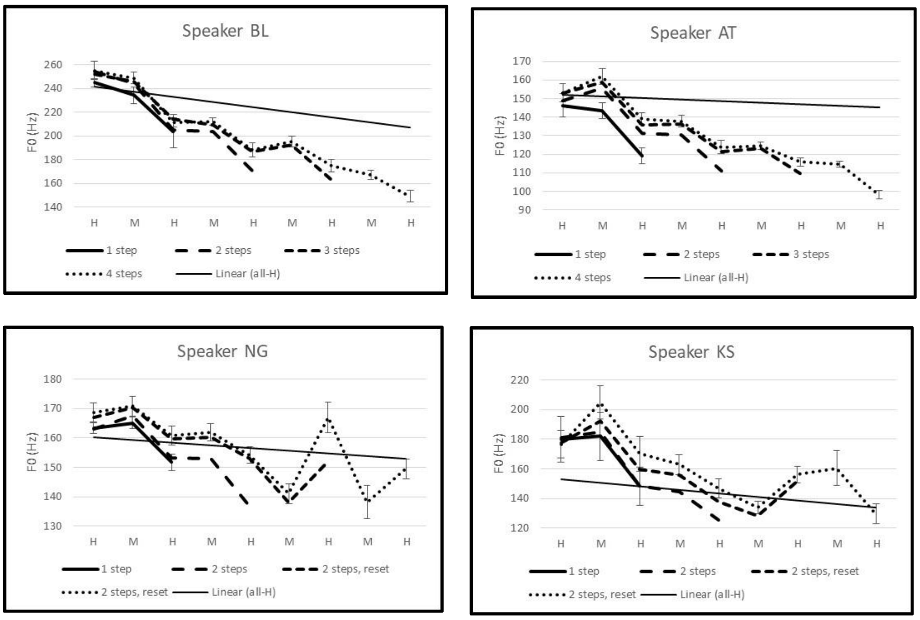
(444) Iterative underlying H–M sequences over 9 TBUs for speaker KS (10 tokens total divided into 3 patterns)

In one token, speaker KS employs the strategy used by speaker NG, canceling non-automatic downstep after 2 steps, and resetting H without reinitiating H spread throughout the rest of the utterance. In another token of the same sentence, he employs the strategy of speakers BL and AT of allowing continuous non-automatic downstep to occur throughout the entire utterance. This is assisted by a significant initial raising of H and a compression of downsteps two and three.

In the other eight tokens, non-automatic downstep is canceled after two steps, but this is a temporary cancellation. Speaker KS does not apply the rule of Tonal spread at this point in the utterance and H is reset above the level of the previous H. However, H spread is reinitiated again after this single reset of H and non-automatic downstep is triggered once again on the last H by a floating M.

We see, then, that speaker KS is eclectic in the strategies he employs in dealing with long utterances of alternating underlying Hs and Ms. In the majority of cases, however, he—like speaker NG—employs strategies that prevent the F_0 of H from dropping significantly below the baseline level of H that is established in all-H utterances.

We have seen that there is anticipatory raising of H involved in the production of non-automatic downstep of H. This anticipatory raising is often most clearly observed on the second surface H of the utterance rather than on the initial H—an effect of peak delay. We can look more closely at H raising in the context of non-automatic downstep by seeing what happens in utterances of alternating underlying Hs and Ms as they increase in length. The following graphs represent data from the utterances (434) through (437). Mean F_0 measurements are obtained from 10 tokens of each utterance.[114] The bars represent 95% confidence intervals for the /HMH/ and /HMHMHMHMH/ utterances.



(445) Increasing iterations of alternating underlying Hs and Ms (avg. 10 tokens)

For speakers AT, NG, and KS, the second surface H in the utterance is where one observes the greatest variance in the extent of H raising. For all speakers, the longest utterance (/HMHMHMHMH/) has the highest F_0 at this second TBU and the shortest utterance (/HMH/) has the lowest F_0 at this point. This is particularly true of speaker KS. It must be noted, however, that because the 95% confidence

[114] Because of a reading error, one token from the /HMH/ sentence of speaker NG was not able to be included. The confidence intervals given take into account this difference in the number of tokens. For speaker KS, the /HMHMHMHMH/ graph represents the eight tokens of his most common pattern of implementation.

intervals either touch or overlap at this point for all speakers except speaker AT, this finding would need to be verified through further studies with greater numbers of utterances.

To conclude this section, we see that there is a fairly clear and consistent distinction between the way speakers BL and AT manipulate H tone within their F_0 range and the way speakers NG and KS do so. Here in the discussion of non-automatic downstep, we see that for speakers BL and AT, the baseline for H does not correlate in any way with any limitations in the F_0 range of production of a H. A H tone can be realized both above and below the H baseline (and even below the L baseline). Speakers BL and AT produce non-automatic downstep of H which is not restricted or qualified in any way by the parameters tested here.

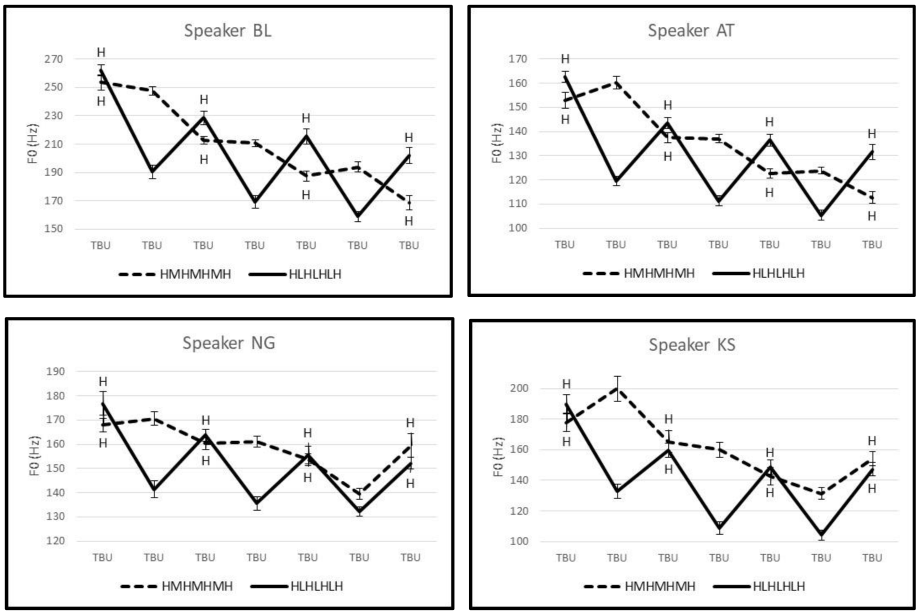
For speakers NG and KS, we see that in longer utterances (minimum of seven TBUs), speakers avoid producing a H significantly below the F_0 baseline for all-H tones. A H can be produced above the baseline for H through anticipatory raising, but these speakers avoid having a H tone fall significantly below it (although speaker KS is eclectic in his manipulation of tone and occasionally does permit pitch patterns that look like those of speakers BL and AT).

According to the stated criteria for downstep in section 7.3.1, the specific label given to the non-automatic downstep of H of speakers NG and KS is delimited, arrested non-automatic downstep. It is downstep that is delimited due to the avoidance of producing a H below a lower limit of F_0. In order to accommodate this constraint in the realization of H, speakers NG and KS both raise the level of H utterance-initially and compress the second downstep in comparison to the first. Despite both of these accommodations, only two downsteps are implemented before the level of H is reset. In this reset, Tonal spread fails to operate and a surface M is realized. This is followed by a H which has a F_0 higher than the H which preceded it. Any subsequent H is again lowered from this newly reset level. In the discussion of these results in 7.3.6, we look at the challenge of understanding how constraints regarding the limitations of the production of H are able to prevent phonological rules from being applied.

