BA thesis Linguistics
University of Amsterdam
Date: 16/07/2025

Tristan Bouverot 14204843
tristan.bouverot@gmail.com
Supervised by Titia Benders
Co-Supervised by Martine Bruil

Ecuadorian Siona Prosody: towards a separation of pitch accent and pre-aspiration

Word count : 8390

UNIVERSITEIT VAN AMSTERDAM

Abstract

This thesis investigates the relationship between pitch accent and pre-aspiration in Ecuadorian Siona. Pitch accent is a word final phenomenon, while pre-aspiration is iterative, falling on every even-numbered syllable when counting from the left-edge of the word. Previous studies on the prosody of Ecuadorian Siona, as well as on its genetic affiliates analyzed the two phenomena either under the undefined umbrella term of 'stress', or assumed pitch accent distribution to be dependent on pre-aspiration. Such an assumption (inspired by Metrical Theory), ignores the fact that pitch accent consistently does not occur on pre-aspirated syllables in words with an odd number of syllables.

In order to investigate the potential relationship between the two phenomena, five analyses were drawn depending on the theoretical approach. The parameters differentiating analyses were the framework used (Metrical Theory/Accent First Theory), the preliminary analysis made of the underlying nature of pre-aspiration (rhythm/foot-mediality) and the number of different foot inventories assumed to exist in Ecuadorian Siona in the case of the foot-medial Metrical theoric account.

Results indicated that the rhythmic Accent First Theoric appeared to be the most parsimonious one, accounting for the possible presence of a secondary accent in Ecuadorian Siona. The implications of such results are that pitch accent and pre-aspiration are completely unrelated phenomena in Ecuadorian Siona.

Acknowledgements

I am very grateful to my supervisor Titia Benders, who supported me and answered my questions throughout the whole process of writing this thesis. I can not thank her enough for taking the time to review my work every week even when my submissions did not come at the agreed time or day. Finally, I cannot understate how helpful it was to me to have a more phonetically oriented linguist constantly trying to poke holes at my phonological theories, without this constant questioning I would not have been able to go in such depth in my analyses.

I am very thankful to my co-supervisor Martine Bruil, for providing me with the necessary literature on Ecuadorian Siona, as well as for being the literature herself in many ways, being the main researcher involved with Ecuadorian Siona. I can't understate my gratitude when it comes to the opportunity Martine has given me to go to Ecuador to work on Ecuadorian Siona myself.

I am grateful for Andrés Salanova and Justin Case, who were both present during my fieldwork in Sototsiaya and taught me a lot about amazonian languages as well as the art of doing fieldwork.

My thanks go as well to the community of Sototsiaya, for hosting us, for answering our endless questions about Siona, and for working hard to make our stay comfortable. In particular I want to thank Miguel for helping organize the trip, and letting charge my phone of laptop in their house in a community where power outlets were quite scarce. My thanks also go to Angela and her family for hosting us next to their house and for cooking for us every day without fault.

I also want to thank Marijn van 't Veer, who made me fall in love with phonology, and introduced me both to Martine and to the subject of pre-aspiration in Ecuadorian Siona. Thank you Ksenia Bogomolets, for taking the time to answer my questions about Accent First Theory by email, despite not knowing me.

I am also very appreciative of my friend Amaury for taking the time to peer review my work, and discuss both of our theses in between games of chess. Lastly, I want to thank my girlfriend Julia, for always supporting me despite the stress of writing her own thesis and drowning under approximately twice the amount of work as I was. Somehow she managed to find the energy and proof-read and correct language mistakes in a paper she did not understand in the slightest.

Table of contents

1 – Introduction

The current paper concerns itself with prosody in Ecuadorian Siona (ES), a western Tukanoan variety of Siona spoken in eastern Ecuador. This thesis is specifically focused on the phenomena of pre-aspiration and pitch accent and their potential relationship. Indeed, although pre-aspiration and pitch accent are both differently realized and conditioned, they are referred to in the literature as a terminologically vague notion of stress (van 't Veer, 2025).

On the surface, 'pre-aspiration' in ES is a phenomenon that occurs on every onset non-glottal obstruent, provided they are in an even numbered syllable counting from the left side of the word, and that it follows a short vowel. Pre-aspiration may occur multiple times a word, or none, meaning it does not fulfill common criteria for definitions of stress such *culminativity* and *obligatoriness* (Hyman; 2006, 2009). 'Pitch accent' in ES on the other hand is observed as a pitch rise word-finally. As such, it always occurs once per word, and fulfills the conditions of culminativity and obligatoriness, making it a stronger candidate for stress in ES following traditional definitions.

A grouping of pitch accent and pre-aspiration into a notion of stress in ES strongly implies that both are related and co-dependent in their assignment, as prosodic theories such as Metrical Theory (Liberman & Prince, 1977) or Accent First Theory (Van der Hulst, 1996, 2009, 2012, 2014a, 2014b; Goedemans & Van der Hulst, 2014; Bogomolets 2020, 2023) predict. van 't Veer et al. (2025) for instance consider pre-aspiration as a property of stressed syllables, which alternates due to the iambic foot structure of ES. Furthermore, they assume that pitch accent is reliant on pre-aspiration for its assignment. However, regular co-occurence between the two phenomena is only partly attested. Words with an even number of syllables feature co-occurence of pre-aspiration with pitch accent (σ.^{h}σ.σ.^{h}σ́), whereas it is not the case on the last syllable of words with an odd number of syllables (σ.^{h}σ.σ.^{h}σ.σ́).

In an attempt at providing a clearer picture of the prosodic situation in Ecuadorian Siona, the current study focuses on answering the following research question: Which theoretical lens provides the most parsimonious analysis of the prosodic relationship between pre-aspiration and pitch-accent in Ecuadorian Siona?

Chapter §2 starts with a discussion on current the state of prosodic research in ES and genetic affiliates as well, followed by a discussion on the underlying nature of pitch accent and pre-aspiration in Chapter §3. We then investigate whether there is, and the nature of the link between pre-aspiration and accent through the lens of two theoretical frameworks. The analyses using Metrical Theory are made in Chapter §4, and the ones using Accent First Theory are made in Chapter §5. Results are discussed in Chapter §6, with first a comparison of the leading analyses from both theoretical frameworks in §6.1, following by a discussion on the limitations of this study and suggestions for future research in §6.2. Finally, a general conclusion is drawn in Chapter §7

2 – The state of the art

2.1 – Accounts of prosody in the Siona dialect group

ES is a variety of the Western Tukanoan language of Siona, spoken in the west of the amazon rainforest in South America (Chacon, 2014). Originally a single language, it is now split into three varieties that now have significantly diverged. Colombian Siona is spoken in Colombia, Peruvian Siona[1] in Peru, and ES in Ecuador. The split is of such importance in fact that Bruil (2014:11) considers ES to be synchronically closer to Sekoya than to Colombin Siona, despite a more distant genetic relationship. Still, Colombian Siona, Sekoya and ES are considered as part of a single dialect continuum, which I call for lack of a better term the 'Siona dialectal group'. One area where the three varieties bear similarities is in the domain of prosody.

In Chacon (2010) as cited in (Wetzels & Meira, 2011), the Siona dialectal group as a whole is referred to as consisting of 'stress accent' languages. This description contrasts with other Tukanoan languages which are described as either 'restricted tonal languages' or 'mixed systems' languages. Accounts on individual varieties do confirm those claims, if in terminologically vague terms.

For Colombian Siona, Wheeler (1970: 20-21), Wheeler (1987: 89-92) and Wheeler & Wheeler (1975: 63-64) all indeed report occurence of stress observable through its selection of allophones. One example of this is a phenomenon that has been analyzed in Wheeler (1987: 90) as unvoiced vocalic lengthening or as pre-aspiration in Wheeler & Wheeler (1975: 64). In both accounts, this phenomenon is said to occur before the stressed syllable, found in the second syllable in the word (1,2). Similarly in Sekoya, Johnson & Peeke (1975) and Johnson & Levinsohn (1990) describe a word-second stress which triggers unvoiced vocalic lengthening in the previous syllable, provided the stressed syllable is unvoiced (3, 4).

(1) Pre-aspiration on stressed syllables in Colombian Siona
/su.ˈsi/
[su.^{h}si]
'(type of nettle)'
(adapted from Wheeler 1975: 64)[2]

(2) Unvoiced vocalic lengthening before stressed syllables in Colombian Siona
/ʔa.ˈp^{h}a.si/
[ʔaḁ.p^{h}a.si]
'sapote (fruit)'
(adapted from Wheeler 1987: 90)

[1] Peruvian Siona has not been thoroughly studied yet, and whether the variety is even still spoken to this day is unclear. As such this variety is not discussed in this thesis.
[2] Examples in this paper are only cited when coming from an external source. Part of the data I use here comes my own fieldwork experience in the community of Sototsiaya, in Ecuador.

(3) Unvoiced vocalic lengthening before stressed syllables in Sekoya
/ʔa.ˈpa.si/
[ʔaa̯.pa.si]
‘sapote (fruit)’
(adapted from Johnson & Levinsohn 1990: 19)

(4) Unvoiced vocalic lengthening before stressed syllables in Sekoya
/ha.ˈso.si.ʔi/
[haa̯.so.si.ʔi]
‘I shot it’
(adapted from Johnson & Levinsohn 1990: 19)

2.2 – Accounts of prosody in Ecuadorian Siona

2.2.1 – Distribution

Ecuadorian Siona possesses two main surface prosodic processes. The first one is pitch accent rise, which occurs once a word in the utimate position (Bruil, 2014). The second process is similar to the one described in Sekoya and Colombian Siona as either unvoiced vocalic lengthening (CV̥V.CV) or pre-aspiration (CV.^{h}CV). In the case of ES, van 't Veer et al. (2025) make a pre-aspiration (CV.^{h}CV) analysis of the phenomenon, arguing that pre-aspiration is a property of stressed syllables.

In this account, pre-aspiration occurs iteratively on onsets of iambic[3] syllables, provided that they are non-glottal obstruents {p, t, tʃ, s, k, k^{w}, (ɲ)} following short vowels. Consonant clusters are forbidden in onsets in ES, and only the glottal stop /ʔ/ is allowed as a coda (Bruil, 2014). Hence onsets always follow either a glottal stop (5), a long vowel (6), or a short vowel, in which case pre-aspiration may occur (7-9), provided the onset is a non-glottal obstruent (10, 11). In following examples, underlying pre-aspiration is represented with underlines and underlying pitch accent is represented with an accent mark, in accordance with the analysis of pitch accent as a type of accent made in §3.1. Following Andrikopoulou et al. (2021), accent is considered as specified in the lexicon. As such, examples in this paper which use of feet to determine accent/stress placement similarly depict feet as underlying.

