UNIVERSITY OF CALIFORNIA

Santa Barbara

The cohesive function of prosody in Ékegusií (Kisii) narratives:

A functional-typological approach

A Thesis submitted in partial satisfaction of the requirements for the degree of Master of Arts in Linguistics

by

Daniel William Hieber

Committee in Charge:

Professor Carol Genetti, Chair

Professor Marianne Mithun

Professor Matthew Gordon

March 2016

The thesis of Daniel William Hieber is approved.

Carol Genetti

Marianne Mithun

Matthew Gordon

January 2016

ACKNOWLEDGEMENTS

Thanks are due first and foremost to Carol Genetti for her patience and guidance as this project has evolved, and for fostering my excitement in a new area of linguistics. I am also thankful to Marianne Mithun, Matthew Gordon, and Carlos Nash for various helpful conversations and feedback on this project. My deepest gratitude goes to Kennedy M. Bosire and Gladys Machogu for welcoming me into their homes, for all their help in setting up recording sessions and finding speakers to work with, and for their inexhaustible drive to breathe new life into the Ékegusií language. Imbuya mono.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1144085. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

ABSTRACT

The cohesive function of prosody in Ékegusií (Kisii) narratives:

A functional-typological approach

by

Daniel William Hieber

This thesis aims to advance the idea that prosody is fundamentally about creating *cohesion*, that is, signaling the "relations of meaning that exist within the text" (Halliday & Hasan 1976:4). Building on research on the cohesive function of prosody by Wichmann (2000) and Wennerstrom (2001), I show how each of the features generally referred to as prosodic are used by speakers to lend cohesion to their discourse by signaling the transitions from one unit of discourse to the next, the relations that hold between those units, and their relative prominence. To accomplish this, I look at six prosodic cues in Ékegusií, a Great Lakes Bantu language of southwestern Kenya with lexical and grammatical tone (Cammenga 2002; Nash 2011). Those cues are pause, vowel elision, prosodic accent, pitch reset, isotony (intonational parallelism), and intonational contour. For each feature, I exemplify the ways in which it demarcates conceptually cohesive units of discourse, and/or signals the relations between one unit of speech and another. I show that when these prosodic cues appear, they create cohesive ties between one segment of discourse and another by signaling where one discourse topic ends and another begins, and indicating how – and how closely – the new discourse topic relates to the old (Couper-Kuhlen 2004; Swerts & Geluykens 1994). Together

with morphosyntactic devices for cohesion, such as anaphoric pronouns and reference, the cohesive ties created by prosody are what give coherence to the text, thus distinguishing it from a random assortment of unrelated utterances (Halliday & Hasan 1976). I conclude by discussing how an understanding of prosody as a means for signaling discourse cohesion complements more interactional approaches to prosody (Barth-Weingarten 2013; Barth-Weingarten & Reber 2010; Couper-Kuhlen & Ford 2004), and provides a language-independent means of examining prosody crosslinguistically, thus laying a foundation for future typological studies.

1. Introduction

This thesis aims to advance an understanding of prosody as the collection of phonetic-phonological cues that give structure to discourse. In particular, I wish to show that the function of prosody is fundamentally about creating *cohesion*, that is, signaling the "relations of meaning that exist within the text" (Halliday & Hasan 1976:4). Building on research on the cohesive function of prosody by Wichmann (2000) and Wennerstrom (2001), I show how each of the features generally referred to as prosodic are used by speakers to lend cohesion to their discourse by signaling the transitions from one unit of discourse to the next, the relations that hold between those units, and their relative prominence. To accomplish this, I qualitatively examine six prosodic features across a collection of spoken narratives in Ékegusií (Kisii), a Great Lakes Bantu language of southwestern Kenya with lexical and grammatical tone (Cammenga 2002; Nash 2011). I show that when these prosodic cues appear, they create cohesive ties between one segment of discourse and another by signaling where one discourse topic (or "idea", to use Chafe's (1988:3) term) ends and another begins, and indicating how – and how closely – the new discourse topic relates to the old (Couper-Kuhlen 2004; Swerts & Geluykens 1994). Together with morphosyntactic devices for cohesion, such as anaphoric pronouns and reference, the cohesive ties created by prosody are what give coherence to the text, thus distinguishing it from a random assortment of unrelated utterances (Halliday & Hasan 1976).

Within Bantu linguistics, there have been relatively few studies of intonation and prosody, or indeed of discourse-level phenomena generally (Nurse & Philippson 2003:40). Most phonetic and phonological research in Bantu concerns word-level tone, and to the extent that prosody is discussed, it is usually treated as a syntactic phenomenon (see

Downing (2013) and Cole (2015:5–8) for overviews). Rarely is prosody in Bantu treated as an independent system operating at the level of discourse with its own functions and structure, which in turn interacts with other linguistic systems such as word-level tones and syntax. The present thesis therefore aims to fill this lacuna with a more fully discourse-oriented approach to prosody in a Bantu language.

This thesis proceeds as follows: I begin by surveying the diverse ways that previous researchers have conceptualized prosody with the aim of showing that, despite this diversity of viewpoints, their research exhibits some common themes in how prosody is defined as well as treated in practice. Prosody is viewed by almost all researchers as a phenomenon that operates 'above the word level', 'beyond the sentence level', or 'at the level of discourse' (Cole 2015; Gordon 2016). The intuition behind prosody as a 'discourse-level' phenomenon, I argue, is that prosodic features are discourse-related because they help signal discourse structure. For researchers that emphasize the suprasegmental rather than discourse nature of prosody (e.g., Fox (2000) and many others), I suggest that what is common to the features generally listed as prosodic – tempo and rhythm, intonation, loudness, phrasal stress/accent, and sometimes voice quality – and what is also the motivation for treating them as a unified phenomenon – is precisely that they are the kinds of phonetic-phonological cues that tend to signify cohesive relations and structure in discourse.