7.3.5 Comparison of automatic and non-automatic downstep

In this section, the final research question (4) from section 7.3.2 regarding downstep is addressed: Is the lowering of the F_0 of H which is attributed to a surface L between Hs equal to that which is attributed to a floating M between Hs?

The graphs below combine information previously seen in sections 7.3.3 and 7.3.4. For the solid line, the data come from 10 tokens each (20 tokens total) of utterances (432) and (433)—one utterance of 7 TBUs and one utterance of 9 TBUs, each with alternating surface H and L tones. For the dashed line, the data come from 10 tokens each (20 tokens total) of utterances (436) and (437)—one utterance of 7 TBUs and one utterance of 9 TBUs, this time each with alternating underlying H and M tones. Exceptionally in the case of speaker KS, the dashed line represents only 15 tokens total. This is because this comparison is taking into account only the most common of the two patterns that he uses for alternating sequences of H and M tones (section 7.3.4).



(446) Comparison of the lowering of H triggered by surface Ls or floating Ms

Once again these graphs highlight the fact that speakers AT and BL show patterns of phonetic implementation that are different than the patterns seen for speakers NG and KS. This has largely to do with the fact that speakers AT and BL have no lower F_0 threshold for the production of H, whereas speakers NG and KS usually do have such a threshold (although speaker KS allows for exceptions).

For speakers AT and BL, the lowering of H that is triggered by a surface L extends for a maximum of two instances of lowering (one larger and one smaller) that clearly exceed the rate of declination. After this, Hs that occur between Ls continue to lower, but at the same rate as declination—although having dropped below the baseline level of all-H utterances, they never return to that level (section 7.3.3). For these same speakers, the lowering of H that is triggered by a floating M is continuous throughout these utterances and brings the level of H to a position lower than the baseline level of all-L utterances (section 7.3.4). This means that the more iterations into the utterance, the greater the disparity between the F_0 of production of H as represented by the solid and dashed lines. After three non-automatic downsteps, the disparity between the F_0 of production of H in these two alternating sequences is highly significant. For these speakers, there is no equivalence in terms of Hz between the lowering of H that is triggered by a surface L and the lowering of H that is triggered by a floating M.

The situation is different for speakers NG and KS. In utterances that have a minimum of seven TBUs, these speakers have a relatively inflexible lower threshold of the F_0 production of H, correlating roughly with the baseline level of H found in all-H utterances. This lower limitation of the F_0 range of H is continually lowering throughout the prosodic unit at the rate of declination. Because these speakers avoid producing a H significantly below this limit, we find that in utterances where there is lowering of H triggered by a surface L and in those where lowering of H is triggered by a floating M, the F_0 levels for H at any given distance into the prosodic unit are quite close. In fact, after two non-automatic downsteps, there is considerable overlap of confidence intervals for the Hs in the two types of utterances. This is because in these longer utterances of both types, two downsteps is the maximum number of downsteps that can be produced before the level of H in these utterances reaches the threshold lower limits for H production.[115]

7.3.6 Discussion of results

We have seen that a significant issue raised by these tests is the question of what defines downstep as distinct from other pitch lowering phenomena. For purposes of comparison, the following is a reiteration from section 7.3.1 of the criteria used here to define downstep.

First, there must be a minimum of two steps down, each representing a decline from the preceding H that exceeds the rate of declination. Because there is anticipatory raising of H, the lowering need not descend below the baseline for an all-H utterance. However, if it does not descend below this baseline, I label it as "delimited downstep", and if it does not continue within the prosodic domain further than the minimum of two steps down, I give it the label "arrested downstep" (Rialland, 2001). In a case of "arrested downstep", there can still be, due to declination, progressive lowering of H beyond the two steps, but the rate of lowering will no longer exceed the rate of declination. Alternatively, there may be an upward resetting of the level of H after the two steps of lowering.

There are other descriptions of this type of situation where downstep is arrested and a phonologically-relevant lowering at an early stage in the utterance does not have lasting effects on the level of subsequent H tones late in the utterance. According to Genzel (2013), the lowering of the pitch F_0 of H in Akan is equivalent whether it is a case of automatic or non-automatic downstep. However, this lowering

[115] It seems that this threshold for the production of H is not as relevant in an utterance of five TBUs where exactly two non-automatic downsteps of H are realized, at which point the utterance ends. There seems to be an element of preplanning that is involved in the establishment of lower thresholds for H production over the course of longer utterances. This is a topic to be studied further.

is gradually offset over the course of a long utterance (of ten syllables) such that by the end of the utterance, the level of H is found to be equivalent to that of a H tone in an utterance of similar length composed only of H tones (Kügler, 2017). (In fact, there is a neutralization of both H and L at the end of the utterance.) The conclusion made is that Akan has 'phonologized declination' (Genzel, 2013).

There is also the observation that in some languages a surface L will cause a lowering of the immediately following H tone in a HLH sequence. In subsequent H tone syllables, however, the level of H will gradually creep up to the level of the initial H. It is concluded that this is a purely phonetic phenomenon (Connell, 2017).

Here in Saxwe, speakers NG and KS employ the strategy of resetting H tone after an initial one or two downward steps of H tone. As a result, in non-automatic downstep, levels of H tone at the end of a long utterance are roughly similar whether the utterance is composed solely of H tone, or whether there is an alternation between H and underlying M.

The results of these tests are in some ways quite surprising given the fact that these four speakers were chosen specifically because they were of similar age, of similar provenance, and because they speak the same dialect of Saxwe. The intention was to establish an understanding of the way that middle-aged speakers of the Saxwe variant spoken in the township of Houeyogbe manipulate pitch within the speaker's F_0 range in realizing certain sequences of tones. Instead, we see two rather divergent global approaches to manipulating tone—each approach employed by two speakers (although speaker KS vacillates occasionally between the two differing global approaches).

These global approaches differ primarily in whether or not in longer utterances there is a lower threshold for the realization of H. For speakers BL and AT, no such threshold is observed. In the case of speakers NG and KS, there is (usually) such a threshold and it corresponds roughly with the declining baseline levels of H in an all-H utterance.

This raises the question of how to view this threshold. It could simply be a strategy of preplanning which helps a speaker to avoid having to descend too low into his or her natural or comfortable range of production of pitch.

Taking into account the discussion of tone features in section 6.2, another possibility is that the distinguishing factor between these two global approaches is whether, in a given speaker's phonology, the distinction between [+upper] and [-upper] is purely relative to what precedes, or whether these values are divided by a more inflexible boundary in terms of F_0 production. In the latter case, the floating or surface feature [-upper] between two [+upper] TBUs will cause a lowering of F_0 so long as this does not result in crossing the relatively inflexible F_0 boundary that divides the [+upper] register from the [-upper] register.

An interesting observation is that for speakers BL and AT, whose "floor" for H is highly flexible, the F_0 lowering that is triggered by a floating M (non-automatic downstep) is clearly not equal in Hz to the lowering that is triggered by a surface L (automatic downstep). If one considers that what comes between the [+upper] register feature in both cases is the [-upper] feature (whether occuring as a floating M or as a surface L), one might expect that these two kinds of lowering would be equivalent.