[3] In this context, ‘iambic syllable’ means ‘even-numbered when counting from the left’. Opposite to ‘trochaic syllable’, which in this context would mean ‘odd-numbered when counting from the left’. More on this in §4

(5) Pre-aspiration is blocked by preceeding glottal stops
|waʔ.ˈti|[4]
/waʔ.tí/
‘machete’

(6) Pre-aspiration is blocked by preceeding long vowels
|saː.ˈsio|
/saː.sío/
‘She will take’
(adapted from Bruil, 2014: 100)

(7) Pre-aspiration may occur after short vowels
|wa.ˈti|
/wa.ʰtí/
‘bad spirit’

(8) Pre-aspiration may occur after short vowels
|o.ˈko|
/o.ʰkó/
‘water’

(9) Pre-aspiration may occur after short vowels
|a.tʃa.si.kɨ.ˈpʔi|
/a.ʰtʃa.si.ʰkɨ.pʔí/
‘The one (m) listening’
(Bruil, 2012: 20120918elicr004)

(10) Pre-aspiration does not occur on glottal obstruents
|ho.ˈpʔo|
/ho.pʔó/
‘the middle’
(van 't Veer et al., 2025)

(11) Pre-aspiration does not occur on glottal obstruents
|ho.ˈtʔo|
/ho.tʔó/
‘flower’
(van 't Veer et al., 2025)

Note that surface [ʰɲ] is also attested in ES, but van 't Veer et al. (2025) analyze it as being the result of a counter bleeding ordering of multiple rules. Indeed, in cases of

[4] Note that instead of using the traditional binary phonetic versus phonological levels of representations indicated respectively by square brackets and slashes, most examples are depicted following notation from the BiPhon model of grammar (Boersma, 2011). In this model, phonology is split into underlying (|xx|) and surface forms (/xx/), and phonetics into articulatory ([xx]) and auditory forms ([[xx]]), allowing for finer layered distinctions between different phenomena: In perception in a given language X, the auditory form [[F1 = 700_{Hz}, F2 = 1300_{Hz}]] becomes surface /a/, and then in turn underlying |aɑ|. Two steps are needed to represent those three stages in representational level, whereas a traditional phonetic layer vs. phonological layer allows only for one step and two representational levels in this case.

pluralization involving the -jã suffix, pre-aspiration can occur on a segment which is then truncated, leaving the intermediate form /ʰjã/. This /ʰjã/ then becomes /ʰɲã/ through regressive nasal harmomny. See (12) for an example of surface [ʰɲ] in the word *toŋa* ‘planks’, derived from *toto* ‘plank’ using the *-jã* pluralizing suffix.

(12) Counterbleeding rule ordering yields explains surface [ʰɲ]

0. Underlying form	/toto/	/toto-jã/
1. Pre-aspiration	to.ʰto	to.ʰto.jã
2. Truncation	–	to.ʰjã
3. Nasal harmony	–	to.ʰɲã
4. Surface form	[to.ʰto]	[to.ʰɲã]

(van 't Veer et al., 2025)

2.2.2 – Accoustics of pre-aspiration

Accoustically, pre-aspiration is described in both Bruil (2014) and van 't Veer et al. (2025) as a post-vocalic glottal fricative [h]. However, data collected during my own fieldwork trip in the Siona speaking community of Sototsiaya in June 2025 seems to indicate instead a broader window of possible realizations. Additionally to a post vocalic [h] realization, pre-aspiration is also realized as a post-vocalic glottal stop [ʔ], as creakiness on the vocal segment itself (preceding the pre-aspirated syllable), and as a general lengthening of the syllable (preceding the pre-aspirated syllable). Glottal stop and creakiness realizations of pre-aspiration seem more commonplace in relaxed speech, and more prominently used by younger speakers. The only phonological conditioning I found was that glottal stops [ʔ] are by far more common than other realizations before /s/ segments.

Furthermore, not all pre-aspirations are equal in acoustic salience. Perceptually, second syllables in words appear relatively more strongly pre-aspirated than fourth syllables, with pre-aspiration being less often reduced to glottal stops or creakiness. An attempt at investigating this acoustically was made using the software Praat (Boersma & Weenink, 2025), however pre-aspiration turned out difficult to observe reliably on spectrograms due to the poor audio quality and ambient noise of the jungle fauna and rain. As such, this claim is of an impressionistic nature, and still needs to be verified. See Figures 1 and 2 for spectrograms of words featuring two pre-aspirated syllables.

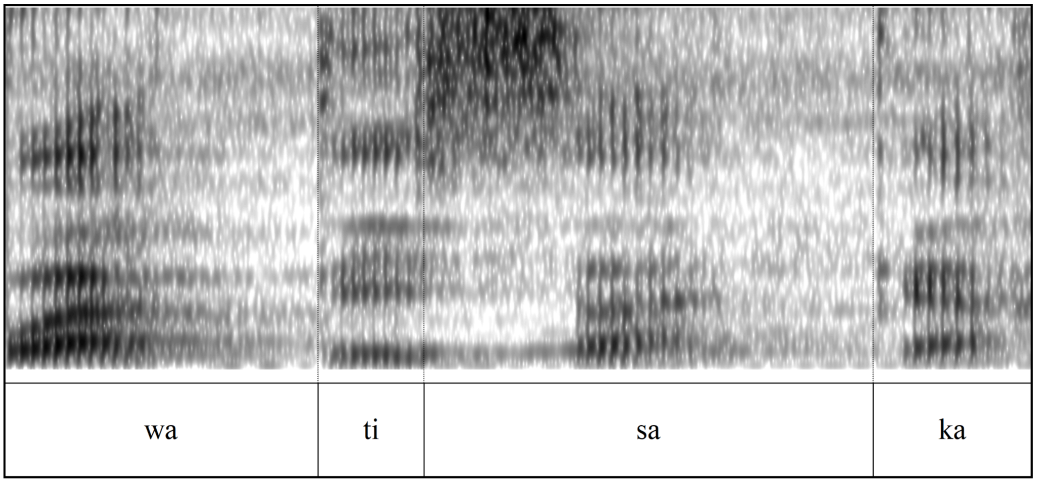


Figure 1: Spectrogram of the word *watisaka*

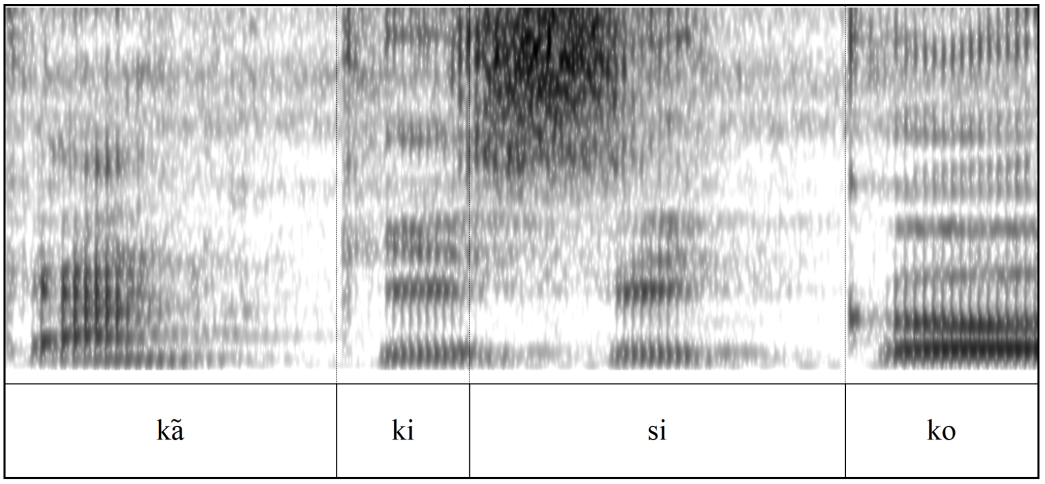


Figure 2: Spectrogram of the word *kãkisiko*

2.2.3 – Pre-aspiration, unvoiced vocalic lengthening or coda epenthesis?

A pre-aspiration analysis (CV.^{h}CV) possesses multiple advantages compared to the unvoiced vocalic lengthening (CV̥V.CV) analysis made in related languages, or a coda-epenthesis (CV^{h}.CV) analysis (van 't Veer et al., 2025). For the purposes of explaining the advantages of pre-aspiration compared to other analyses, both the unvoiced vocalic lengthening and coda epenthesis are considered the same, referring to them using the latter term, as they have the same implications.

The first advantage of a pre-aspiration analysis is that it more readily explains the failure of the glottal obstruents {p^{ʔ}, t^{ʔ}, k^{ʔ}, k^{wʔ}} to trigger it than the other two analyses. Indeed, the blocking of pre-aspiration by glottal obstruents could be explained by the existence of a constraint blocking the linking of multiple laryngeal specifications to one root node (van 't Veer et al., 2025). Both unvoiced vocalic lengthening and coda epenthesis fail to address this issue.

The second advantage of a pre-aspiration analysis compared to unvoiced vocalic lengthening or coda-epenthesis is that it respects ES's bimoraic stem constraint (Bruil, 2014; van 't Veer et al., 2025), stating that all stems must be at least and at most two moras long. On the other hand, the unvoiced vocalic lengthening analysis can be rejected as it leads to violations of the constraint (13), and so too can the coda epenthesis analysis, as it posits an

ambiguous status of epenthesized [h] with regards to moras. Indeed, following a coda epenthesis analysis, the epenthesized [h] does not appear to count as moraic when following short vowels (14), but does following long vowels (15). Considering the absence of an explanation for the contradicting status of epenthesized [h] in regards to moras, and the violations of the bimoraic stem constraint by the unvoiced vocalic lengthening, the pre-aspiration analysis is considered superior by default.

(13) Unvoiced vocalic lengthening violates the bimoraic constraint
|wa.ˈti|
*/waḁ.tí/ (3 mora stem)
‘Bad sprit’

(14) Epenthesized [h] does not count as moraic when following short vowels
|sa.ˈsio|
/sa^{h}.sío/ (2 mora stem)
‘She will go’

(adapted from Bruil, 2014: 100)

(15) Epenthesized [h] counts as moraic when following long vowels
|sa:.ˈsio|
/sa:.sío/ (2 mora stem) */sa:.^{h}sío/ (3 mora stem)
‘She will take’

(adapted from Bruil, 2014: 100)

2.2.4 – The relationship assumption

Comparatively, pitch accent has received considerably less attention than pre-aspiration. It is briefly mentioned in Bruil (2014) in a section on stress and tone, and van 't Veer et al. (2025) state in a footnote that pitch accent being word final would be problematic for a trochaic analysis compared to the iambic analysis of ES they make. Though it is not explicitly said, such a statement indicates that the authors adopt a Metrical Theoric view of stress assignment, meaning that the foot level phenomenon of pre-aspiration is linked to the word level phenomenon of pitch accent. Coupled with their attribution of iambic pre-aspiration to stress, the picture drawn of ES prosody is that ‘stress’ is equivalent to iambic pre-aspiration, and pitch accent depends on said ‘stress’ for its assignment. It is of my opinion that the terminology employed is unfortunately quite vague. As such the next chapter focuses on defining the phenomena under the scope of this study, before testing the assumption that pitch accent is indeed linked to pre-aspiration using Metrical Theory in §4 and Accent First Theory in §5.