The thesis then proceeds to look at six prosodic cues of relevance to Ékegusií, and their realization in the narratives that constitute my data: pause, vowel elision, prosodic accent, pitch reset, isotony (intonational parallelism), and intonational contour. For each feature, I exemplify the ways in which it demarcates conceptually cohesive units of discourse, and/or signals the relations between one unit of speech and another. I conclude by discussing how

an understanding of prosody as a means for signaling discourse cohesion complements more interactional approaches to prosody (Barth-Weingarten 2013; Barth-Weingarten & Reber 2010; Couper-Kuhlen & Ford 2004), and provides a language-independent means of examining prosody crosslinguistically, thus laying a foundation for future typological studies.

2. Themes in Prosody Research

The ways in which prosody researchers have conceptualized their object of study are diverse to say the least. While nearly all researchers regard intonation, tempo, and loudness as central, the perspective from which they approach these features varies. The different emphases on prosody have included its suprasegmental nature (Fox 2000), its hierarchical structure (Cruttendon 1986; Ladd 2008; Nespor & Vogel 2007), its syntactic functions (Downing 2013; Selkirk 1986), its cognitive underpinnings (Chafe 1994), its use in interaction (Couper-Kuhlen & Selting; Barth-Weingarten & Reber 2010; Szczepek Reed 2011), its affective functions (Gussenhoven 2004:71–96), its relation to discourse structure (Swerts 1994), and many others (see Cole (2015) for an excellent review of these and other approaches to prosody, and Fónagy & Bérard (2006) for an example of its manifold functions). These varied foci have led researchers to add other phenomena to the list of features considered prosodic, from low-level phonological structure (morae, onsets, etc.) to voice quality (creaky phonation), so that the list of prosodic features changes from researcher to researcher and language to language. Thus while there is general agreement on at least some of what prosody *involves*, there is hardly any agreement on what prosody *is* (Fox 2000:1, 9).

It is perhaps for this reason that definitions of prosody are often list-like, without always discussing what makes the features in that list cohere. This technique for defining prosody is

used by Wennerstrom (2001:4), for example: “prosody is a general term encompassing intonation, rhythm, tempo, loudness, and pauses, as these interact with syntax, lexical meaning, and segmental phonology in spoken texts.” Like many authors, Wennerstrom goes on to discuss the different functions of prosody, but does not explicitly address what binds them together as an object of study. We are still left with the question of what makes features as disparate as timing and pitch part of a unified phenomenon.

Still, there are some common themes to list-definitions, and much to be gleaned from considering which features are included in them, as well as how prosody is framed in the subsequent discussions. For starters, the very existence of variability across prosody research on different languages keys us into an important point: prosodic features are language-specific. As Himmelman & Ladd (2008:253) note, “there is no way of knowing ahead of time how the phonetic features loosely referred to as “prosodic” – pitch, duration, and so on – are going to be put to phonological use in any given language.” The phonetic cues that signal phonemic distinctions in one language may have purely prosodic functions in another, and vice versa. Creaky phonation is one example of this. In some languages, creaky phonation is a phonemic property of segments operating at the lexical level. In other languages, creaky phonation correlates with phrasal boundaries, and so operates at the level of discourse and prosody. Whether a given phonetic feature is prosodic is a language-specific fact. This point has led several researchers on prosody to set aside terms like *intonation unit* in favor of ones less biased towards one particular prosodic feature, such as the *prosodic unit* of Genetti & Slater (2004).

A second theme in prosody research is a general acknowledgement that prosody is something which functions ‘at the level of discourse’, or at some sort of postlexical level

(Cole 2015:2; Himmelman & Ladd 2008). This is evident from the way that prosody is always closely associated with intonation in the literature. (e.g., Cruttendon 1986:1ff; Jun 2005). Indeed, the one feature that is incontrovertibly referred to as prosodic by all researchers is intonation, and the one universally-agreed upon function of intonation is the demarcation of larger stretches of discourse (e.g., Cole 2015; Fónagy & Bérard 2006; Himmelman & Ladd 2008; Jun 2005; Wennerstrom 2001:7). If prosody does any one thing, it is this. Moreover, even those researchers whose definition of prosody includes low-level phonological features such as syllable structure recognize a principled distinction between these "purely phonological" features and those that interact with other areas of the grammar like discourse (Nespor & Vogel 2007:3).

Related to the discourse function of prosody is its interactional function, another theme emphasized by many of those working in conversation and discourse analysis (Barth-Weingarten & Reber 2010; Ogden 2002; Park 2002; Szczepek Reed & Raymond 2013). Szczepek Reed (2011:13), for example, defines prosody as all the "interactionally relevant, suprasegmental aspects of talk." Much of the research in this perspective draws on earlier work by Gumperz (Gumperz 1982; Gumperz 1992), which introduces the notion of *contextualization cues*, i.e. the cues that enable a listener to interpret the context and meaning of a speaker's discourse. Prosodic cues are taken to be one of the important contributors to contextualization cues. What this and the above theme suggest is that prosody is most centrally about the role that phonetic-phonological features have in shaping discourse.