Perhaps the reason for this difference is that the lowering triggered by a floating [-upper] tone can be the single element that marks a difference in meaning and is therefore phonologically contrastive. It has to exceed the rate of declination to be perceived by the hearer as distinct from the lowering of declination. The lowering triggered by a surface [-upper] tone is not phonologically contrastive and therefore does not have to exceed the rate of declination. Therefore, regardless of whether the sequence of register features is the same, these two types of lowering are implemented differently in the phonetic component in Saxwe. Their implementation is so different that in the case of speaker BL, we cannot even conclude (given the criteria established for downstep) that there is automatic downstep of H, whereas non-automatic downstep of H is clearly implemented for this speaker. This is very different from languages such as Chumburung in which the two types of downstep are shown to produce equal measurements of lowering in F_0 (Snider, 1998).

A final issue is how to explain the observation that for some speakers (such as NG and KS), Tonal spread may fail to occur in an utterance when its operation would result in bringing the level of the following H below the threshold of ideal production. This may be another indication that the predetermined inflexible boundary between the [+upper] and the [-upper] registers is part of the phonology of some speakers and not merely a product of the phonetic implementation. There are several arguments in support of this idea. First, it makes sense that a phonological parameter (such as whether register is defined in a relative manner or not) could be part of the conditioning environment for the application of a phonological rule such as Tonal spread. Second, the anticipatory failure of Tonal spread to occur cannot be easily explained by physical restrictions that impose themselves at the time of failure; the threshold for H is not at the lower limits of a speaker's F_0 range.

What is needed is further research within the Saxwe population in order to confirm whether this flexible/inflexible boundary between registers is truly a parameter that is relevant for all speakers and to what degree the observations noted here might be tied to other confounding factors, such as syntactic structures.

I turn now to lowering and other aspects of phonetic implementation having to do with L tone.

7.4 Phonetic implementation relating to L tone

In section 7.3, non-automatic downstep of H is described as the lowering of the level of H triggered by a floating M. Here in this section, the question explored is whether a floating M also triggers a lowering of the level of L. Stated otherwise, I explore whether there is non-automatic downstep of L in Saxwe. In the utterances used here, the floating M is present because a preceding H or L tone has spread onto an underlying M TBU, causing it to be delinked (section 3.2).

Recall from section 7.3.3 that pitch traces from alternating sequences of H and L TBUs reveal that there is no automatic downstep of L. In such an utterance, there is only a single instance of lowering of L beyond that which could be attributed to declination. Therefore, by the criteria I have laid out, we cannot claim that there is automatic downstep of L. Given this background, we turn to the question of whether there is non-automatic downstep of L.

7.4.1 Research question and utterances recorded

The following testing seeks to answer the research question: Is the lowering of the F_0 of L attributed to a floating M between Ls greater than that which could be attributed to declination?

In order to answer this question, the following set of eight sentences were created and recorded. Each sentence was recorded ten times by each speaker, and sentences were reshuffled between each pass.

(447) /$^{\text{M-}}$ télà sɔ̃́/
The tailor left. (3 TBUs - /HLH/)

(448) /$^{\text{M-}}$ télà ò sɔ̃́/
The tailor has left. (4 TBUs - /HLLH/)

(449) /$^{\text{M-}}$ télà ò gbɔ̃̀ sá/
The tailor already returned a while back. (5 TBUs - /HLLLH/)

(450) /$^{\text{M-}}$ télà gàgà ò gbɔ̃̀ sá/
The tall tailor already returned a while back. (7 TBUs - /HLLLLLH/)

(451) /$^{\text{M-}}$ télà gàgà gbĩ̀gbɔ̃̀ mɔ̃̀ drɔ̃̀ ké/
The tall tailor's return still remains [to be]. (9 TBUs - /HLLLLLLLH/)

(452) /$^{\text{M-}}$ télà nā̃ gbɔ̃̀ fí/
The tailor will return right away. (5 TBUs - /HLMLH/)

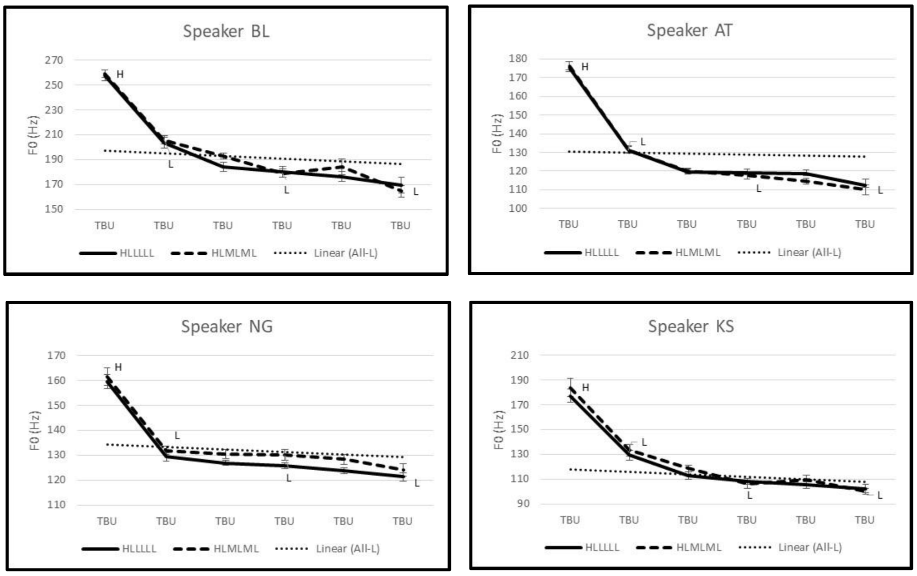
(453) /$^{\text{M-}}$ télà nā̃ ɦũ̀ ōɦɔ̃̀ fí/
The tailor will open the door right away. (7 TBUs - /HLMLMLH/)

(454) /$^{\text{M-}}$ télà nā̃ ɦɛ̃̀ āɦà mɛ̃̄ gbɔ̃̀ fí/
The tailor will bring that beverage back right away. (9 TBUs - /HLMLMLMLH/)

The sections below describes how the measurement obtained from these recordings were used.

7.4.2 Results

In the following graphs, all utterances begin with a H tone. The bold lines show the averages for repeated underlying Ls following the H. The dashed lines show the averages for alternating underlying Ls and Ms following the H. Each point in the graph is a mean calculated from 20 tokens.[116] These 20 tokens include 10 tokens of a 7 TBU-utterance and 10 tokens of a 9-TBU utterance. For the repeating Ls, data are taken from utterances (450) and (451), and for the sequences of alternating Ls and Ms, data are taken from utterances (453) and (454). Only data from the first six TBUs are included in these graphs; all utterances ended with a H tone as a frame, but it is the non-Hs that are in focus here. The bars represent 95% confidence intervals. In addition, a linear regression line from the all-L baseline data in section 7.2 is included.