3 – The objects of study

In the analyses regarding ES prosody made in the following two chapters, two surface phenomena are discussed: pitch accent and pre-aspiration. In an effort to bring terminological clarity, the current chapter concerns itself with identifying and defining those two prosodic phenomena, under the terms of prominence for pitch accent, and of either prominence or foot mediality for pre-aspiration, depending on the analysis.

Prominence as used in the current paper is defined in Terken & Hermes (2000) as “a linguistic entity is prosodically prominent when it stands out from its environment by virtue of its prosodic characteristics”. Prominence is an underlying relative property of a syllable compared to others, which is phonetically implemented in language-specific ways[5] (Cangemi & Bauman, 2020). Stress is for instance the most well known case of surface realization of prominence, where one syllable is fortified and thus made more salient than others within the domain of the word.

In the case of pitch accent, I argue in this chapter that the phenomenon can best be described as surface realization of prominence at the word-level, then called *accent*. In the case of pre-aspiration however, I argue that two analyses are equally applicable. The first analysis is that of *rhythm*, which I define as alternating sub-lexical prominence (σσ̲σσ̲σ). The second analysis of pre-aspiration applied is that of a *foot medial* phenomenon, without the need to involve prominence ($\sigma^h\sigma$)($\sigma^h\sigma$).

3.1 – Pitch accent as accent

Pitch accent is a term often used to describe prosodic characteristics of languages, yet one that is rarely defined. Hyman (2009) indeed states that a great issue surrounding discussions on pitch accent is that concepts of prototypes and definitions are usually conflated: Definitions include minimal requirements for a phenomenon to be considered as ‘X’, whereas a prototype is rather a list of the characteristics most easily attributable to ‘X’. Pitch accent as Hyman argues is an example of a phenomenon which is often analyzed in terms of prototypes but rarely defined, unlike stress and tone.

The main reason for a general lack of definitions, is that pitch accent is a system which finds itself in many ways in between stress and tone, sharing characteristics with both (McCawkley, 1970). Similarly, Hyman (2009) demonstrates that pitch accent prototypes possess characteristics which are not unique, but also found in either stress or tonal systems, questioning the need for a separate prosodic category. As such, recent attempts have been conducted at redefining pitch accent. In this section,the two dominant ones are examined. Borrowing Van der Hulst’s (2011) terminology, they are referred to as the “extreme accent view” and the “extreme tone view”.

3.1.1 – The extreme tone view

Developping on the idea that pitch accent possesses characteristics also found in both stress and tonal systems, Hyman (2009) advocates for a view of pitch instead as a “defective tonal system”. Indeed, according to the author, the notion of pitch accent should be altogether discarded, arguing that a tonal analysis is always possible instead. Hyman (2009) addresses four characteristics common to stress systems which are also found in pitch accent analyses, arguing that those do not exclude a tonal analysis.

The four characteristics usually attributed to pitch accent or stress, yet still analyzable in terms of tone according to Hyman are as follows. The first two, *obligatoriness* and *culminativity* are common criteria used to define stress (Hyman 2006). The former means that tone must occur at least once per word and the latter means that tone must occur at most once per word. Taken together, those two characteristics mean ‘every word features

[5] Common phonetic correlates of prominence include among others intensity, duration or phonotactic restrictions.

only one tone'. The third characteristic discussed is *privativity*, meaning that the studied system features a tone vs. no tone contrast, rather than a high tone vs. low tone contrast. Finally, the fourth characteristic is called *metricality*, meaning that tones are regularly attracted to specific positions in the word (e.g. penultimate syllable).

The extreme tone view thus allows for a tonal analysis of phenomena previously described as pitch accent, through a broadening of the notion of 'tone'. Tones in this view may be sensitive word-level phonology, through assessment of the presence/absence of tones in the word (obligatoriness/culminativity)[6] as well as position within the word (metricality).

3.1.2 – The extreme accent view

In opposition to the extreme tone view, Van der Hulst (2011, 2012) analyses pitch accent as a type of accent, claiming that any language which features a binary pitch contrast can be analyzed in terms of accent, even when multiple or no syllables are marked in a word (Tokyo Japanese: Kubozono 2011; Somali: Hyman 1981, Banti 1988; Haya: Hyman & Byarushengo, 1984). The extreme accent view thus appears on the surface to violate the above mentioned constraints of culminativity and obligatoriness. This issue is solved through the analysis of pitch and stress as two different surface realizations of accent. Accent on the other hand, is the underlying form sharing by both pitch accent and stress accent. The culminativity and obligatoriness restrictions in this view solely apply to stress accent but not necessarily to pitch accent (adherence to the culminativity and obligatoriness constraints is a language-specific feature).

Further, Van der Hulst (2012) addresses the issue of descriptions of pitch as a correlate of stress, which could arguably make it a phonetic correlate of stress accent instead of pitch accent. He argues however that previous studies (Mol and Uhlenbeck, 1956; Fry, 1955) that have argued that pitch can be a realization of stress did so only because examined words were produced in isolation. However, when uttered in out-of-focus positions pitch was not found as a major cue for stress at all. Thus only languages that specifically use pitch as a main cue to making a syllable salient can be considered pitch accent languages.

As a summary, the extreme accent view places pitch accent as a surface realization of accent, on the same level as stress accent. Doing so, pitch accent does not breach the culminativity and obligatoriness constraints, as it is argued that they only apply to stress accent. Thus, following the extreme accent view, any language which features a binary pitch contrast can be analyzed in terms of accent. Table 1 sums up the differences of stress accent and pitch accent following the extreme accent view.

[6] Hyman's argument for 'privativity' seems to be that privative systems lead to violations of obligatoriness/culminativity constraints, making a stress analysis impossible.

Underlying form	Phenomenon	Accent	
Surface form	Phenomenon	Stress accent	Pitch accent
	Phonetic correlates	Duration, intensity, phonotactic restrictions..	Fundamental frequency (F0)
	Culminativity/obligatoriness	Y	Y/N

Table 1: Differences between stress accent and pitch accent following the extreme accent view.

3.1.3 – Evaluation

Both of the above described views offer strong arguments for a tonal or accentual analysis instead of a 'middle point' pitch accent one. In order to account for the great variety in pitch accent systems (some featuring culminativity, some featuring obligatoriness, some featuring both, some featuring none), the extreme tone view has to posit that tone can be obligatory and culminative. On the other hand, the extreme accent view has to posit that accent is not necessarily obligatory and culminative.

However, I believe that a tone view of pitch accent lacks motivation in some respect. Indeed, rewording a quote from Van der Hulst (2011), one could ask: What makes, within the context of a word, a predictable and non-contrastive *tone/pitch* different enough from for instance *length* or *tenseness* to warrant a whole different kind of analysis other than an accentual one? If one of the latter two properties were found word-finally in a language without serving any contrasts, an accentual analysis would almost always prevail. In the absence of a different behavior specific to tones, an accentual analysis should prevail by default due to its functional identicality to other types of accents.

As Bruil (2014: 86) states, pitch accent is strictly word-final, and as such it is not contrastive. Hence, I argue that an accentual view in ES is applicable, and should be, as it provides a unified framework through which comparison with other accent types is possible. On the other hand, a tonal view is unable to study the relationship of pitch accent with other phenomena to the same extent.

3.2 – Pre-aspiration: rhythm or foot mediality?

The second prosodic phenomena found in ES this thesis focuses on is pre-aspiration, previously analyzed as stress in van 't Veer et al. (2025). As a reminder, pre-aspiration only occurs on even numbered syllables (counting left to right), no matter the length of the word, and on non-glottal obstruents provided they follow a short vowel (16-19). Unlike pitch accent which can best be analyzed way as accent, pre-aspiration in ES is a phenomenon which can arguably be analyzed either as rhythm or as foot mediality.

(16) Rhythmic analysis of pre-aspiration, two-syllable word

σσ
|to.ˈto|
/to.^{h}tó/
'board'

(adapted from Bruil, 2014: 105)

(17) Rhythmic analysis of pre-aspiration, three-syllable word
σσ̲σ
|sa.si̲.ˈpʔi|
/sa.ʰsi.pʔí/
‘He will/can go’

(adapted from Bruil, 2014: 197)

(18) Rhythmic analysis of pre-aspiration, four-syllable word
σσ̲σσ̲
|kʷaʔ.ku̲.ma.ˈka̲|
/kʷaʔ.ku.maˑʰká/
‘...When it was semi-cooked’

(adapted from Bruil, 2014: 165)

(19) Rhythmic analysis of pre-aspiration, five-syllable word
σσ̲σσ̲σ
|a.tʃa̲.si.ki̲.ˈpʔi|
/aˑʰtʃa.siˑʰkɨ.pʔí/
‘The one (m) listening’

(adapted from Bruil, 2012: 20120918elicr004)

3.2.1 – Pre-aspiration as rhythm

One analysis of pre-aspiration that can be made is that it is the phonetic correlate of rhythm: All even-numbered syllables are rhythmically prominent, and said prominence is phonetically realized when the syllable begins with a non-glottal obstruent and follows a short vowel, formalized in a rule based way in (20). Rhythmic prominence is depicted by an underline on the segment.

(20) Rule based depiction of phonetic realization of rhythmic pre-aspiration
C̲[+OBSTRUENT, +CONS., -CONSTR.GLOT] → ʰC̲ / v[+SHORT] _

3.2.2 – Pre-aspiration as foot mediality

The other analysis of pre-aspiration that can be made is of a foot medial phenomenon. Prosodic feet are usually binary groupings of syllables traditionally used to explain accent assignment in Metrical Theory (Liberman & Prince, 1977; Hayes 1981, 1995). Within the foot, one syllable is more prominent than the others and thus heads the whole foot. However, following Culhane (2023), the foot is a prosodic structure not only relevant for accent assignment, but also a domain in itself with its own phonotactic restrictions and within which certain processes may apply without the need to refer to prominence. It is unclear at this stage whether such ‘phonotactic’ feet are identical to Metrical Theoric ‘headed’ or ‘rhythmic’ feet, or if languages might allow for both types of feet to exist alongside each other with different distributions.

Examples of foot-based phonotactic restrictions include, languages such as Central Lembata (Fricke 2020: 42–49), Meto (Edwards 2020: 108), and Waima’a (Culhane 2022)

which all allow consonant clusters only in foot edges, whereas vowel sequences only are allowed foot internally. Waima'a also feature foot-sensitive reduplication, where first syllables of word final feet is reduplicated and placed to the left of the foot. Consider for instance |ber(tama)| → /bertatama/. As an example of a foot-internal process, Culhane analyzes Dagbani debuccalization (Hudu 2018: 208) as a foot medial process, such as in /pí.**s**í.tá/ → [(pí.**h**í).tá], foot depicted in between parentheses. The phonological foot may also select affixes. Pitjantjatjara for instance features two inflectional classes, one used for words with an even number of syllables, which are entirely footed, the other used for words featuring an odd number of syllables which are only partially footed (Wilmoth and Mansfield, 2021).