There is also a long tradition in prosody research, going back to at least Jakobson, Fant & Halle (1951), of associating prosody with suprasegmentals (Cole 2015:2). While later research then refined the understanding of prosody to include only those suprasegmental

features that operate at the phrasal or discourse level (see above), it is still the case that much of the work on prosody considers only suprasegmental features when examining prosody, often explicitly so. At the same time, several prosody researchers have pointed out that, if prosody is intended to encompass those phonological features which operate at the phrasal or discourse level, then a focus on just suprasegmental features risks missing other important phonological cues to discourse segmentation (Fox 2000:2; Himmelman & Ladd 2008:249; Ladd 2008:5). Cole (2015:2), for example, reviewing numerous studies on prosody, points out that, "At the same locations where prosody is expressed through suprasegmental features we also often observe segmental effects, for example, on the acoustic parameters that encode voicing, manner or place of articulation." She then discusses how American English consonants show variation depending on their phrase-level prominence (e.g. whether a syllable is accented), or how segments may become glottalized at phrasal edges (Dilley, Shattuck-Hufnagel & Ostendorf 1996).

Another way in which prosodic phrasing may be signaled by non-suprasegmental cues is through phonological domains (Ladd 1986:312; Nespor & Vogel 2007:5). The domain of application of many segmental processes is a phrase or other discourse-level unit rather than a morpheme or word. The boundaries of these phonological domains contribute to the segmentation of discourse into cohesive units. In many American English dialects, for example, the realization of /t/ as a flap [ɾ] at the beginning of unstressed non-initial syllables applies within prosodic phrases but not across phrase boundaries (Nespor & Vogel 1986). To take another example, in the Chitimacha language (isolate, Louisiana), certain words contain a glottal stop consonant in the ultimate syllable phrase-finally but not phrase-medially (Swadesh 1946:316). As will be seen, Ékegusií too signals prosodic phrasing in part by

means of a segmental phonological process of vowel elision, where elision applies phrase-internally but not at phrasal boundaries. Whatever prosody is, then, it seems it must encompass non-suprasegmental features of speech as well.

To further illustrate the non-intonational means by which languages systematically structure discourse, let us examine a case of particular relevance to Ékegusií. In Bantu languages, it is common to disallow H tones on phrase-final syllables, a phenomenon known as *non-finality* (Nurse & Philippson 2003:64–65). Strategies that languages exhibit for avoiding final H tones include deleting the final tone or shifting the tone leftward, among others. The non-finality of Bantu H tone is thus another example of how the phonological domain for a phonological feature typically functioning at the word level (tone) can serve as one additional phonological cue that helps delimit prosodic phrases at the discourse level. In these cases, tone has both a lexical/grammatical function and a prosodic function. For other languages, however, tone has no prosodic function whatsoever. The lexical tones operate independently of prosody. However, the particular *phonetic* realization of these lexical tones is indeed affected by the intonational contour they happen to reside under, so that tones are shifted upwards in the presence of rising intonation contours or high registers, and lowered in the presence of falling contours or low registers. It is as though the intonational contour is overlaid on the tones, and the phonetic realization of those tones calibrated to the intonational baseline. I term these types of prosodic systems *overlay systems*. Dena'ina (Dene/Athabaskan) exhibits this kind of system (Lovick & Tuttle 2012:141–144), as does Ékegusií.

Finally, let us consider the privileged place of intonation in prosody research. Because of the central nature of intonation to the prosodic systems of the majority of the world's

languages, the strong association of prosody with intonation has largely stood the test of time. But there are languages where intonation seems to do very little prosodic work: according to Lin (2009:140), phrasal boundaries in rGyalrong (Tibeto-Burman) are not determined through overlaying broad intonational contours on the phrase, as we see in most languages. Instead, prosodic phrasing is realized through changes to the lexical tones on the word in phrase-final position. This is a subtle but important distinction: in more canonical prosodic systems, prosodic effects on pitch are realized through changes to broad intonational contours that apply over an entire phrase. But for rGyalrong and other languages like it, prosody affects pitch through changes to the tones on individual segments. These are *tonal* cues to prosodic phrasing rather than *intonational* ones. In another case, Tao (1996) reports for Mandarin that prosodic phrasing is realized primarily through pitch declination across the phrase and changes to the lexical tones, while Kratochvil (1998) states that much of the prosodic work in Beijing Mandarin is accomplished through phonation rather than pitch. Donlay (2015:201) makes the case most explicitly for Khatso (Sino-Tibetan), a language with 8 tones: "Because Khatso is a tonal language and tone contours are part of the lexical tone system, the functional load on pitch is too great to allow it to also play an extensive role in prosody. Thus, boundary tone, pitch reset and tune are not employed in Khatso IUs [intonation units]. Instead, IUs rely on cues that do not involve pitch – that is, pause, lag, rush and intensity." Since prosodic phrasing in Khatso does not depend on pitch, it seems that even intonation cannot be taken as criterial for prosody.

Nonetheless, because of the dominant role that suprasegmentals play in prosody, it became popular, particularly in the generativist tradition, to treat any non-paradigmatic suprasegmental features as prosodic, so that even sub-lexical phonotactic phenomena like

syllable organization are considered prosodic (especially influential here is Nespor & Vogel (1986; 2007)). Out of this tradition grew what is now called the *prosodic hierarchy*. This idea of a hierarchical phonological structure to mirror the type of hierarchical structure seen in syntax is for these researchers a fundamental insight of prosodic research (Cole 2015:2). However, from the perspective of more discourse-oriented approaches to prosody like those outlined above, this hierarchy is not so much a prosodic one as simply phonological, where some levels of the hierarchy function prosodically and others do not. Those levels of the hierarchy which function at the discourse level, or otherwise influence discourse structure, would be considered prosodic, while lower levels like syllables are simply a matter of phonotactics – unless of course certain syllabic or phonotactic patterns contribute to the identification of units of discourse, which can certainly be the case. The essential point, from a discourse perspective, is that it is not participation in the phonological hierarchy itself that makes a feature prosodic; rather, it is the extent to which that feature plays a role in discourse.