(455) Iterative underlying Ls or L–M sequences over 5 TBUs (avg. 20 tokens)

---

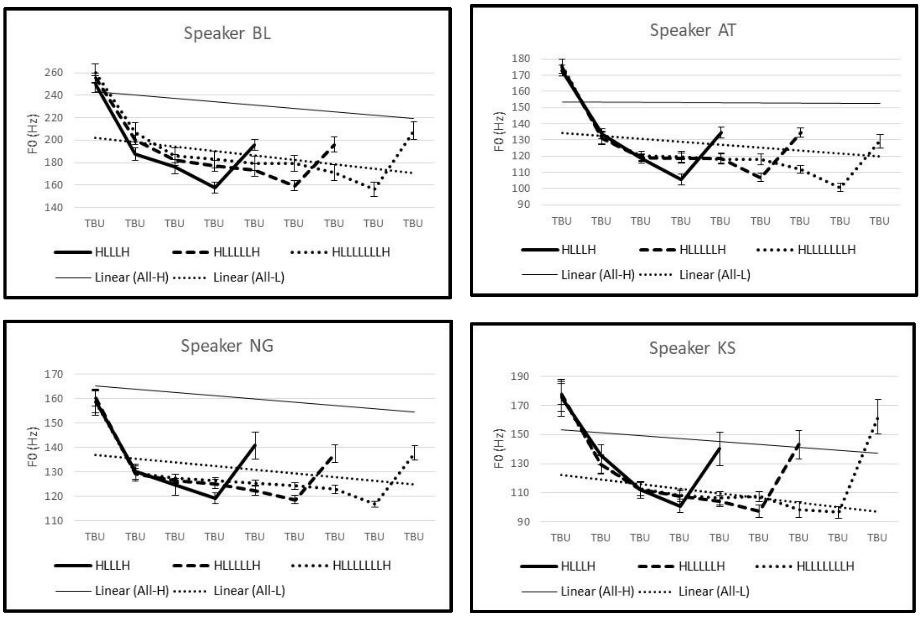
[116] Because of a reading error, speaker NG had only 19 tokens of the utterance with the sequence of Ls. The 95% confidence intervals take this into account.

In the alternating sequences of underlying Ls and Ms, L spreads to the M TBU, delinking this M. However, we do not see the same progressive down-stepping of pitch triggered by these floating Ms that we see when there is a floating M between underlying Hs (section 7.3.4). For all speakers, the /HLMLML/ sequence has confidence intervals that overlap at multiple places with the confidence intervals for the /HLLLLL/ sequence. For speakers AT and KS, the two pitch traces are essentially overlapping. Speakers BL and NG show some slight differences in the pitch traces of the two /HLMLML/ and /HLLLLL/ sequences, but by the last L TBU shown in these graphs, pitch levels are not significantly different from each other.

The conclusion is that just as there is no evidence of automatic downstep of L in Saxwe, there is also no evidence of non-automatic downstep of L in Saxwe. Floating M does not trigger downstep between L tones.

There is one interesting observation to be made, which is that the first L that follows the utterance-initial Hs in these utterances may be considered to be slightly raised in F_0. It is as of the third TBU that lowering happens at a stable rate parallel to the line of declination for all-L utterances. This is a phenomenon that could be described as reverse peak delay. One could label it more generally as "target achievement delay". Here, the target level of F_0 for L is achieved in the second L syllable following a H rather than the first. This can be seen most clearly for speakers AT and KS.When we include the final H tone from sentences (447) through (454) in a graph, we see another kind of lowering phenomenon which is worth noting. This lowering is not related to a floating M tone, but has to do with the final H TBU which was included in these sentences as a frame. In the following graphs, the three pitch traces represent 10 tokens each of a /HLLLH/, /HLLLLLH/, and /HLLLLLLLH/ utterance.[117] The vertical bars represent 95% confidence intervals.

[117] Speaker NG, because of reading errors, had only 9 tokens each of the /HLLLH/ and /HLLLLLLLH/ utterances.



(456) Increasing numbers of L TBUs between two Hs (avg. 10 tokens)

The most interesting thing to note from these graphs is that there is a dip in F_0 at the last L TBU in each utterance which immediately precedes the utterance-final H. Instead of continuing at the same rate of declination as the preceding Ls, this L is lowered at a steeper rate.

This is an anticipatory phonetic effect, like the raising of H before following L–H sequences (section 7.3.3). It is also, like the anticipatory raising of H, a dissimilation of tones which enhances tonal distinctions. Interestingly, while dissimilatory processes may be responsible for the anticipatory lowering of L *before* H, we do not see dissimilatory lowering of L *after* H—at least not when there are multiple Ls following this H. We have just seen that the first L that follows the utterance-initial Hs in these utterances may be slightly raised in F_0 with reference to the stable level of declining Ls that occurs afterwards in the utterance.

This means that for these speakers, the three Ls in a /HLLLH/ sequence are realized at three different positions with respect to the all-L baseline. For speaker AT, the first L in a /HLLLH/ utterance is at the L baseline, the second L is below the baseline at a stable point with reference to longer utterances, and the third L is lowered even further below the baseline in anticipation of the final H. For speaker KS, the first L is above the L baseline, the second is at the L baseline, and the third is below the baseline.

7.4.3 Discussion of results

We see in this section that there is no evidence of non-automatic downstep of L in Saxwe. Floating M does not trigger downstep between L tones. Downstep can only be triggered by a [-upper] register appearing between [+upper] registers. Since M and L are both [-upper] (differing only in the feature [+/- Raised]), alternating M–L sequences and all-L sequences are realized in approximately the same way. Lowering occurs in these contexts, but only at a rate consistent with declination.

While there is no lowering of L triggered by a floating M, there is another lowering-related phenomenon associated with L tones, which is that when multiple L TBUs precede a H, the last of these Ls is lowered beyond the level of lowering attributable to declination. This lowering of L before H is an anticipatory dissimilatory process—one which bears some similarities to the raising of H before L (section 7.3.3). Both are anticipatory processes, both involve dissimilation, and both occur in the sentences used for this testing between the two TBUs at the outermost edges of the IP.

It would be interesting to test whether, if the H were not utterance-final, this single instance of lowering of L before H would still be clearly observed. This would help to answer the question whether dissimilation in pitch implementation is most significant at the junction of two TBUs located either at the beginning or at the end of the IP.

Having treated several topics related to pitch implementation at the utterance level, I turn now to pitch implementation at the word level, focusing specifically on V.C(C)V nouns.

7.5 The most common tone patterns of V.C(C)V nouns

This section discusses the phonetic implementation of monomorphemic nouns (section 3.7). In this section, I concentrate on the nouns whose initial vowel is M; this excludes the /L.H/ noun tone pattern (section 3.7.8). I also exclude the extremely rare /M.H ^{M}/ tone pattern. The remaining six noun tone patterns are examined here: /M.H/, /M.M/, /M.L/, /M.M ^{H}/, /M.L ^{H}/, and /M.LH/.

For each of these six tone patterns, ten nouns were chosen to be recorded. These nouns are listed in Appendix D. They were chosen primarily for their ease of recognition from the French translation, but they also display the consonant-tone correspondences that are noted throughout this study—the /M.L/, /M.L ^{H}/, and /M.LH/ tone patterns most commonly have a depressor consonant (voiced obstruent) in their syllable onsets and the other tone patterns have a non-depressor consonant (voiceless obstruent, sonorant, /ɖ/, or /b/) in their syllable onsets.