In order to apply a foot analysis to ES, I first define the concept of foot medial process, left ambiguous in Culhane (2023): A foot medial process is a process which only targets segments at the juncture between neighboring syllables within the same foot. Consider for instance the following word in a language X, foot segments are depicted in bold: (C**V**.**C**V).(CV**C**.**C**V).

In the case of ES, parsing feet from left to right ($\sigma^{h}\sigma$)(σ) instead of right to left *(σ)($\sigma^{h}\sigma$) (21), a foot medial analysis correctly accounts for the alternating occurrence of pre-aspiration (22-25). Final syllables of words with an odd number of syllables are depicted as fully footed for visual purposes. Finally, the foot-medial view of pre-aspiration is formalized in a rule-based manner in (26).

(21) Right to left foot parsing does not predict pre-aspiration accurately
*|(σ)(σσ)|
*/(ka).(ka.^{h}sió)/
'She will enter'
(adapted from Bruil 2011)

(22) Foot medial analysis of pre-aspiration, two syllable word
|(σσ)|
/(to.^{h}tó)/
'board'
(adapted from Bruil, 2014: 105)

(23) Foot medial analysis of pre-aspiration, three syllable word
|(σσ)(σ)|
/(sa.^{h}si).(p^{ʔ}í)/
'He will/can go'
(adapted from Bruil, 2014: 197)

(24) Foot medial analysis of pre-aspiration, four syllable word
|(σσ)(σσ)|
/(k^{w}aʔ.ku).(ma.^{h}ká)/
'...When it was semi-cooked'
(adapted from Bruil, 2014: 165)

(25) Foot medial analysis of pre-aspiration, five syllable word
|(σσ)(σσ)(σ)|
/(a.^{h}tʃa).(si.^{h}kɬ).(p^{ʔ}í)/
'The one (m) listening'
(adapted from Bruil, 2012: 20120918elicr004)

(26) Rule based depiction of foot medial pre-aspiration
C[+OBSTRUENT, +CONS., -CONSTR.GLOT] → ^{h}C / (v[+SHORT]._)

To clarify, the rhythmic and foot structure approaches are considered thus far equally applicable to ES. In order to determine which approach is more beneficial, pitch accent needs to be added into the equation and explore its relationship with pre-aspiration.

3.3 – Summary

In this chapter, clarified the nature of the two phenomena under the scope of this study was clarified. The first one is the word final pitch accent, which is analyzed here as the surface realization of accent instead of as a tone. The second is pre-aspiration, which occurs on non-glottal obstruents in even numbered syllables following short vowels, which can be both analyzed as surface rhythm, or as a foot medial process. Table 2 summarises the conclusions on the nature of the phenomena under this study.

	Pre-aspiration		Pitch accent
Surface distribution	Non-glottal consonants following short vowels. On even numbered syllables.		Ultimate syllable
Analysis	Surface realization of rhythm (σσ̲σσ̲)	Foot medial process (σ^{h}σ)(σ^{h}σ)	Surface realization of accent

Table 2: The two studied prosodic phenomena, including two different analyses of pre-aspiration

4 – Metrical Theory

The current chapter concerns itself with analyses of the relationship between pre-aspiration and pitch-accent through the framework of Metrical Theory. Though not explicitly stated, it appears this is the theoretical framework employed in van 't Veer et al. (2025).

4.1 – Theoretical background

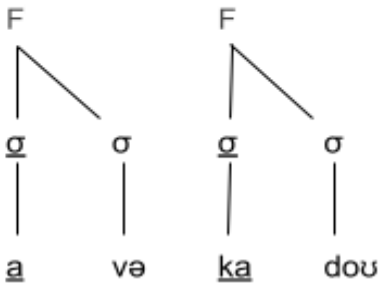
Metrical Theory (Liberman & Prince, 1977; Hayes 1981, 1995) (or 'MT') is a theory of accent assignment[7] where words' accent assignment is based on rhythm. In order to do so, rhythm is computed in MT using headed (or rhythmic) feet, meaning that one syllable is more

[7] Traditionally worded rather as a theory of stress assignment. However I apply throughout the paper Van der Hulst's terminological split between stress and accent, originally absent from MT discussions.

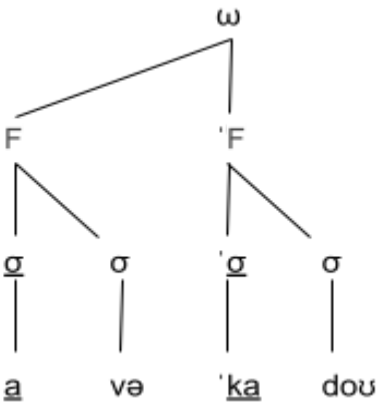
prominent than the other(s) within the foot (unlike the phonotactic foot which introduced in the previous section). After parsing syllables into headed feet, primary accent arises through promotion of a foot head. The prominent syllable within the foot thus goes from heading only its foot to heading the whole word as well. In this view, accent is considered a mere manifestation of rhythm (Hayes, 1995: 8). Example (27) shows how accent assignment happens in two steps in MT with the English word ‘avocado’. Rhythmic prominence is underlined, while word-level prominence is represented with the traditional accent mark. In the first step the syllables of [a.və.ka.doʊ] are paired into left-headed feet, (a.və) and (ka.doʊ), without a determined primary accent yet. In the second step, the head |ka| of the (ka.doʊ) foot is promoted into heading the whole word, yielding /a.və.ˈka.doʊ/.

(27) Two steps of accent assignment

Step 1: Foot parsing

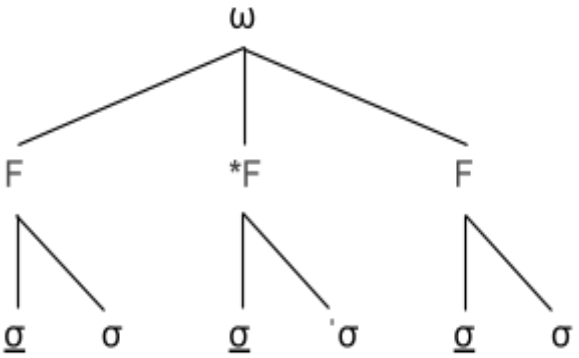


Step 2: Foot promotion



Most important for a successful MT analysis, is that accent in this framework is must depend on rhythm. In other words, this means that there cannot exist a word featuring a syllable that is prominent at the word level, but not at the foot level. This theoretical impossibility is exemplified in (28), with a word featuringa grammatically impossible foot, marked with an asterisk.

(28) A foot cannot feature primary accent on a non-rhythmically prominent syllable.



4.2 – MT parameters

Languages differ in the parameters they use in order to parse syllables into feet and feet into words. This section lists parameters used in MT based on Gussenhoven & Jacobs (2017). Additionally, I introduce a split of available parameters into 'foot parsing' and 'accent placement' subgroupings, inspired by AFT's 'accent domain' and 'accent placement' distinction. Finally, I add the new parameter PROMINENCE (Y/N) in order to differentiate phonotactic feet from headed feet. Parameters which may be inactive are depicted in between parentheses. When active, parameters may either have a binary (Y/N) or directional (R/L) value.

(29) List of all major MT parameters

Foot Parsing	(DIRECTIONALITY) (R/L)
	(DOMINANCE) (IAMBIC/TROCHAIC)
	(EXTRAMETRICALITY) (Y/N)
	ITERATIVITY (Y/N)
	PROMINENCE (Y/N)
	WEIGHT-TO-ACCENT (Y/N)
Accent Placement	(BOUNDED) (Y/N)
	(EDGE) (R/L)

The foot parsing parameters, relevant for rhythm assignment, are the following: (DIRECTIONALITY) (R/L) refers to the direction of foot parsing in a language, either right to left (L) or left to right (R). (DOMINANCE) (IAMBIC/TROCHAIC) refers to whether feet are right-headed (IAMBIC) or left-headed (TROCHAIC). (EXTRAMETRICALITY) (Y/N) refers to whether finally parsed syllables are footed and thus accentable (Y) (σσ)(σσ)(ˈσ) or not (N). Extrametrical syllables are called degenerate feet, and non-metrical syllables are called stray syllables. ITERATIVITY (Y/N) refers to whether foot parsing occurs throughout the whole word (Y), or only in the foot bearing primary accent (N). I add the parameter PROMINENCE (Y/N), which refers to whether feet are prominent or not, allowing for a distinction between phonotactic feet (N) and headed feet (Y). Finally, WEIGHT-TO-ACCENT (Y/N) refers to whether feet construction is weight sensitive (Y) or not (N)

The accent placement parameters are the following: BOUNDED (Y/N) refers to whether accent is positioned close to word edges (Y), or if it may occur anywhere in the word (N). EDGE (R/L) refers to whether primary accent falls on the right-most (R) or left-most (L) foot in the word.

4.2 – A metrical account of Ecuadorian Siona

4.2.1 – The rhythmic view within MT

Following the rhythmic view of pre-aspiration, foot structure in ES is iterative, and feet are parsed left to right and are iambic (van 't Veer et al., 2025). Accent placement does not

appear to be sensitive to syllable weight and extrametrical syllables are allowed, in accordance with the strictly word-final distribution of pitch accent (Bruil, 2014: 86). Finally, accent is bounded and found on the right edge of the word. See Table 3 for the parametric values used in ES.

Foot Parsing	DIRECTIONALITY (R) DOMINANCE (IAMBIC) EXTRAMETRICALITY (Y) ITERATIVITY (Y) PROMINENCE (Y) WEIGHT-TO-ACCENT (N)
Accent Placement	BOUNDED (Y) EDGE (R)

Table 3: MT parameters in ES, rhythmic account of pre-aspiration

This considered, words with an even number of syllables (or 'even-syllabled words') follow the pattern (σσ̲)(σσ̲)...(σˈσ̲) (30, 31). In this case, the MT assumption of accent occurring on prominent rhythmic beats is fulfilled.

(30) Accent occurs on rhythmic beats in even-syllabled words
(σˈσ̲)
|mo.ˈtor|
/mo.ʰtór/
'motor'

(Bruil, 2014: 104)

(31) Accent occurs on rhythmic beats in even-syllabled words
(σσ̲)(σˈσ̲)
|kʷaʔ.ku.ma.ˈka|
/kʷaʔ.ku.ma.ʰká/
'...When it was semi-cooked'

(adapted from Bruil, 2014: 165)

In words with an odd number of syllables (or 'odd-syllabled words') however, an analysis of the relationship between accent and rhythm is less straightforward. Considering that in MT accented syllables have to be rhythmically prominent, one would expect to find pre-aspirated syllables even on degenerate feet, which bear pitch accent, as in the following (σσ̲)(σσ̲)...(ˈσ̲). However, when considering the data, it appears that pitch accent does not co-occur with pre-aspiration in those cases, as seen in (32, 33).