A final theme in prosody research is the need to consider its affective (attitudinal, emotional) functions in addition to its more structural ones. It is often noted that a speaker's emotional state affects features considered prosodic, such as loudness and pitch. However, these modulations are considered by some to be *extralinguistic* or *paralinguistic*, in that they convey psychological or sociolinguistic information rather than linguistic information such as the structure of words or discourse (Gussenhoven 2004; Ladd 2008; Laver 1994). On this issue, researchers are divided or uncertain as to whether these affective functions should also be encompassed within the domain of prosody, with some researchers excluding the affective functions of prosody (Ladd 2008; Laver 1994:21–23), and others embracing it (Ford &

Couper-Kuhlen 2004; Inoue 2006:199; Selting 2010). Some maintain the distinction, but acknowledge it as an important function of prosody nevertheless (Gussenhoven 2004; Wennerstrom 2001), while others note that the distinction is impossible to maintain in principle (Fox 2000:271). To complicate matters, Reed (2011) points out that much of what is often called extralinguistic or paralinguistic in prosody is deployed by speakers in interactionally relevant ways, and thus is relevant to the discourse as well. Thus, in addition to their lexical and prosodic functions, features like pitch can also have broader semiotic functions that extend beyond the realm of linguistic semantics narrowly conceived.

This fact complicates the study of prosody insofar as pitch is involved, making the discourse functions of pitch in particular difficult to operationalize. This does not by any means imply that linguists can ignore the affective, semiotic functions of prosody in order to focus on the more narrowly linguistic functions – quite the opposite. The researcher interested in the linguistic functions of prosody alone just needs to operationalize the study of prosodic features in a way that accounts for their affective functions, or at the very least makes it possible to compare prosodic features across different affective states. For example, in many of the stories examined as part of this study, the speaker's baseline pitch rises steadily throughout the narrative, presumably due to increasing levels of excitement as the story nears its climax. For this and other reasons, it is therefore helpful to examine pitch in relative measurements (e.g. semitones or percent pitch change from one vowel to the next) than absolute ones (Hertz). This strategy will be adopted at various points in the present study. So while it is good to be attentive to the emotive functions of prosody, linguistic descriptions of prosody will also need to find a way to abstract away from these affective modulations of pitch.

3. Prosody as Cohesion

Given the diversity of perspectives on prosody just outlined, and the different ways prosody is realized crosslinguistically, it is perhaps no surprise that appropriately general definitions of prosody are difficult to come by. How does one decide when a given feature is functioning prosodically in a language and when it is not? Put differently, on what basis should we consider a feature prosodic? In order to answer this question, a language-general definition of prosody (a *comparative concept* in the sense of Haspelmath (2010)) is required, one that captures the functional underpinnings of prosody rather than attempting to list its various realizations. Such a functional definition would be sufficiently flexible as to accommodate various language-specific implementations. The present section attempts to advance this kind of functional definition.

We have seen above that prosody is crucially linked to discourse, even if this discourse function is not always viewed as its essential characteristic. I want to suggest, however, that the discourse function of prosody is indeed its defining characteristic. What makes us intuitively perceive features like intonation, pause, voice quality, and others as related is that they share in the function of structuring discourse. More precisely, what it means for prosody to be a 'discourse-level' phenomenon is that it helps create textual cohesion by lending additional structure to the text. Therefore I suggest the following definition of prosody:

I. *Prosody is the set of phonetic and phonological cues that create cohesion in discourse by marking transitions between units of speech, the relations that hold between them, and their relative prominence*

It will be the goal of §5 to provide evidence in support of this claim, and show how each of the prosodic features of Ékegusií contributes to the cohesive structure of the text.

Cohesion may be thought of as part of what distinguishes a text from a collection of random utterances, such that, "Where the interpretation of any item in the discourse requires making reference to some other item in the discourse, there is cohesion." (Halliday & Hasan 1976:11). Cohesion is therefore constructed through collections of *ties*, or connections made between two or more items in discourse (Halliday & Hasan 1976:3). Anaphoric reference is a classic example of a kind of tie, since the anaphoric element creates an explicit tie to a prior point in the utterance through coreference with it. In addition, any kind of structure (whether morphosyntactic, prosodic, discourse-level, or other) definitionally contributes to cohesion as well, because structure implies ties between the component parts of that structure: "In general, any unit which is structured hangs together so as to form text. All grammatical units – sentences, clauses, groups, words – are internally 'cohesive' simply because they are structured. The same applies to phonological units, the tone group, foot and syllable. Structure is one means of expressing texture" (Halliday & Hasan 1976:6–7).

As Wichmann (2000:74) points out, "The view that prosody can have a cohesive function in discourse is not new, and has most commonly been discussed in relation to information status, involving the concept of 'given and new information'." Indeed, even Halliday & Hasan (1976:6), in their seminal work on cohesion, note in passing that, "certain types of grammatical cohesion are in their turn expressed through the intonation system." Despite this, few researchers have explicitly examined the connection between prosody and cohesion since. The most thorough treatments are by Wennerstrom (2001:7), who views prosody as constituting a "grammar of cohesion" that "contributes information about connections among constituents in discourse", and Wichmann (2000:74), who devotes an entire chapter to the cohesive functions of prosody in English, with the observation that prosody serves to indicate

"the close conceptual relatedness of successive utterances." Many other researchers, while not explicitly focused on cohesion, nonetheless demonstrate the cohesive function of prosody well. In her review of prosody research, Cole (2015:9) notes that many studies have shown that, "when prosodic cues are available, listeners do appear to make use of them in detecting discourse boundaries and in evaluating the degree of discourse cohesion or juncture between successive utterances in a discourse." Work showing how prosody works to connect large topical units of discourse can also be found in, among others, Brown (1977), Genetti (2011a), Ladd (1986; 1988), Swerts & Geluykens (1994), and Yule (1980). Other studies showing how prosody signals the relatedness between successive units, usually in connection to pitch reset, are Brazil (1985), Couper-Kuhlen (2004), Ladd (1986; 1988), Pierrehumbert & Hirschberg (1990).