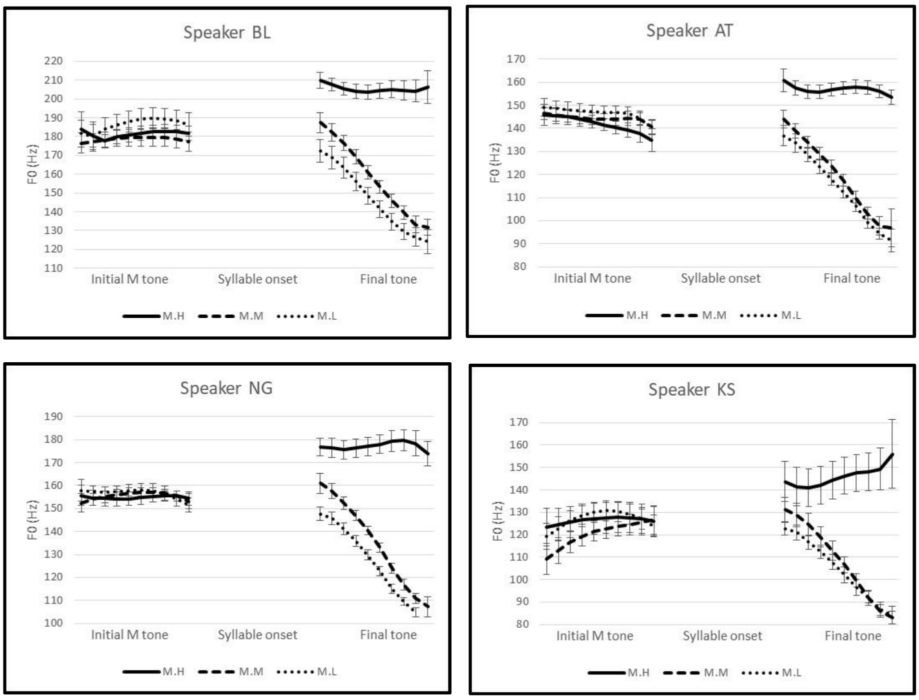
All sixty nouns were shuffled and a single pass was made through the words. Each word was pronounced two times. For speaker AT, this meant that there was a total of 20 tokens for each noun token pattern.

A variable that I had not predicted was that there was some interspeaker variation regarding the noun tone pattern assigned to certain of these nouns; the areas of divergence had largely to do with the tone pattern /M.L ^{H}/. Words that speaker AT pronounced with this tone pattern were not necessarily pronounced with that tone pattern by other speakers. This meant that for all speakers except AT, one or more of the six tone patterns had more than 20 tokens and the /M.L ^{H}/ had less than 20 tokens. The calculation of the 95% confidence intervals takes into account these differences. The following are the numbers of tokens for the speakers where they deviated from 20: speaker BL—22 of the /M.LH/ pattern, 22 of the /M.L/ pattern, 16 of the /M.L ^{H}/ pattern; speaker KS—22 of the /M.L/ pattern, 22 of the /M.M/ pattern, 16 of the /M.L ^{H}/ pattern; speaker NG—22 of the /M.L/ pattern, 16 of the /M.L ^{H}/ pattern, 2 of the /L.H/ pattern.[118]

7.5.1 The /M.H/, /M.M/, and /M.L/ tone patterns

The following graphs show the phonetic realization of the /M.H/, /M.M/, and /M.L/ tone patterns for all four speakers.

[118] Speaker BL did not include among words of the /M.L ^{H}/ pattern: [oɦjã] 'corn weevil' and [omlɛ̃] 'fishhook'. Speaker KS did not include among words of the /M.L ^{H}/ pattern: [omlɛ̃] 'fishhook' and [aɦa] 'side'. Speaker NG did not include among words of the /M.L ^{H}/ pattern: [oɦjã] 'corn weevil' and [aɦa] 'side'.



(457) Comparison of /M.H/, /M.M/, and /M.L/ tone patterns (avg. 20 tokens)

What we see very clearly from these graphs is the effect that the final $L_{\%}$ IP boundary has on utterance-final M and L tones. In section 3.5, I discuss the utterance-final lowering or downglide on any underlying M or L tone that does not have a floating H or boundary tone H following it. This utterance-final lowering occurs even when the underlying M is realized H because of Tonal spread (72). The mechanism which explains this final lowering or downglide is the right edge $L_{\%}$ IP boundary which links to the final M or L ([-upper]) TBU prior to Tonal spread.

In these graphs, we see again the observation made in section 7.2.2 that if there is an uneven distribution of H, M and L within the F_0 range of a speaker, it is M and L which are more closely localized within the range and H which is more distanced from M. This is true even before the lowering due to the IP boundary comes into play and is most clearly seen for speakers BL and AT. In fact, for speaker AT, 95% confidence intervals are overlapping during the realization of M and L both at the beginning of these final vowels and throughout their entire duration.[119]

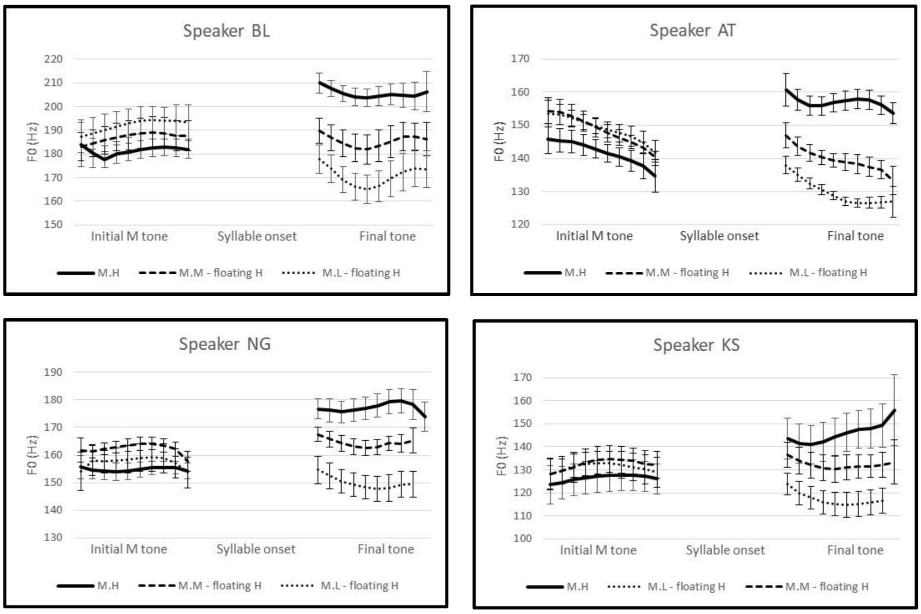
[119] The large confidence interval bars in the tenth measurement of the final vowel are due to the fact that Praat was often unable to produce a F_0 measurement at this final point in the duration of the vowel, and therefore confidence calculations are based on a small sample size.

As M and L are lowered over the course of the utterance-final TBU, the distinction between these tones becomes more neutralized so that for speakers BL, AT, and KS, there is overlapping of the 95% confidence intervals by the end of the duration of the final vowel even if there is no overlapping at the beginning of the vowel.

When the final $L_{\%}$ IP boundary is not in play, the outcome is quite different. This is what we see in the next section.

7.5.2 The /M.H/, /M.M ^{H}/, and /M.L ^{H}/ tone patterns

The following graphs show the surface realizations of the /M.H/, /M.M ^{H}/, and /M.L ^{H}/ tone patterns.



(458) Comparison of /M.H/, /M.M ^{H}/, and /M.L ^{H}/ tone patterns (avg. 20 tokens)

Because of the floating H tones on the /M.M ^{H}/ and /M.L ^{H}/ patterns, the final $L_{\%}$ IP boundary does not link to the utterance-final TBU and there is no final lowering on the M and L vowels like that seen in (457). In this context, we can see more clearly that there are three distinct levels for H, M and L, and that this distinction cannot simply be attributed to phonetically-driven consonantal influences because the contrasts remain relevant throughout the duration of the final vowel.