(32) Pre-aspiration does not occur on pitch accented syllables in odd-syllabled words
(σσ)(ˈσ)
*/a.ʰpa.ʰsí/
/a.ʰpa.sí/
‘sapote (fruit)’

(Bruil, 2014: 102)

(33) Pre-aspiration does not occur on pitch accented syllables in odd-syllabled words
(σσ)(ˈσ)
*/ka.ʰka.ʰsíó/
/ka.ʰka.síó/
‘She will enter’

(Bruil 2011)

Based on (31) and (32), it appears that the MT prediction of pitch accent occuring on pre-aspirated syllables in ES is not verified when only examining the surface form. However, absence of surface pre-aspiration does not necessarily mean absence of underlying rhythm, since it is already established that certain phonological contexts block phonetic realization. Assuming underlying presence of rhythm on degenerate feet, two possible explanations for lack of surface realization of pre-aspiration on degenerate feet are discussed below, and then rejected.

The first possible explanation for lack of surface rhythm on degenerate feet is that pitch accent interacts with pre-aspiration and blocks its phonetic implementation. This explanation can be rejected since word final syllables are attested in the language as can be seen in (34).

(34) Pitch accent does not block pre-aspiration
|a.ˈkɨ|
/a.ʰkí/
'the one (m) from'

(Bruil 2014: 103)

The second possible explanation is that for rhythm to be realized on the surface, the preceding syllable needs to be weak, additionally to the other requirement that the onset of the syllable needs to be a non-glottal obstruent following a short vowel (35, 36). This explanation solves the discrepancy observed in odd-syllabled words.

(35) Realization of rhythmic prominence is blocked when the preceeding syllable is also rhythmically prominent
(σσ)(ˈσ)
|ka.ka.ˈsio|
/ka.ʰka.síó/
‘She will enter’

(adapted from Bruil 2011)

(36) Realization of rhythmic prominence is blocked when the preceeding syllable is also rhythmically prominent

(σσ̲)(ˈσ̲)
|a.pa̲.ˈsí̲|
/a.ʰpa.sí/
‘sapote (fruit)’

(Bruil, 2014: 102)

However, it must be noted that conditioning based on a neighboring syllable’s prominence instead of just one’s own is, to the extent of my knowledge, empirically unsupported. Furthermore, such a concept seems to go against MT’s foot based explanation for rhythmic prominence, as syllables in MT are prominent only relatively to other syllables within the same foot. Hence, requiring the preceding syllable to be rhythmically weak across foot boundaries in ES would be contradictory with the premises of the very framework that is being used. This considered, it appears that the fact that pre-aspiration is never observed in degenerate feet cannot readily be explained following an MT rhythmic account.

4.2.2 – The foot medial view within MT

Following the foot medial view (Culhane, 2023), pre-aspiration is a foot medial phenomenon instead of a sub-lexical prominence phenomenon. However, considering that MT requires rhythm in order to compute accent, the possibility that ES does not feature rhythm at all is excluded. As such two analyses are possible. The first one is that phonotactic feet which govern pre-aspiration are also headed, as in (σʰσ̲)(σʰσ̲)(ˈσ̲). The second possibility is that ES features two-foot inventories: one phonotactic foot inventory $(\sigma^{h}\sigma)_{P}(\sigma^{h}\sigma)_{P}(\sigma)_{P}$, as well as one headed foot inventory $(\sigma\underline{\sigma})_{H}(\sigma\underline{\sigma})_{H}(\text{ˈ}\underline{\sigma})_{H}$. In both of those cases, foot headedness would be underlying only, meaning that rhythm has no phonetic substance. This can be argued to be a drawback from a computational perspective, as rhythm would merely be a theoretical tool without which accent cannot exist. Theories like Accent First Theory in this case are computationally more efficient, as they do not require rhythm in order to place accent (more on this is §5)

Following the one-foot inventory account, accent assignment and rhythm assignment is done using the same parametric values as in Table 3. Final syllables are always both accented and rhythmically prominent at once. This time however, pre-aspiration is computed foot medially regardless of prominence, and as such final syllables in odd-syllabled words do not constitute an issue (37, 38). Foot medial segments are depicted in bold.

(37) Pre-aspiration and accent in an even-syllabled word, one-foot inventory account

(σʰσ̲)(σˈʰσ̲)
|(t͡sɨ̶.**t͡sɨ**).(d**a**.ˈ**k**a̲)|
/(t͡sɨ.ʰt͡sɨ).(da.ʰká)/
‘cold water’

(38) Pre-aspiration and accent in an odd-syllabled word, one-foot inventory account
(σ^{h}σ)(ˈσ)
|(k**a.k**a).(ˈsio)|
/(ka.^{h}ka).(ˈsíó)/
'She will enter'

(adapted from Bruil 2011)

Following the two-foot inventory account (phonotactic foot and headed foot), accent assignment is done through headed feet. Similarly to the one-foot account, prominence in headed feet would be underlying only. The simplest parametric configuration for headed feet (the one where the least amount of parameters had to be specified) is shown in Table 4, following the pattern σ(σˈσ)$_{H}$ or σσ(σˈσ)$_{H}$. Multiple parameters can be omitted, as parsing a single foot on the right edge of the word is sufficient to describe accent falling on the ultimate syllable.

Foot Parsing	DIRECTIONALITY (L)
	DOMINANCE (IAMBIC)
	ITERATIVITY (N)
	PROMINENCE (Y)
	WEIGHT-TO-ACCENT (N)

Table 4: MT parameters for headed feet in ES, two-foot inventory account

For phonotactic feet, responsible for pre-aspiration, the parametric values are shown in Table 5. Most notably, accent placement parameters are inactive, and PROMINENCE is negative.

Foot Parsing	DIRECTIONALITY (R)
	ITERATIVITY (Y)
	PROMINENCE (N)
	WEIGHT-TO-ACCENT (N)

Table 5: MT parameters for phonotactic feet in ES, two-foot inventory account

Examples (39, 40) showcase how the two-foot inventory correctly accounts for both even-syllabled words or odd-syllabled words. For visual purposes, phonotactic feet and headed feet processes are represented separately.

(39) Pre-aspiration and accent in an even-syllabled word, two-foot inventory account

(39a) Headed feet	(39b) Phonotactic feet
$(\sigma\underline{\sigma})_H(\sigma'\underline{\sigma})_H$	$(\sigma^h\sigma)_P(\sigma^h\sigma)_P$
$\|(\widehat{\text{tsɨ}}.\widehat{\underline{\text{tsɨ}}}).(\text{da}.'\underline{\text{ka}})\|$	$\|(\widehat{\text{ts}}\textbf{ɨ}.\widehat{\textbf{ts}}\text{ɨ}).(\text{d}\textbf{a}.\textbf{k}\text{a})\|$
$/(\widehat{\text{tsɨ}}.\widehat{\text{tsɨ}}).(\text{da.ká})/$	$/(\widehat{\text{tsɨ}}.^h\widehat{\text{tsɨ}}).(\text{da}.^h\text{ka})/$

→ /$\widehat{\text{tsɨ}}.^h\widehat{\text{tsɨ}}$.da.^{h}ká/
'cold water'

(40) Pre-aspiration and accent in an odd-syllabled word, two-foot inventory account

(40a) Headed feet	(40b) Phonotactic feet
$(\underline{\sigma})_H(\sigma'\underline{\sigma})_H$	$(\sigma^h\sigma)_P(\sigma^h\sigma)_P$
$\|(\underline{\text{ka}})_H.(\text{ka}.'\underline{\text{sio}})_H\|$	$\|(\text{k}\textbf{a}.\textbf{k}\text{a})_P.(\text{sio})_P\|$
$/(\text{ka})_H.(\text{ka.sió})_H/$	$/(\text{ka}.^h\text{ka})_P.(\text{sio})_P/$

→ /ka.^{h}ka.sió/
'She will enter'

(adapted from Bruil 2011)

4.2.3 – Conclusion

To summarize, the rythmic view of pre-aspiration offered a weak description of the distribution of pitch accent and pre-aspiration in ES. This explanation required final syllables of odd-syllabled words to be sensitive to the level of prominence of previous syllables, across foot boundaries. Such an explanation is considered weak as it is both empirically unattested (to the extent of my knowledge), and contradictory with theorical assumptions of MT.

On the other hand, the foot-medial view of pre-aspiration offered a stronger description of the distribution of pitch accent and pre-aspiration in ES. In this case, two analyses co-exist. The one-foot inventory account assumes that there can only be one-foot inventory in ES. Feet are headed in order to bear pitch accent, however pre-aspiration is not a product of foot-headedness, but rather a product of foot mediality. The two-foot inventory account assumes on the other hand that pre-aspiration and pitch accent are computed using two separate foot inventories. Pre-aspiration is computed using phonotactic feet, whereas pitch accent is computed using headed feet.

Thus far, the one-foot inventory account of foot-medial pre-aspiration is favored. It does not suffer from both theoretical and empirical flaws as the rhythmic account does, and computationally it is simpler than the two-foot inventory account, as it needs to parse only one set of foot instead of two.

5 – Accent First Theory

The current chapter concerns itself with analyses of pre-aspiration and pitch-accent through the framework of Accent First Theory. This is a more recent prosodic framework which attempts to resolve common issues with Metrical Theory such as reliance of accent on rhythm for its assignment.

5.1 – Theoretical innovations

The main attempt at separating accent and rhythm into different algorithms considered in this thesis is Accent-First Theory, or 'AFT' (Van der Hulst, 1996, 2009, 2012, 2014a, 2014b; Goedemans & Van der Hulst, 2014; Bogomolets 2020, 2023). AFT succeeds Metrical Theory and deviates from it in one major regard.

The most important theoretical change brought by AFT compared to MT is that, as the name of the theory suggest, word accent is computed before rhythm. MT on the other hand computes rhythm first, and then primary accent is promoted from it. This reverse ordering to that of MT is motivated by the observation that accent is stored lexically in a multitude of cases, whereas rhythm always seems to be a post-lexical process (van der Hulst, 2012). This reverse ordering has one major implication: While in MT, rhythm is required in order to have accent, AFT requires any kind of accent in order to have rhythm. The way this functions is that rhythm in AFT requires an 'anchor', from which it alternates. This anchor can either be the primary or secondary accent (Bogomolets, 2020, 2023)[8].

The fact that rhythm may be anchored on a syllable that is not the primary accent is one of the main draws of AFT, as it permits us to describe languages where primary accent does not co-occur with rhythmic beats, such as is the case in ES. Example (41a, b) from Passamaquoddy and (42a, b) from South Conchucos Quechua showcase primary and secondary accent as respective anchors for rhythm: Rhythm only binarily alternates from primary accent in Passamaquoddy and secondary accent in South Conchucos Quechua. Rhythmically prominent syllables are underlined, primary accent denoted by upper accent marks, and secondary accent evidenced by lower accent marks.

(41) Primary accent is the anchor for rhythm in Passamaquoddy

(41a) /ˌteh.sah.kwa.pa.sol.ˈti.ne/ 'Let's walk around on top.'