Cohesion is thus a common theme among studies of prosody, even when not explicitly acknowledged. The approach advocated here brings the connection between prosody and cohesion to the fore, such that a certain phonetic-phonological feature should be considered prosodic whenever it is functioning to add structure and cohesion to the text. In this respect the present approach is similar to that of Selting (2010:5), who states that "all suprasegmental phenomena that are constituted by the interplay of pitch, loudness, duration and voice quality can be understood as prosodic, as long as they are used – independently of the language's segmental structure – as communicative signals." The definition given here represents a synthesis of more cognitively-oriented, "intonation-as-information-flow" (Couper-Kuhlen 2015:84–85) approaches to prosody like that of Chafe (1979; 1988; 1993; 1994), Croft (1995), Du Bois (1992), and Mithun (1996), and more interactionally-oriented, "intonation-as-contextualization-cues" (Couper-Kuhlen 2015:85) approaches like that of Couper-Kuhlen

& Ford (2004), Szczepek Reed (2011), and Szczepek Reed & Raymond (2013). This latter research builds on Gumperz' (Gumperz 1982; Gumperz 1992) notion of *contextualization cues*, the 'empirically detectable signs' that allow speakers to make "inferences about what is being said in interaction or more generally about what is 'going on'," with prosody being foremost among these cues (Couper-Kuhlen & Selting 1996:13). Couper-Kuhlen (2015:85) explains that, "In this approach contextualization cues, and consequently prosodic phenomena, are not seen as accidental or fortuitous nor as automatic reflexes of cognitive and affective states. They are thought to have their own systematicity." And like the approach advocated here, intonation is taken as just one among a number of potential contextualization cues. It is for this reason that, "there has been a subtle shift away from the study of "intonation" to the study of *prosody* and discourse. [This] school of thought thus actually deserves to be called *prosody-as-contextualization cue*." (Couper-Kuhlen 2015:85).

What context is being constructed by speakers through their use of prosody? As speakers talk, they are constantly foregrounding new or semi-active information, while simultaneously backgrounding old information (Chafe 1988:22). This fact is also noted by Wennerstrom (2001:70), who states, "Turn by turn, participants interpret each new utterance in the context of the mental representation that they have constructed so far, taking into account the prior text" and much other social and contextual information. The backgrounded information becomes the prior context, the backdrop against which the new information is situated. Therefore, the speaker must frequently signal when they are transitioning from one discourse topic to the next, and in what way the new topic should be considered in relation to the old. Prosodic cues play a crucial role in this continual task that speakers have of foregrounding

new topics in the discourse while simultaneously backgrounding old ones (Chafe 1994; Kumpf 1987).

One of the ways prosody accomplishes this task of backgrounding and linking segments of speech is, counterintuitively, through the creation of prosodic breaks. A great deal of prosody research has focused on the way that prosodic features combine to indicate the location and relative strength of such boundaries or breaks (Barth-Weingarten 2013; Brierley & Atwell 2005; Downing 2010; Genetti & Hieber 2015; Karlsson, Svantesson & House 2014; Myers 1996; Oliveira 2003; Swerts 1997; Yabin & Aijun 2003). It is important to recognize, however, that prosodic breaks or disjunctures are every bit as cohesion-building as prosodic cues that mark continuity (e.g. continuing terminal contours) or parallelism (e.g. isotony). A prosodic break is what indicates for the listener that the preceding discourse topic is done, and can now be treated as a bounded unit, against which the upcoming topic will be situated. This function is pointed out by Chafe (1988:39) for intonation units: "Another common function of an intonation unit is to provide what can be called an "orientation" for a preceding or following clause."

Additionally, while the strength of a break in the discourse signals the degree of discontinuity between one segment of speech and another, the flip side of this fact is that it also signals their degree of cohesiveness. Weaker breaks indicate a stronger cohesive tie between units of speech than stronger breaks. Finally, it is precisely these breaks that help create hierarchical structure in the discourse (Wennerstrom 2001:96). As we will see in §5, stronger breaks signal the ends of larger units of discourse (*prosodic sentences*), while weaker but still strong breaks signal the ends of smaller units (*prosodic phrases*). Because all structure is cohesion-building, these breaks therefore contribute to the overall cohesion of the

text, rather than just its fragmentation. ‘Boundaries’, then, work to add coherence to the discourse every bit as much as more overt links between utterances, such as isotony or anaphora.

In sum, the approach to prosody outlined here views prosody as the phonetic and phonological cues that structure discourse through the creation of cohesive ties between units of speech. It will be the aim of the remainder of this thesis to provide evidence from Ékegusií in support of this perspective. In §4 I describe the data and methods used for this study, and then I proceed in §5 to describe the cohesive functions of a variety of prosodic cues in the data.