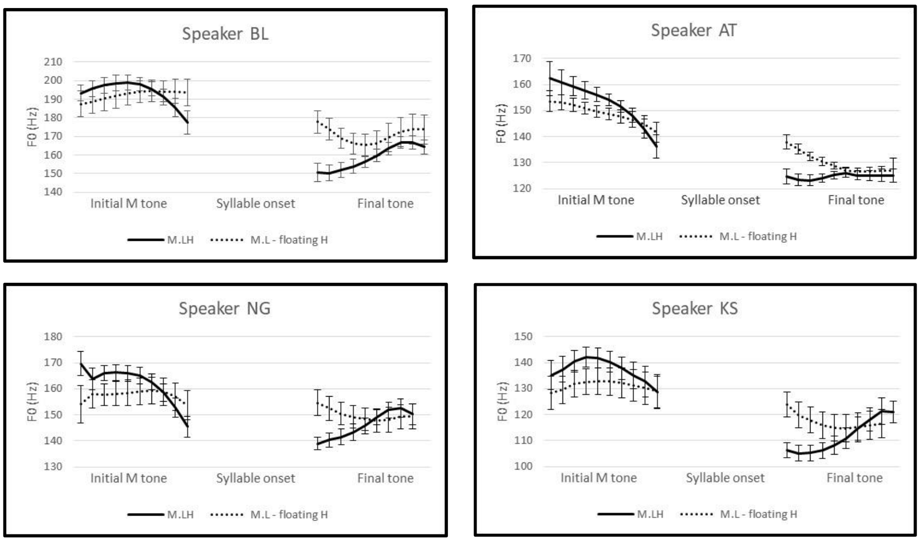
In these graphs, there is very little overlap of confidence intervals in the final vowel, and the overlap that exists is in the final measurements taken in the vowel. The large confidence intervals on the final measurement are due to the smaller numbers of F_0 readings provided by Praat at this point in the duration of the final vowel, and in some cases (speakers NG and KS), mean data points and confidence intervals were simply left off the final point of measurement because there were fewer than four data points available for these calculations.

We also see in these graphs some indication that there is a lowering of initial M before H in the M.H tone pattern. Although there is still a great deal of overlap of confidence intervals, we can see that the trend for all four speakers is that when it precedes H in the M.H pattern, M is realized lower in F_0 than it is for the other tone patterns shown in these graphs. This is an anticipatory phonetic process of dissimilation.

7.5.3 The /M.LH/ and /M.L ^{H}/ tone patterns

The /M.LH/ tone pattern is one of the more complex noun tone patterns in Saxwe. In section 3.7.5, we see that the LH contour in the /M.LH/ tone pattern is simplified—either by delinking the H (following L or M), or by deleting the L (following H). In the /M.LH/ noun, the contour is simplified by delinking the H. This makes the phonetic output sound very much like that of the /M.L ^{H}/ tone pattern.

In the derivations proposed for noun tone patterns in section 3.7, the difference between the phonetic output of /M.LH/ and /M.L ^{H}/ nouns lies in the fact that the floating H of the /M.L ^{H}/ tone pattern is deleted before the final output from the phonology, whereas the floating H which is a result of Contour simplification remains present in the output from the phonology. This accounts for the fact that there are some subtle differences in the F_0 traces. This can be seen in the following graphs.



(459) Comparison of /M.LH/ and /M.L ^{H}/ tone patterns (avg. 20 tokens)

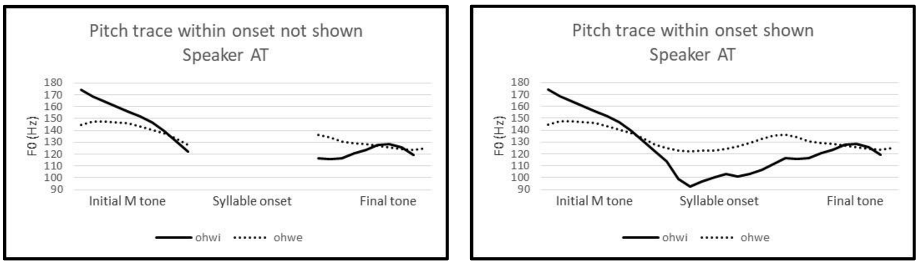
By about 60% into the duration of the final vowel, all speakers have largely overlapping confidence intervals for the /M.LH/ and /M.L ^{H}/ tone patterns, and for the remainder of the duration of the vowel, the F_0 traces are fairly close to each other. However, in the first half of the duration of this vowel, there is a difference in the pitch traces; for all speakers, the final vowel of the /M.LH/ tone pattern begins at a lower pitch than does the final vowel of the /M.L ^{H}/ tone pattern. The pitch then rises so that midway through the vowel, the two pitch traces have overlapping confidence intervals and start to be indistinguishable. This is the upglide or rising pitch that I represent in the notation of surface forms with the superscript [^{R}] as in [M.L ^{R}].

I do not have an explanation for why a final floating H in the output from the phonology triggers extra-low F_0 levels for a brief duration in the surface L vowel of the /M.LH/ pattern. However, these extra-low F_0 levels seem to be related to another observation—this time regarding the implementation of the initial M in the /M.LH/ tone pattern. This initial M is raised midway in the duration of the initial vowel in comparison with the M of the /M.L ^{H}/ tone pattern. Following this, the F_0 level slopes downward as it approaches the syllable onset.

The difference between surface realizations of these two tone patterns may be evident already in the production of the consonantal onset. We can see this if we look at the near-minimal tone pair /ōɦwǐ/ 'knife' and /ōɦwè ^{H}/ 'sun/fish'.[120] The

[120] The sound /ɦw/ functions as a single phoneme in Saxwe (section 1.4.3).

following pitch traces are averaged from two tokens of each word and come from the speaker AT.

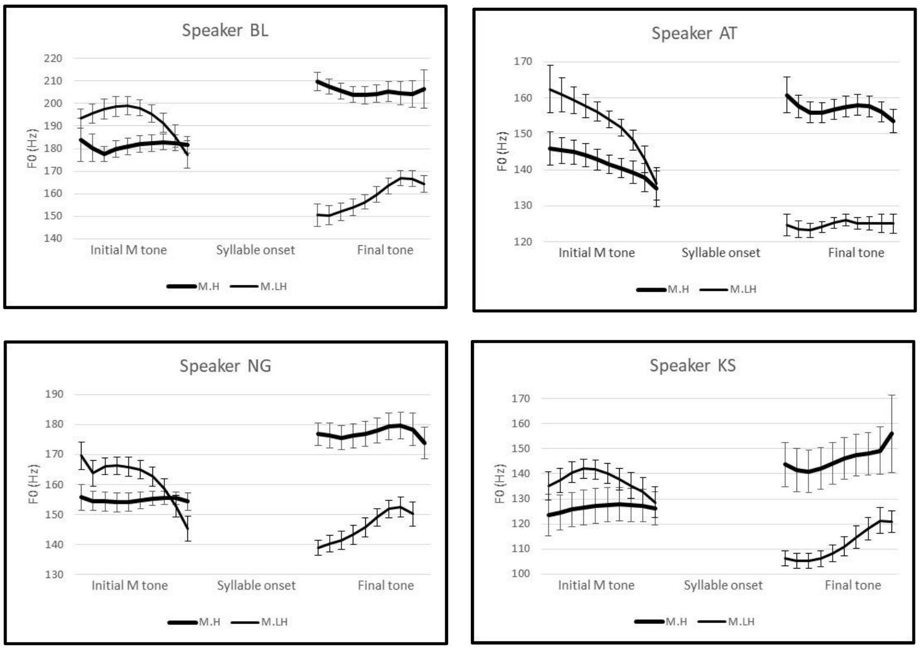


(460) /M.LH/ and /M.L ^{H}/ tone patterns in /ōɦwǐ/ and /ōɦwè ^{H}/ (2 tokens each)

These words share the same syllable onset but are realized differently due to their different underlying tone patterns. When we look at the pitch trace over the duration of the onset, we see that even during the articulation of the consonant onset, the pitch traces are dissimilar. This is an observation that needs to be explored in further studies of large data sets with words controlled for the consonant quality of the onset.