(41b) /ˌwi.coh.ke.ˈke.mo/ 'He helps out'

(Passamaquoddy, adapted from LeSourd, 1988: 140, 143)

(42) Secondary accent is the anchor for rhythm in South Conchucos Quechua

(42a) /ˈtu.ʃu.ku.ˌna.qa̰/ 'dancers'

(42b) /ˈwa.ra:.ka.mun.qa.ˌna.tʃḭ/ 'Hopefully it will appear at dawn'

(South Conchucos Quechua, adapted from Hintz, 2006: 4)

5.2 – AFT Parameters

In order to describe accent and rhythm assignment in languages, AFT proposes two sets of parameters: one for primary and secondary accent, one for rhythm. This section is entirely devoted to AFT parameters and explaining how they function.

[8] Previous accounts consider rhythm as alternating from either accent or an edge-prominent syllable found at the polar opposite of the accent. Bogomolets (2020, 2023) instead considers edge-prominence as a type of secondary accent, and thus rhythm as alternating from either primary or secondary accent.

5.2.1 – Accent parameters

Accent assignment parameters are divided into two types. The accent domain parameters select a domain inside the word within which the accent will be placed according to the accent placement parameters. Similarly to MT, AFT parameters when active take either a binary (Y/N) value, or a directional (R/L) value.

(43) List of all accent parameters in AFT

Accent Domain	(BOUNDED) (R/L)
	(SATELLITE) (R/L)
Accent Placement	WEIGHT-TO-ACCENT (Y/N)
	(SELECT) (R/L)
	(DEFAULT) (R/L)
	OBLIGATORINESS (Y/N)
	CULMINATIVITY (Y/N)

The accent domain parameters are defined as follow: (BOUNDED) (R/L) creates a bisyllabic domain at a right or left edge of the word. When inactive, this parameter means that accent may occur anywhere within the word, and not only on the edges. (SATELLITE) (R/L), when active, assigns an extra syllable (called a satellite) to the right or left edge of the domain, which then becomes part of the domain. Satellites on word-edges are called external satellites and cannot be accented, whereas word internal satellites on the other hand are fully accentable. The AFT notion of satellites can be taken as equivalent to MT's notions of stray syllables when the satellite is external, and of degenerate feet when the satellite is internal.

The accent placement parameters are as follow: WEIGHT-TO ACCENT (Y/N) determines whether heavy syllables attract accent. If positive, all heavy syllables are considered accentual candidates. (SELECT) (R/L) assigns the accent to the right-most or left-most accent candidate edge of the domain if there are multiple accentual candidates in competition. (DEFAULT) (R/L) assigns the accent to the right or left edge of the domain if there are no candidates for accentuation. While absent from previous AFT accounts, Bogomolets (2020, 2023) adds Hyman's (2006, 2009) criteria of culminativity and obligatoriness as the binary parameters CULMINATIVITY (Y/N) and OBLIGATORINESS (Y/N). When positive, CULMINATIVITY is the driving force behind accent selection process from (SELECT). Both CULMINATIVITY and OBLIGATORINESS allow for more detailed descriptions of pitch accent languages which may feature multiple or no accents.

As an example, Table 6 displays the parameters used in Passamoquoddy, and (44, 45) display the accent selection process in the the Passamaquoddy word *tehsahkwapasoltine*. Examples are depicted using traditional AFT grid notation: square

brackets represent word edges, parentheses represent an accentual domain, ‘x’ marks represent accent candidates.

	Primary Accent	Secondary Accent
Accent Domain	BOUNDED (R)	BOUNDED (L)
Accent Placement	WEIGHT-TO-ACCENT (N)	WEIGHT-TO-ACCENT (N)
	DEFAULT (L)	DEFAULT (L)
	OBLIGATORINESS (Y)	OBLIGATORINESS (Y)
	CULMINATIVITY (Y)	CULMINATIVITY (Y)

Table 6: Primary and secondary accent parameters in Passamaquoddy

(44) Primary accent selection in Passamaquoddy

x
[σ σ σ σ σ (σ σ)]
|teh.sah.kwa.pa.sol.ˈti.ne|
‘Let’s walk around on top’

(45) Secondary accent selection in Passamaquoddy

x
[(σ σ) σ σ σ σ σ]
|ˌteh.sah.kwa.pa.sol.ti.ne|
‘Let’s walk around on top’

5.2.2 – Rhythm parameters

The second set of parameters proposed by AFT are the rhythm parameters. They are less numerous than the accent parameters, and differ from MT rhythm computation mainly in that they do no make use of feet to determine rhythm. Consider (46) for the complete list of rhythm parameters.

(46) List of all rhythm parameters in AFT

Rhythm parameters	ANCHOR (PRIMARY/SECONDARY)
	NONFINALITY (Y/N)
	LAPSE (Y/N)

The first parameter is posited in Van der Hulst (2014) and Bogomolets (2020, 2023) as RHYTHM (POLAR/ECHO). RHYTHM determines whether rhythmic beats are assigned to syllables alternating from the primary accent (ECHO) or secondary (POLAR). However, I consider this parameter to be terminologically inexplicit, hene this parameter will be referred

to instead as ANCHOR (PRIMARY/SECONDARY)[9]. This term was is already used as a parameter in Bogomolets (2020: 28, 2023), co-exisiting with RHYTHM. However, 'echo' will still find usage here, not as a parameter, but as a term meaning 'to alternate binarily/ternarily from'. LAPSE (Y/N), has the rhythmic beat alternate every three syllables if positive, or every two syllables if negative. NONFINALITY (Y/N) determines whether the last syllable is extrametrical. If positive, it is extrametrical and can not be rhythmically prominent, if negative, it can be rhythmically prominent.

As an example, Table 7 displays the parameters used in Passamoquoddy, and (47) displays the accent selection process in the the Passamaquoddy word *tehsahkwapasoltine*. For the sake of completeness, all accents are also represented. Subscript numbers are used to differentiate primary from secondary accents. Rhythm's anchor is depicted in bold, and rhythmic prominence underlined on the syllables.

Rhythm parameters	ANCHOR (PRIMARY)
	NONFINALITY (N)
	LAPSE (N)

Table 7: Rhythm parameters in Passamaquoddy

(47) Rhythm selection in Passamaquoddy

x_2 $\mathbf{x_1}$
[σ σ σ σ σ (**σ** σ)$_1$]
[(σ σ)$_2$ σ σ σ σ σ]
|ˌteh.sah.kwa.pa.sol.ˈti.ne|
'Let's walk around on top'

5.3 – An AFT account of Ecuadorian Siona

In this section, ES prosody is analyzed through the lens of AFT. In the first sub-section, accent computation is described using AFT parameters which are the same regardless of the view of pre-aspiration adopted. In the second sub-section, pre-aspiration is anayzed as rhythm, thus, AFT parameters will be used. In the third sub-section, pre-aspiration is analyzed in terms of foot mediality instead.

5.3.1 – Accent computation

Considering that accent parameters do not differ between both of the following analyses of pre-aspiration, this first section is dedicated towards accent parametric values. Indeed, it

[9] The name of the values (POLAR) and (ECHO) are inherited from van der Hulst (2014b), where it was considered that rhythm alternates from the primary accent (echoing it), or an edge prominent syllable (situated at the polar opposite of accent). However, since Bogolomets (2020, 2023) argues that the (ECHO) parameter instead makes rhythm alternate from the primary accent, edge prominence being a case of secondary accent, it becomes unclear what those two values refer to: To an uninformed (ECHO) could as easily mean echoing the primary orr the secondary accent. Likewise both primary and secondary accents are usually found near word edge, obscuring the meaning of (POLAR) as a parametric value.

appears that in order to describe ES accent assignment, only five parameters have to be specified: BOUNDED (R), (DEFAULT) (R), WEIGHT-TO-ACCENT (N), OBLIGATORINESS (Y) and CULMINATIVITY (Y) (see Table 8). This is based on the fact that that pitch accent in ES occurs without exception on the ultimate syllable of the word, without any sensitivity to weight. Examples (48-50) illustrate the accent computation process with words of varying lengths, pre-aspiration omitted from transcriptions.

	Primary Accent
Accent Domain	BOUNDED (R)
Accent Placement	WEIGHT-TO-ACCENT (N) DEFAULT (L) OBLIGATORINESS (Y) CULMINATIVITY (Y)

Table 8: Primary accent parameters in ES

(48) Primary accent computation

x
[(σ σ)]
|o.ˈko|
‘water’

(adapted from Bruil, 2012: 20120912elicr007)

(49) Primary accent computation

x
[σ (σ σ)]
|sa.si.ˈpʔi|
‘He will/can go’

(adapted from Bruil, 2014: 197)

(50) Primary accent computation

x
[σ σ σ (σ σ)]
|a.tʃa.si.kɨ.ˈpʔi|
‘The one (m) listening’

(adapted from Bruil, 2012: 20120918elicr004)

5.3.2 – The rhythmic view within AFT

Following a rhythmic view of pre-aspiration, the first two parameters for rhythm are relatively simple to determine. NONFINALITY in ES is set to (N), as ultimate syllables may be pre-aspirated (48). Likewise, LAPSE is set to (N) due to the binary alternation of pre-aspiration in ES.

Complicating the case of ES rhythm assignment however is the ANCHOR (PRIMARY/SECONDARY) parameter. Indeed, as mentioned in §5.1, rhythm in AFT needs to be anchored either to primary or secondary accent, from which it then alternates. Anchoring rhythm to primary accent yields an accurate analysis for words featuring an even number of syllables, as can be observed in (51, 52). However, similarly to the MT analysis, rhythm does not echo primary accent in words featuring an odd number of syllables, as shown in (52, 53).

(51) Rhythm echoes primary accent in even-syllabled words
|pʔe.te.tʔa.ˈka|
/pʔe.ʰte.tʔa.ʰká/
‘duck soup’

(52) Rhythm echoes primary accent in even-syllabled words
|ŋa.ˈse|
/ŋa.ʰsé/
‘toucan’

(53) Rhythm does not echo primary accent in odd-syllabled words
|a.tʃa.si.kɨ.ˈpʔi|
/a.ʰtʃa.si.ʰkɨ.pʔí/
‘The one (m) listening’
(adapted from Bruil, 2012: 20120918elicr004)

(54) Rhythm does not echo primary accent in odd-syllabled words
|pɨ.pɨ.ˈri|
/pɨ.ʰpɨ.rí/
‘condor’

However, rhythm does not necessarily need to be anchored on primary accent, but can be anchored on a secondary accent instead (ANCHOR (SECONDARY)). As stated in §2.2.2, pre-aspiration is more salient on the second syllable than on, for instance, the fourth syllable. This relative difference in salience could mean that the second syllable hosts not only rhythmic prominence like other even-numbered syllables, but also a secondary accent[10]. In this case, if rhythm is anchored to secondary accent, the distribution of rhythmic pre-aspiration is easily representable as binarily alternating from the peninitial secondary accent realized also as pre-aspiration, but stronger. Pitch accent in this view is completely unrelated to pre-aspiration.