4. Data & Methods

The data for this study consist of a collection of 10 texts which I recorded in the Kisii region of Kenya during a fieldwork trip in the summer of 2014. These texts are part of a larger collection of stories, conversations, expositions, and speeches at community gatherings, in both audio and video formats, recorded as part of a community language revitalization project. As such, the ten texts selected for inclusion in this study were chosen because they are taken by community members to be exemplary of the traditional folktale genre. All ten were told by a single speaker, who community members acknowledge as one of the best storytellers in the area. This ensured a fair amount of consistency across the texts in terms of structure and content. They were narrated entirely in Ékegusií, and transcribed and translated using ELAN (Max Planck Institute for Psycholinguistics) with the help of a second native speaker. The folktales revolve around the actions of a recurring set of anthropomorphized animals, and thus contain a fair amount of reported speech and character

voicing. Many of the stories also contain short, one-stanza songs, but these sections were not included in the present analysis.

Once the stories were transcribed, they were annotated for a number of the prosodic features including pause, breath, prosodic lengthening, creaky phonation, vowel reduction, isotony (intonational parallelism), and overall contour shape. Unlike nearly all previous researchers on prosody, however, I do not demarcate 'intonation units', 'prosodic phrases', 'intonational phrases' and the like, because I wish to remain agnostic regarding the status of the 'boundaries' that separate them. Most prosody researchers working with natural discourse (as opposed to constructed sentences) have noted the convergent nature of prosodic cues (Barth-Weingarten 2013; Chafe 1994:60; Cole 2015; Du Bois 2014; Himmelman & Ladd 2008:252; Tao 1996:41; Wennerstrom 2001:28), and many have commented on the fact that prosodic boundaries come in different degrees of strength (Amir, Silber-Varod & Izre'el 2004; Barth-Weingarten 2013; Genetti 2011; Himmelman & Ladd 2008:252; Ladd 1988; Lovick & Tuttle 2012:313–314; Swerts 1997; Yufang & Bei 2002:4). All writers on prosodic phrasing acknowledge that no one cue is sufficient to establish prosodic constituency.

In spite of this fact, the practical necessity of discourse transcription requires that the researcher make a decision whether or not to mark a boundary at any given point in the discourse, reducing a complex, gradient phenomenon to a mere binary one. The identification of prosodic units and their boundaries is also a point of some contention among prosody researchers because of its potential for methodological circularity: the same phonetic-phonological cues cannot be used as a diagnostic for identifying prosodic phrases if prosodic phrases are then used to establish a correlation between those phonetic-phonological cues and prosodic structure (Barth-Weingarten 2013; Swerts 1997). As such, different researchers

have attempted to establish prosodic constituency 'from the ground up' or in a 'bottom-up' fashion, i.e. in an independently valid way (Barth-Weingarten 2013; Swerts 1997; Yufang & Bei 2002). One way this has been addressed is through crowdsourcing the identification of prosodic phrases to both experts and non-experts: Oliveira (2003) involved seven experts in intonation and discourse analysis in the annotation task, while Swerts (1997) has 38 non-linguists segment texts into 'paragraphs' (see also Lin (2009)). Swerts then looked at the extent of agreement among different transcribers as a heuristic for the strength of each given boundary. However, this approach seems to merely crowdsource the circularity, so to speak, and does not address the basic issue.

A different approach is taken by Barth-Weingarten (2013): rather than focusing on prosodic units or even boundaries per se, she instead examines the 'bumps' or 'cuts' in the flow of speech, which she terms *cesuras*. She establishes the independent validity of cesuras by identifying them not through their phonetic-phonological properties, but through their interactional relevance to the discourse. If prosodic segmentation has any function in the discourse, then speakers and listeners alike must be aware of cesuras at some level, and deploy them in interactionally relevant ways. Barth-Weingarten therefore determines the placement of cesuras by where conversational participants (attempt to) come into the discourse, on the idea that participants take up a turn at perceived breaks. The speech prior to these *incomings*, as they are termed, is then examined to determine the phonetic-phonological cues that the listener may have been picking up on, and in particular what cues tend to cluster at these incomings.

Barth-Weingarten's study neatly avoids circularity while simultaneously highlighting the discourse-based nature of prosody. A similar 'bottom-up' approach will be adopted here,

where the assorted features to be investigated will be annotated as-is, wherever they happen to occur in the data, mid-utterance or not. Only then will convergence of the different features be examined. It should be noted, however, that even this 'bottom-up' approach is not entirely objective, since the researcher (me) must still decide which phonetic-phonological properties to investigate for correlations to the prosodic boundary. As discussed above, there is no way to know in advance whether a given phonetic-phonological feature will be prosodic in any given language. Contra more positivistic approaches to science, there is no way this decision can be made from the 'bottom-up' on the basis of the data alone. Ultimately the linguist's subjective judgment of the relevant features, built through a process of induction over many instances of the phenomenon, has to be taken into account. The definition of prosody proposed here at least provides a functional basis on which this decision can be made: does the particular phonetic-phonological feature contribute to the structure and the cohesion of the discourse? Therefore some of the phonetic features selected for examination in this study were included on the basis of this heuristic and my impressionistic sense of the data. Other features were included because they are known to be relevant to the prosody of related or neighboring languages, or common typologically. Most are motivated in both ways. Not every feature was found to be prosodically relevant, as we will see.