7.5.4 Dissimilatory effects on M in the /M.H/ and /M.LH/ tone patterns

Anticipatory dissimilation is a recurring phenomenon in the phonetic implementation of Saxwe tone. When looking at nouns produced in isolation, it is useful to compare the anticipatory raising of M before a [L ^{H}] vowel to the anticipatory lowering of M before a [H] vowel. This is what is shown in the following graphs.



(461) Comparison of /M.H/ and /M.LH/ tone patterns (avg. 20 tokens)

The surface realization of the initial underlying M of nouns in isolation varies considerably with respect to F_0. In these graphs, the 95% confidence intervals for the two realizations of M do not overlap during most of the duration of the vowel. These extremes (within a range of roughly 20 Hz) in the production of M can both be explained as anticipatory dissimilation. In the case of the /M.H/ pattern, M is realized at a relatively low F_0 before the following H. In the case of the /M.L ^{H}/ tone pattern, M is realized at a relatively high F_0 before the extra-low pitch produced early in the final vowel as a realization of the L–floating H combination. The result is that for speakers BL, AT, and KS, the initial M of the /M.L ^{H}/ tone pattern is realized at a F_0 value that is not too far from the F_0 value of the H in the /M.H/ pattern.

7.5.5 Discussion of results

In this section examining the six most common noun tone patterns (/M.H/, /M.M/, /M.L/, /M.M ^{H}/, /M.L ^{H}/, and /M.LH/), my principal objective is to illustrate by way of F_0 measurements the surface patterns that are described in earlier chapters of this study. In the course of doing so, certain trends become apparent.

First, there is a default utterance-final $L_{\%}$ IP boundary that links to a final M or L ([-upper]) TBU unless this TBU is followed by a floating or boundary H.

When the $L_{\%}$ IP boundary is linked to final M or L, the difference between these two tones is masked to some degree and 95% confidence intervals for the two means overlap, particularly toward the end of the duration of that final vowel (457). However, when this IP boundary is prevented from associating to the final TBU because of the presence of a floating H, we see in (458) three distinct levels of F_0 for H, M, and L.

Another observation made here is that the upgliding [M.L^R] surface form that is derived from a /M.LH/ underlying tone pattern can be distinguished from the non-falling [M.L°] surface form derived from /M.L^H/ by two relatively small pitch differences. First, the F_0 of the initial M of the /M.LH/ pattern is raised and then begins to trail downward as it approaches the consonantal onset. Second, following the consonantal onset, the F_0 of the surface L which follows begins extra-low and glides upward so that by the end of the vowel the F_0 measurements of the [M.L^R] and [M.L°] surface forms are largely overlapping. What happens within the consonantal onset is a topic that deserves further study and must be tested using a larger sample size of near-minimal pairs controlled for equal occurrences of specific depressor consonants. There are indications that the F_0 levels produced during the duration of the depressor consonant are manipulated in order to help distinguish between the /M.LH/ and /M.L^H/ tone patterns—specifically to distinguish between the phonological output of the former which includes a floating H and the output of the latter which does not.

In (461), we see that M is realized at a relatively low F_0 before the H of the /M.H/ tone pattern and that in the /M.LH/ tone pattern, M is realized at a relatively high F_0 before the extra-low realization produced for the L–floating H combination. In these contexts, tonal oppositions are emphasized through anticipatory pitch dissimilation.

7.6 Summary: aspects of the phonetic implementation

The purpose of this chapter is to give instrumental support for claims made in this study regarding certain aspects of tonal implementation—including those made about the existence or non-existence of automatic and non-automatic downstep, as well as those made about the surface differences between underlying V.C(C)V tone patterns. The results described in this chapter also point to future areas where further and more comprehensive testing could be carried out to clarify aspects of the phonetic implementation of tone in Saxwe.

The four speakers recorded for this testing were intentionally chosen in an effort to have test subjects of similar sociolinguistic background. However, despite their similar linguistic background, there is interspeaker variation in details such as how the downstep of H is implemented, as well as the assignment of particular underlying tone patterns to certain of the V.C(C)V nouns.

One useful starting point for the description of the phonetics of tone is to establish baseline values for utterances of all-H, all-M, and all-L tones for each speaker. From the baseline values for these four speakers, it is clear that declination is relevant in the realization of all three tones. Taking into account the results from all four speakers, we see that the rates of declination are not consistently greater for one tone than for any other (section 7.2). Speakers vary in how they space all-H, all-M, and all-L utterances within their F_0 range; these are either spaced evenly or with a greater gap in F_0 between all-H and all-M than between all-M and all-L. This latter proximity between M and L may reflect the fact that these two tones are [-upper] register, whereas H is [+upper].

In examining the question of whether there is automatic downstep of either H or L in Saxwe (section 7.3), I include some discussion as to how precisely downstep is defined. Automatic downstep of H is typically described as the iterative lowering of the level (or the "ceiling") of H triggered by a surface L (Clements, 1979; Connell, 2011; Stewart, 1993). However, there is less consensus about the specifics—including whether the lowering must be in excess of lowering that can be attributed to declination (something that may be assumed but is not always stated explicitly). Although there are other ways to characterize the relationship between downstep and declination (Downing & Rialland, 2017a), the criteria chosen here are that in order to qualify as downstep, the lowering observed must be in excess of that attibutable to declination. Furthermore, there must be at least two iterative steps in order for the term "downstep" to be used.

The finding here is that for all four speakers, there is an anticipatory raising of H utterance-initially in consecutive H–L sequences. This anticipatory raising creates the space for a large first step down. However, whether there is a second downstep of the level of H is less clear among some of these speakers, and there is clearly no third downstep of the level of H. Therefore, downstep of H, if it is a reality for some speakers, is arrested after two steps. In fact, speaker AT—who most clearly seems to have a second downstep in longer utterances—follows his second downstep in longer utterances by a resetting of the next H at a F_0 higher than that of the previous H. We can say that judging from the speakers tested here, there is for some of the population an iterative automatic downstep of H, but it is an arrested downstep limited to two steps. This is perhaps understandable given the fact that the lowering of H in consecutive H–L sequences does not represent a phonologically salient distinction.

There is no justification in these data for claiming a phenomenon of automatic downstep of L in Saxwe. For the speakers tested here, there is in consecutive H–L sequences a maximum single instance of lowering of the level of L beyond that which could be attributable to declination.

A further observation made in the course of studying automatic downstep is that for speakers NG and KS, the declining baseline utterances for all-H and all-L utterances roughly delimit the lower threshold of the F_0 of production for H and L tones in long utterances of H–L sequences. For speakers BL and AT however, in consecutive H–L sequences, Hs are realized at a position lower in F_0 than the all-H baseline—although after the first or second instance of lowering, Hs and Ls are then produced in a pattern of decreasing F_0 roughly parallel to the lines of declination. One could say that the frequency range for production of Hs and Ls has shifted downward, but once shifted, remains in line with standard declination rates. This difference in implementation strategies highlights the need to define what constitutes downstep.