See Table 9 for an overview of the primary/secondary accent and rhythm parameters in ES, and examples (55-58) for an analysis of pitch accent as primary accent, and pre-aspiration as alternating from the peninitial secondary accent, also realized as secondary accent.

[10] It is at this stage unclear whether a hierarchy between primary and secondary accent is needed at all. Pitch accent and pre-aspiration do not appear to share any phonetic content, and there is no proof of greater perceptual salience of one compared to the other. For consistency purposes, pitch accent will still be considered a realization of primary accent, and pre-aspiration (on second syllables only) will be considered a realization of secondary accent, but no claims are made as to whether one is more prominent than the other.

	Primary Accent	Secondary Accent	Rhythm
Accent Domain	BOUNDED (R)	BOUNDED (L)	NONFINALITY (N)
Accent Placement	WEIGHT-TO-ACCENT (N) DEFAULT (L) OBLIGATORINESS (Y) CULMINATIVITY (Y)	WEIGHT-TO-ACCENT (N) DEFAULT (L) OBLIGATORINESS (Y) CULMINATIVITY (Y)	LAPSE (N) ANCHOR (SECONDARY)

Table 9: Primary/secondary accent and rhythm parameters in ES

(55) Pitch accent is a realization of primary accent, pre-aspiration is a realization of rhythm anchored on the secondary accent, two syllable word

x$_{1, 2}$
[(σ σ)$_1$]
[(σ **σ**)$_2$]
|ŋa.ˈse|
/ŋa.^{h}sé/
‘toucan’

(56) Pitch accent is a realization of primary accent, pre-aspiration is a realization of rhythm anchored on the secondary accent, three syllable word

x$_2$ x$_1$
[σ (σ σ)$_1$]
[(σ **σ**)$_2$ σ]
|pɨ.pɨ.ˈri|
/pɨ.^{h}pɨ.rí/
‘condor’

(57) Pitch accent is a realization of primary accent, pre-aspiration is a realization of rhythm anchored on the secondary accent, four syllable word

x$_2$ x$_1$
[σ σ (σ σ)$_1$]
[(σ **σ**)$_2$ σ σ]
|pˀe.te.da.ˈka|
/pˀe.^{h}te.da.^{h}ká/
‘duck soup’

(58) Pitch accent is a realization of primary accent, pre-aspiration is a realization of rhythm anchored on the secondary accent, five syllable word

x_2 x_1
[σ σ σ (σ σ)$_1$]
[(σ **σ**)$_2$ σ σ σ]
|a.tʃa.si.ki.ˈpʔi|
/a.^{h}tʃa.si.^{h}kɨ.pʔí/
'The one (m) listening'

(adapted from Bruil, 2012: 20120918elicr004)

5.3.3 – The foot medial view within AFT

Following a view of pre-aspiration as a foot structure phenomenon instead, there is no need for us to posit the existence of rhythm. Existence of accent despite absence of rhythm in ES would not be unique, as such a thing is attested in world languages like Sibutu Sama (Gordon, 2002), which features primary and secondary accent but no rhythm (59).

(59)	a.	bisˈsala	'talk'
	b.	ˌbissaˈlahan	'persuading'
	c.	ˌbissalaˈhanna	'he is persuading'
	d.	ˌbissalahanˈkami	'we are persuading'

(Gordon 2002: 505)

From this perspective, accent is computed with the parameters expressed listed in Table 8, whereas pre-aspiration is computed foot medially (60, 61). It is important to note that the notion of 'foot' is not rejected by AFT, it is merely regarded as unnecessary to accent/rhythm computation. As such, usage of phonotactic feet does not clash with the theoretical basis of AFT. Foot medial segments in the examples are depicted in bold.

(60) Pre-aspiration and accent in a word with an even number of syllables

(60a) Pre-aspiration computation

(σ^{h}σ)(σ^{h}σ)
|(k^{w}a**ʔ.k**u).(m**a.k**a)|
/(k^{w}aʔ.ku).(ma.^{h}ka)/
'...When it was semi-cooked'

(60b) Accent assignment

x
[σ σ (σ σ)]
|k^{w}aʔ.ku.ma.ˈka|
/k^{w}aʔ.ku.ma.ká/
'...When it was semi-cooked'

(adapted from Bruil, 2014: 165)

(61) Pre-aspiration and accent in a word with an odd number of syllables

(61a) Pre-aspiration computation

(σ^{h}σ)(σ)
|(k**a.k**a).(sio)|
/(ka.^{h}ka).(sio)/
'She will enter'

(61b) Accent assignment

x
[σ (σ σ)]
|ka.ka.ˈsio|
/ka.ka.ˈsíó/
'She will enter'

(adapted from Bruil 2011)

5.3.4 – Conclusion

To summarize, both the rhythmic and foot-medial views in AFT consider pitch accent and pre-aspiration as completely unrelated. Following a rhythmic view of pre-aspiration within AFT, peninitial syllables host a secondary accent phonetically realized as a stronger pre-aspiration than rhythm, which is anchored on the secondary accent.

The foot medial view of pre-aspiration within AFT on the other hand viewed pre-aspiration as a foot-medial phenomenon, whereas pitch accent was computed separately. This analysis' only weakness is that it does not adress or explain why peninitial syllables are more strongly pre-aspirated than other rhythmically prominent syllables. For this reason, the rhythmic analysis is favored over the foot-medial one.

6 – Discussion

6.1 – Evaluation

In the previous two chapters, two main analyses of the relationship between pre-aspiration and pitch accent were identified from an initial total of five analyses. The leading analysis for MT was the foot-medial, one-foot inventory analysis. It posited that pre-aspiration was a foot-medial phenomenon, and that pitch accent on the other hand was a manifestation of accent which was computed via rhythm, which was underlying only, .

The leading analysis for AFT was the rhythmic analysis. It posited that pitch accent was the phonetic correlate of primary accent, and that pre-aspiration was surface rhythm. This rhythm was anchored on a peninitial secondary accent, realized as a more salient pre-aspiration.

The AFT rhythmic analysis possesses two advantages over the MT foot-medial, one-foot inventory analysis. First, it takes more data into account through considering peninitial syllables as secondary accents due to the greater perceptual salience of their pre-aspirations relative to other pre-aspirated syllables. Second, the AFT analysis considered is representationally simpler than the MT one, as all the elements it makes use of (primary accent, secondary accent, rhythm) have phonetic correlates. On the other hand, the leading MT analysis makes use of a rhythm that has no phonetic correlates, and that only serves to compute accent. See Table 10 and 11 for a comparison between the two analyses.

Underlying form	Primary accent	Secondary accent	Rhythm
Surface form	Pitch accent	Pre aspiration (stronger)	Pre aspiration (weaker)

Table 10: Modules required for the AFT rhythmic analysis

Underlying form	Accent	Rhythm	Foot mediality	
Surface form	Pitch accent	∅	Pre aspiration (Stronger)	Pre aspiration (Weaker)

Table 11: Modules required for the MT foot medial, one-foot inventory analysis

Considering the above discussed advantages of the AFT rhythmic analysis over the MT foot medial, one-foot inventory, it appears that a rhythmic analysis of pre-aspiration through the lens of AFT is the most optimal analysis of the (non-existent) relationship between pitch accent and pre-aspiration. It is referred to further in the discussion as the 'leading analysis'.

.

6.2 – Limitations and future research

6.2.1 – Phonetic implementation of accent and rhythm

The analyses conducted in this thesis are all, but especially the AFT rhythmic one, reliant on the quality of collected data. Research on ES itself is very recent compared to other varieties of Siona dialectal group, and thus obtaining good quality recordings of the languages upon which more solid analyses can be based has proven to be a challenge. This is the core issue with the analysis of ES prosody that is made in this thesis: It is reliant on the claim that second syllables are more strongly pre-aspirated than other syllables. So far, this claim has not yet been supported by accoustic data, but by impressionistic data, due to the aforementioned often poor quality of audio recordings.

Furthermore, this claim can only be partially tested, as to the extent of my (limited) knowkedge, there are no words in ES which feature more than two pre-aspirated syllables due to the strict conditions under which it appears, and the limited length of words in ES. As such, only the relative difference between 'pre-aspirated $syllable_1$' and 'pre-aspirated $syllable_2$' can be investigated. In the absence of a 'pre-aspirated $syllable_3$', the possibility that every pre-aspirated syllable is less phonetically salient than the preceding one cannot yet be ruled out. This leads us to multiple questions:

Is the first pre-aspirated syllable in ES more salient by virtue of being the anchor syllable, or by virtue of being the secondarily accented syllable?

Is rhythm cross-linguistically something that emanates from its anchor syllable and decreasing in salience the further away from the accent the rhythmic beat is, or does salience stay stable no matter the distance?

Are accents cross-linguistically necessarily more phonetically salient than rhythmic beats? So far, our definitions for accent and rhythm only include the domains (word vs. sub-lexical) within which prominence of one syllable is compared to that of other syllables, without regard to their phonetic realizations.

For instance, in a given language with peninitial accent and rhythm anchored to it, which phonetic salience pattern associated with accent and rhythm would be expected from (62-65)? In the case of ES, currently available data only allows us to reject the pattern seen in (64).

(62) Stable phonetic salience of rhythm, more salient accent

σ	ˈσ	σ	σ	σ	σ	
	1		0.5		0.5	Relative phonetic salience

(63) Decreasing phonetic salience of rhythm, more salient accent

σ	ˈσ	σ	σ	σ	σ	
	1		0.6		0.4	Relative phonetic salience

(64) Stable phonetic salience of rhythm, equally salient accent

σ	ˌσ	σ	σ	σ	σ	
	1		1		1	Relative phonetic salience

(65) Decreasing phonetic salience of rhythm, equally salient accent

σ	ˌσ	σ	σ	σ	σ	
	1		0.7		0.4	Relative phonetic salience

Finally, a decision was made to continue referring to pre-aspiration accent in ES as a secondary accent, and pitch accent as a primary one, despite the lack of evidence for such a hierarchy. This was done mostly for consistency purposes, both with the rest of the paper and existing literature. To my knowledge there is no work delving into the relationship between different accents in languages, and on whether they necessarily have to be ranked into primary/secondary and more. Further research needs to be conducted on the subject. If hierarchies are not necessary in languages, referring to pitch accent and pre-aspiration accent in ES as $accent_{PI}$ and $accent_{PRE}$ instead of primary and secondary accents would be more justified.

.

6.2.2 – The pitch accent/pre-aspiration relationship in ES

The fact that the most parsimonious (as well as the second-most) analysis points towards a separate treatment of pitch accent and pre-aspiration is also supported from diachronic perspectives. Indeed, pre-aspiration and pitch accent have been argued to have different historical origins in ES. van 't Veer et al. (2025) and Chacon (2014) favor the hypothesis that pre-aspiration arose as a reflex of geminates based on cross-linguistic sources of pre-aspiration, as well as distribution of geminates and pre-aspiration in related western Tukanoan languages. Pitch accent on the other hand is considered to be an inherited feature from Proto-Tukanoan, since Tukanoan languages broadly feature it (Wetzels & Meira, 2011).