Another reason I avoid transcribing prosodic boundaries is that, technically, the different prosodic features that might constitute a 'boundary' do not actually line up temporally. We tend to conceptualize prosodic boundaries as happening at a single point in time, but in reality the variety of features which are typically taken to constitute those boundaries occur over a span of several seconds or more. In my Ékegusií data, for example, creaky phonation often precedes or follows a pause, while a slowing in the rate of speech signals that a pause

or some other sort of break or transition might be coming up. How can one decide where the 'boundary' is, if the features that constitute it in fact span several seconds? This problem forces us to reconsider the notion of 'boundaries' and 'cesuras', in favor of more online, interactional models where speakers are constantly signaling (either intentionally through primary cues like intonation, or unintentionally through secondary cues like creaky phonation) what they're doing at any given point in the discourse. Therefore I wish to remain agnostic about prosodic boundaries, focusing more on the cohesive functions of different prosodic features. More useful for present purposes is to recognize the existence of *transition zones*, where speakers employ multiple signals to indicate a transition from one section of the discourse to the next.

Because the individual prosodic features each involve different decision-making processes for when and how to annotate them, the discussion of those methods is reserved for the particular subsection in which that prosodic feature is discussed. I now turn to the examination of each of these prosodic features.

5. Prosodic Cues to Cohesion

This section examines six prosodic cues in Ékegusií, and shows whether and how those features create cohesive ties between sections of the discourse. Each feature was selected because it either stood out to me in my transcriptions as a potential correlate to discourse structure, or because it is a known correlate to discourse structure in other languages, and especially Bantu languages. The cues to be examined are pause, vowel elision, prosodic accent, pitch reset, isotony, and type of intonational contour. Other potential cues that could be investigated, but that will not be examined in this study, are pitch range, intensity/amplitude, vowel reduction, rate of speech, phrase-final lengthening, creaky

phonation, and many more. In general, I have left aside more 'secondary' prosodic features, i.e. features that, for Ékegusií at least, seem to be purely a matter of phonetics and the physiology of vocal production rather than something that is actively manipulated by speakers for discourse purposes. Creaky phonation, for example, appears predictably at the ends of utterances ending at the bottom of the speaker's pitch range, and in repairs and restarts. It does not appear to be a primary cue for prosodic phrasing. In other languages, however, including some varieties of English, creaky phonation deserves to be considered a primary prosodic cue because it is a salient sociolinguistic variable, or because it is manipulated in discourse-relevant ways.

5.1. Pause

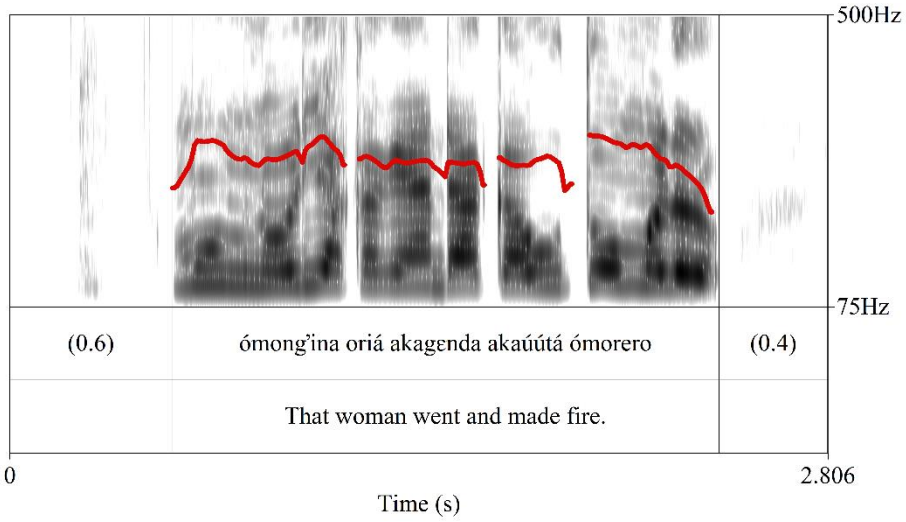
Pauses are generally agreed to be one of the most important cues to prosodic phrasing, with many studies reporting strong correlations between pauses and prosodic boundaries (Amir, Silber-Varod & Izre'el 2004; Downing, Mtenje & Pompino-Marschall 2004; Genetti & Slater 2004; Lin 2009; Lovick & Tuttle 2012; Scheutze-Coburn, Shapley & Weber 1991). Moreover, the placement and relative length of pauses is manipulated by speakers to regulate information flow (Swerts & Geluykens 1994). The length of the pause is also found to correlate with the strength of the break in the narrative (Swerts 1997; Vaissière 1983). Many other prosodic cues are found in the vicinity of pauses as well (Oliveira 2000; Swerts 1994; Swerts 1997; Yabin & Aijun 2003; Yufang & Bei 2002). Pauses are therefore central to understanding how it is that prosody lends coherence to the discourse.

For this study, pauses were measured from the end of the visible activity in the waveform corresponding to a stretch of speech to the beginning of the next, and rounded to the nearest

tenth of a second.[1] Breaths were not counted as speech, and therefore many pauses contain a quick breath within them (more on the prosodic role of breath below). Very occasionally, a micropause (0.2 seconds or less) will occur within an utterance; these are usually hesitations or repairs.