In studying the phenomenon of non-automatic downstep of H in Saxwe, we see that there is interspeaker variation on this issue, even for speakers chosen specifically for their similar sociolinguistic background (section 7.3.4). Non-automatic downstep in Saxwe is triggered by a floating M between Hs. In underlying iterative H–M sequences, H normally spreads to the M TBU, delinking the M and creating a floating M between Hs. Speakers BL and AT—those speakers whose single-tone baselines do not correlate with any limitations of production for those tones—allow for four downsteps of H, triggered each time by a floating M. It is possible that more downsteps would be possible; four was the highest number that was tested for in this study. After four downsteps, the final level of H for these speakers is below the all-L baseline.

However, speakers NG and KS only permit two non-automatic downsteps of H, after which downstep is arrested and the level of H is reset by means of the speakers' failure to spread H tone. This in turn results in a failure to create the floating M tone which would trigger downstep. Once again in the non-automatic downstep of H, we see that there is anticipatory raising of the initial H in preparation for subsequent lowering. It is because of this anticipatory raising that speakers NG and KS are able to realize two non-automatic downsteps before reaching their threshold of production for H.

The difference between these two groups of speakers seems to be whether or not there is a preferred lower threshold for the realization of H in long utterances (provisionally defined as an utterance of seven or more TBUs). The distinguishing factor between these two global approaches to tone implementation seems to be whether, in a given speaker's phonology, the distinction between [+upper] and [-upper] is purely relative to what precedes, or whether these values are divided by a predefined boundary. In the latter case, the floating or surface feature [-upper] between two features [+upper] will cause a lowering of F_0 so long as this does not result in crossing the threshold that divides the [+upper] register from the [-upper] register. This factor (relative *vs.* delimited boundaries between [+upper] and [-upper]) may be part the speaker's phonology rather than simply an issue of phonetic implementation. Having this limitation on the F_0 production of H means that there is a conflict involved in allowing the phonological rule of Tonal spread to operate if the result is that a floating M will subsequently trigger the realization of H at a F_0 level that is below the preferred threshold.

In comparing the non-automatic and automatic downstep of H (or at least the initial lowering of H we see in consecutive H–L sequences, since not all speakers have an iterative lowering of H in these sequences), we see that for speakers BL and AT, whose "floor" for H is relative, the two kinds of lowering are not at all equivalent in terms of F_0 (section 7.3.5). That is, the lowering that one observes in underlying H–M sequences (non-automatic downstep) is not equal in Hz to the lowering that one observes in surface H–L sequences. Automatic downstep of H, if it exists, does not exceed two steps for these speakers, whereas non-automatic downstep can continue lowering for at least four steps (the maximum tested here). This means that the closer to the end of the utterance, the greater the difference between the Hs in these two kinds of sequences.

For speakers who have a threshold limit for H, however, there is a fairly close correspondence between the F_0 of Hs when comparing long utterances of underlying H–M sequences to long utterances of surface H–L sequences. This is because the anticipatory raising of H is a reality in both types of utterances, and the threshold on the lowering of H is also a reality in both types of utterances. There can only be two steps down in either non-automatic or automatic downstep (if the latter exists) for these speakers. For this reason, it is labeled as delimited downstep.

The fact that non-automatic downstep of H (at least for the first two steps down) is clearly a relevant phonetic implementation strategy for all speakers is understandable given that non-automatic downstep reflects contrastive differences in meaning. Automatic downstep of H does not reflect contrastive differences in meaning, and therefore it is not surprising to find that this kind of downstep of H is not as clearly a relevant phonetic implementation strategy for all speakers.

Finishing off the topic of downstep, we see that just as automatic downstep of L is not a phenomenon observed in Saxwe, it is also clear that non-automatic downstep of L is not found in Saxwe either (section 7.4). There is no downstepping of the level of L triggered by a floating M in Saxwe. The fact that downstep (automatic or non-automatic) of L does not exist in Saxwe is not entirely surprising if we consider that downstep in Saxwe occurs when there is alternation between [+upper] and [-upper] registers. Both M and L are [-upper] register and therefore there is no such alternation when either a surface or floating M appears between Ls.

In further discussion of the realization of L, there are two types of phenomena which are worth summarizing. One phenomenon that is observed is seen in utterances where an initial H is followed by multiple Ls; there is a delay in achieving the target F_0 of L such that this target is reached for some speakers on the second L following the H rather than the first. This means that the first L following the H may appear slightly raised in F_0 compared to the more stabilized levels of L TBUs which follow. Another phenomenon that is observed is an anticipatory lowering of L before an utterance-final H (section 7.4.2).

Before turning to the discussion of noun tone patterns, I note that there are several generalizable trends of Saxwe phonetic implementation strategies that can be observed in multiple places in this chapter. The first trend we can label as anticipatory dissimilatory F_0 movement. This includes the raising of an utterance-initial H before a L–H sequence and the lowering of a L before the final H in an utterance.

Another trend we can generally refer to as target achievement delay. This label covers both the peak delay seen for utterance-initial Hs when the surface form begins with multiple Hs (seen in the all-H utterances of section 7.2.3 as well as the surface realizations of /HMH…/ utterances in section 7.3.4), as well as the delay in achieving the targeted F_0 of L after an initial H. Because the term peak delay indicates that we are dealing with a peak, target achievement delay covers both delays in attaining peak F_0 targets as well as delays in attaining trough F_0 targets.

We now turn to a summary of the phonetic implementation of the most common V.C(C)V noun tone patterns: /M.H/, /M.M/, /M.L/, /M.M ^{H}/, /M.L ^{H}/, and /M.LH/. We see clearly in section 7.5 the effects of the utterance-final $L_{\%}$ IP boundary as it links to a final M or L ([-upper]) tone, creating a downward falling pitch trace utterance-finally. In fact, the difference between these two tones is masked to some degree by this utterance-final $L_{\%}$ IP boundary as 95% confidence intervals for M and L values overlap.

The $L_{\%}$ IP boundary tone does not link to a final M or L tone if it is followed by a floating H in the underlying form. It is then that we observe three distinct levels of F_0 for H, M and L—levels which remain distinct throughout the duration of the vowel and not simply during the initial period of the vowel during which onset consonant quality may have an effect.

One result of the testing here was to distinguish instrumentally between the upgliding [M.L^R] surface form that is derived from a /M.LH/ underlying tone pattern and the the non-falling [M.L°] surface form derived from /M.L^H/. There are two subtle differences in F_0 traces that distinguish these. First, the F_0 of the initial M of the /M.LH/ pattern is raised and then begins to trail downward as it approaches the consonantal onset. Second, following the consonantal onset, the F_0 of the surface L which follows begins extra-low and glides upward so that by the end of the vowel the F_0 measurements of the [M.L^R] and [M.L°] surface forms are largely overlapping. What happens within the consonantal onset is a topic that deserves further study as it appears that there may be, as part of the realization of the underlying tone pattern, manipulation of F_0 levels within the duration of the consonant itself.

We also see within V.C(C)V noun tone patterns some anticipatory dissimilatory F_0 trends not unlike those seen in the testing of entire utterances. We see that M is realized at a relatively low F_0 before the final H of the /M.H/ tone pattern, but M is realized at a relatively high F_0 before the extra-low realization of the following L–floating H sequence which is the output from the /M.LH/ tone pattern. In each of these cases, what we see is a dissimilation of adjacent tones utterance-initially or utterance-finally through the anticipatory raising or lowering of F_0. This tendency to emphasize tonal oppositions occurs in all of these cases in the first two or last two TBUs of an IP.