Such a conclusion of pitch accent and pre-aspiration being unrelated contrasts with prosodic assumptions made in both ES and related languages (ES: van 't Veer et al., 2025; Chacon, 2010 as cited in Wetzel & Meira, 2011; Sekoya: Johnson & Peeke, 1975; Johnson & Levinsohn, 1990; Colombian Siona: Wheeler, 1970, 1987; Wheeler & Wheeler 1975). In those works, both phenomena were grouped together under an all too vague umbrella term of ‘stress’ without regards to distributional differences between the two. The leading analysis put forward in this thesis underlines the importance of separating prosodic phenomena into different components, which do not necessarily have to be related.

AFT in its current form appears more able than MT to tackle deconstructions of the ‘stress umbrella’ in the case of ES as was demonstrated in this paper, but in other languages as well[11]. In the case of this thesis, it was mostly thanks to Bogomolets’ (2023) contribution to the subject of secondary accent. MT literature on the other hand tends to conflate both secondary accent and rhythm into the notion of ‘secondary stress’, as can be seen in works such as Apoussidou (2007), Hayes (1981, 1995) or Gussenhoven & Jacobs (2017).

[11] See Van der Hulst (2012) and Bogomolets (2023).

7 – Conclusion

In this paper, the relationship between pre-aspiration and pitch accent was analyzed in ES in an attempt to answer the question ‘Which theoretical lens provides the most parsimonious analysis of the prosodic relationship between pre-aspiration and pitch-accent in Ecuadorian Siona?’. Considering pitch accent as a realization of accent, and pre-aspiration as either surface rhythm or phonotactic foot mediality depending on the analysis, their relationship was compared through the frameworks of MT and AFT Theory respectively.

A total of five analyses were drawn depending on the theoretical position taken. After comparison, it became apparent that the rhythmic analysis through the AFT framework was the most parsimonious one of the relationship between pre-aspiration and pitch accent Ecuadorian Siona. Such an analysis entailed that pre-aspiration and pitch accent do not in fact have any relationship, unlike what previous works imply, and that they should be described as separate phenomena instead, a conclusion supported by diachronic evidence.

Importantly, findings from this thesis raise a multitude of questions regarding the cross-linguistic relationship of phonetic implementations of accents and rhythm, questions that can only be answered through more research. Future research must also include better quality recordings of ES in order to investigate the subject of secondary accent more thoroughly, and an expansion of prosodic researchusing an AFT framework, which appears better suited at separating different prosodic phenomena than MT is in its current state.

References

Andrikopoulou, A., Protopapas, A., Arvaniti, A., & Gauthier, I. (2021). Lexical Stress Representation in Spoken Word Recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *47*(6), 830–851. https://doi.org/10.1037/xhp0000929

Apoussidou, D. (2007). *The learnability of metrical phonology*. LOT.

Banti, G. (1988). Two Cushitic Systems: Somali and Oromo Nouns. In van der Hulst, H. and Smith N. (eds), *Autosegmental Studies on Pitch Accent* (Vol. 11). De Gruyter, Inc.

Boersma, P. (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In Benz, A., & Mattausch, J. (Eds). *Bidirectional Optimality Theory*. John Benjamins Publishing Company.

Boersma, P. & Weenink, D. (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.39, retrieved 13 July 2025 from https://praat.org

Bogomolets, Ksenia. (2020). Lexical Accent in Languages with Complex Morphology. 10.13140/RG.2.2.21703.73120.

Bogomolets, K. (2023). Deconstructing Secondary Stress. In J. Weijer (Ed.), *Representing Phonological Detail. Part II Syllable, Stress, and Sign* (pp. 69-98). Berlin, Boston: De Gruyter Mouton. https://doi-org.proxy.uba.uva.nl/10.1515/9783110730081-004

Bruil, M. (2011). Field notes.

Bruil, M. (2012). Documentation of Ecuadorian Siona. http://hdl.handle.net/2196/00-0000-0000-000D-EA53-3 (6 September, 2023).

Bruil, M. (2014). *Clause-typing and evidentiality in Ecuadorian Siona*, vol. 350. Utrecht: LOT Dissertation series.

Cangemi, F., & Baumann, S. (2020). Integrating phonetics and phonology in the study of linguistic prominence. *Journal of Phonetics*, *81*, Article 100993-. https://doi.org/10.1016/j.wocn.2020.100993

Chacon, T. (2014). A Revised Proposal of Proto-Tukanoan Consonants and Tukanoan Family Classification. *International Journal of American Linguistics*, *80*(3), 275–322. https://doi.org/10.1086/676393

Culhane, K. (2022). The phonology and typological position of Waima'a consonants. *Oceanic Linguistics, 61*, 528–559. https://doi.org/10.1353/ol.2022.0005.

Culhane, K. (2023). The prosodic foot beyond prosodic prominence: a preliminary survey. *Linguistic Typology, 27*(2), 313–339. https://doi.org/10.1515/lingty-2022-0039

Edwards, O. (2020). *Metathesis and unmetathesis in Amarasi*. Berlin: Language Science Press.

Fricke, H. (2020). Traces of language contact; The Flores-Lembata languages in eastern Indonesia. *Wacana*, *21*(1), 156–167. https://doi.org/10.17510/wacana.v21i1.878

Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768.

Gordon, M. (2002). A Factorial Typology of Quantity-Insensitive Stress. *Natural Language and Linguistic Theory*, *20*(3), 491–552. https://doi.org/10.1023/A:1015810531699

Gussenhoven, C., & Jacobs, H. (2017). *Understanding phonology* (4th ed). Hodder Arnold.

Hayes, B. (1981). *A metrical theory of stress rules*. Indiana University Linguistics Club.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press, Chicago.

Hudu, F. (2018). Asymmetries in the phonological behaviour of Dagbani place features: Implications for markedness. *Legon Journal of the Humanities*, *29*(2), 197-. https://doi.org/10.4314/ljh.v29i2.8

Hintz, D.M. (2006). Stress in South Conchucos Quechua: A phonetic and phonological study. *International Journal of American Linguistics 72*(4). 477–521.

Hyman, L.M., Byarushengo, E.R. (1984). A model of Haya tonology. In Goldsmith, J., & Clements, G. N. (eds) *Autosegmental Studies in Bantu Tone: Vol. Band 3*. De Gruyter, Inc.

Hyman, L.M. (1981). Tonal accent in Somali. *Studies in African Linguistics 12*, 171-203.

Hyman, L.M. (2006). Word-prosodic typology. *Phonology*, *23*(2), 225–257. https://doi.org/10.1017/S0952675706000893

Hyman, L.M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences (Oxford)*, *31*(2), 213–238. https://doi.org/10.1016/j.langsci.2008.12.007

Goedemans, R., & Van der Hulst, H. (2014). The separation of accent and rhythm: evidence from StressTyp. In van der Hulst, H. (Ed) *Word stress : theoretical and typological issues*. Cambridge University Press.

Johnson, O.E., & Levinsohn, S. H. (1990). *Gramatica Secoya* (Vol. 11). Quito: Instituto Lingüístico de Verano.

Johnson, O.E., & Peeke, C. (1975). *Unidades fonémicas en la palabra Secoya Estudios fonológicos de lenguas vernáculas del Ecuador*. Quito: Ministerio de Educación Pública del Ecuador.

Kubozono, H. (2011). Japanese Pitch Accent. *The Blackwell companion to phonology*, 1-29.

LeSourd, P.S. (1988). *Accent and syllable structure in Passamaquoddy*. Cambridge, MA: Massachusetts Institute of Technology Doctoral dissertation.

Liberman, M., Prince, A., (1977). On stress and linguistic rhythm. *Linguistic inquiry 8*, 249–336.

McCawley, J.D., (1970). *Some Tonal Systems that Come Close to Being Pitch Accent Systems but do not Quite Make it*, Chicago Linguistic Society.

Mol, H., Uhlenbeck, E.M., (1956). The linguistic relevance of intensity in stress. *Lingua 5*, 205–213.

Terken, J., & Hermes, D. (2000). The perception of prosodic prominence. In M. Horne (ed.), *Prosody: Theory and Experiment*. Dordrecht: Kluwer.

Wheeler, A. (1970). *Grammar of the Siona language, Colombia, South America*. Berkeley: University of California press

Wheeler, A. (1987). *Gantëya Bain, El Pueblo Siona del río Putumayo Colombia*, tomo I. Lomalinda: Editorial Townsend.

Wheeler, A., & Wheeler, M. (1975). *Fonémica Siona (Tucano Occidental) Estudios fonológicos sobre lenguas vernáculas del Ecuador*. Quito: Ministerio de Educación Pública del Ecuador

Van der Hulst, H. (1996). Separating primary accent and secondary accent. In Goedemans, R., Van der Hulst, H., & Visch, E. (eds.), *Stress Patterns of the World. Part I*, 1–26. (HIL Publications 2). The Hague: Holland Academic Graphics.

Van der Hulst, H. (2009). Brackets and grid marks or theories of primary accent and rhythm. In Raimy, E., & Cairns C.E. (eds.), *Contemporary Views on Architecture and Representations in Phonological Theory*, 225–245. Cambridge, MA: MIT Press.

Van der Hulst, H. (2010). Word accent: Terms, typologies and theories. In: Van der Hulst H., Goedemans R., and Van Zanten E. (eds.). *A survey of word accentual systems in the language of the world*. Berlin: Mouton de Gruyter

Van der Hulst, H. (2011). Pitch accent systems. *The Blackwell companion to phonology*, 2, 1003-1026.

Van der Hulst, H. (2012). Deconstructing stress. *Lingua 122*. 1494–1521.

Van der Hulst, H. (2014a). The study of word accent and stress: Past, present, and future. In Van der Hulst, H. (ed.), *Word Stress: Theoretical and Typological Issues*, 3–55. Cambridge: Cambridge University Press.

Van der Hulst, H. (2014b). Representing rhythm. In Van der Hulst, H. (ed.), *Word Stress: Theoretical and Typological Issues*, 325–365. Cambridge: Cambridge University Press.

van 't Veer, M., Bruil, M., & Damonte-Matveienko O. (2025). Pre-aspiration in Ecuadorian Siona. In Kuznetsova, N., Anderson, C. & Easterday, S. (eds.). *Structural, typological, evolutionary, and social dimensions*. Berlin: Language Science press.

Wetzels, L. & Meira, S. (2011). A Survey of South American stress systems. In Van der Hulst H., Goedemans R. & Zanten E. (ed.), *A Survey of Word Accentual Patterns in the Languages of the World*. Berlin, New York: De Gruyter Mouton. https://doi-org.proxy.uba.uva.nl/10.1515/9783110198966.1.313

Wilmoth, S. & MansfieldJ. (2021). Inflectional predictability and prosodic morphology in Pitjantjatjara and Yankunytjatjara. *Morphology 31*. 355–381. https://doi.org/10.1007/s11525-021-09380-y.