How do pauses lend cohesion to the text? One simple way is by indicating that the segment of speech demarcated by pauses should be treated as a cohesive unit. Consider the following example:

Figure 1. Pauses demarcating a cohesive unit of speech

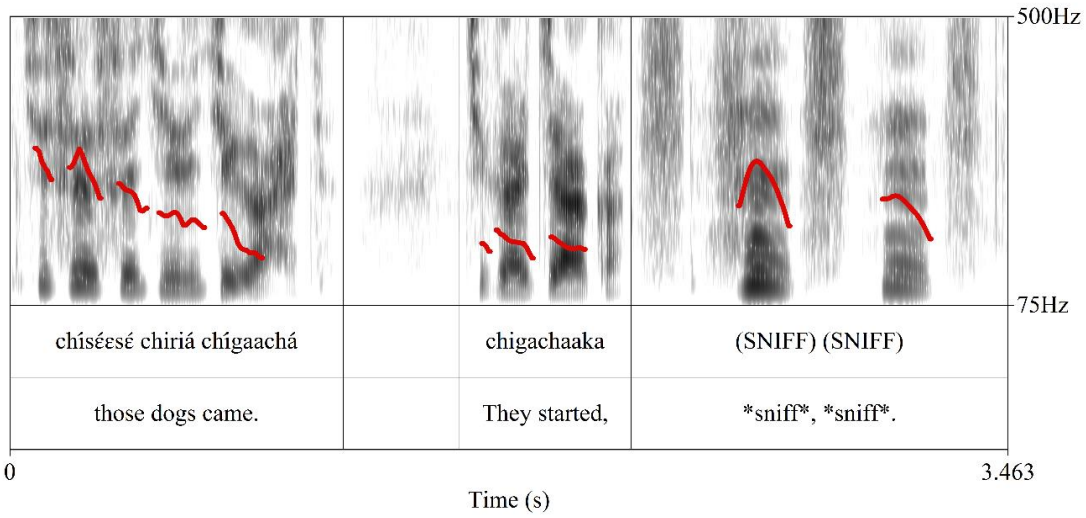


It is very common for an utterance in these narratives to contain a sequence of two or more verbs in the Consecutive Past tense, marked with *-ka-/-ga-*, indicating that the action follows the preceding one. This can be seen in Figure 1 in the two verbs *akagɛnda akaúútá* 'she went and made (fire)'. Given that the translation is 'she went and made (fire)' rather than 'She left. Then she made fire.', it seems the utterance is being construed as a contiguous action.

[1] A greater degree of precision in rounding is of course possible, but for the phenomena of interest here 100 milliseconds was sufficiently precise for capturing any relevant trends.

However, when two verbs in the Consecutive Past are separated by a pause, they are usually translated as separate actions rather than a single contiguous one. Compare the example in Figure 2 with that of Figure 1 above.

Figure 2. Pauses indicating that two events are conceptualized as separate



The fact that this sequence is translated with a sentence boundary rather than with a serial verb construction like 'those dogs came and started *sniff*' suggests that the two verbs are being construed as distinct actions. In both examples, there is no noticeable pitch reset between the verbs, so the presence or absence of pause is the sole indicator of whether the two consecutive verbs should be construed as one cohesive event or two.

Interestingly, in these stories the speaker seems to draw on the cohesive function of pause for dramatic effect as well. Consider the full text of the first story in the dataset below, where each line indicates either a unit demarcated by pause or a pause itself. The length of each pause is given in parentheses. In addition, the length of each utterance is provided in the far right column. Sections of the story are labeled with the canonical parts of the narrative schema (Labov & Waletsky 1967).

(1) Ómoíséké Ɔ́sɔɔkɛ́rɛ́tɛ́ Chísɛ́ɛsɛ́ 'A Girl Who Got Married to Dogs'

#	Ékegusií	English	Narrative Function	sec
1	Mogano ngóóchá índe.	May I, Story, come?	Introduction of narrative	1.5
2	(0.1)			
3	Mogano ínchúó.	Story, come.	Introduction of narrative	0.9
4	(0.1)			
5	Ómoíséké	A girl	Introduction of participants	0.7
6	(0.05)			
7	n'áarɛ́ngɛ́	was there	Introduction of participants	0.8
8	(0.3)			
9	ɔ́sɔɔkɛ́rɛ́tɛ́ chísɛ́ɛsɛ́.	who was married to dogs.	Introduction of participants	1.3
10	(0.6)			
11	Ómoíséké óyo,	This girl,	Introduction of participants	1.1
12	(0.1)			
13	ɛ́kɛrɔ́ asɔɔkɛ́rá chísɛ́ɛsɛ́,	when she got married to dogs,	Introduction of participants	1.5
14	(0.4)			
15	táatágete ábaíbórí báyɛ bagɛnda ɔ́rɔrɔ.	she didn't want her parents to go there.	Introduction of participants	2.0
16	(0.6)			
17	Rítukó ɛ́rimɔ́,	One day,	Complicating action	0.9
18	(0.1)			
19	mama ɔ́mwabɔ́ akamobétéreria kabisa, "Níntágeté góóchá.	her mother persisted forcefully, "I want to come.	Complicating action	2.9
20	(0.3)			
21	Ínché kɔ́rɔra áasé gwásɔɔgɛ́tɛ́ góchia.	I come and see where you got married."	Complicating action	1.7
22	(0.6)			
23	Mama ɔ́mwabɔ́ akaímóka ésabaria	Her mother set on the journey	Complicating action	1.6
24	(0.05)			
25	akagɛnda.	and went.	Complicating action	0.5
26	(0.6)			
27	Ɛ́kɛrɔ́ agɛnda aíká ɔ́rɔrɔ,	When she went and reached there,	Complicating action	1.4
28	(0.5)			

29	akamɔtɛ́ɛ́bíá, “Égento ɔ́gɔkɔ́rá ígó ókwebisa írɔngɔ.”	she [the daughter] told her, “What you do is hide yourself in the ceiling.	Complicating action	2.5
30	(0.6)			
31	N'arió chísɛ́ɛsɛ́ echió chítáache gókoria.	Then those dogs will not be able to eat you up.”	Complicating action	1.6
32	(0.5)			
33	Mama ɔ́mwabɔ́ akaímóká akebisa írɔngɔ.	The mother went and hid in the ceiling.	Complicating action	2.2
34	(0.4)			
35	Ɛ́kɛrɔ́ ébisa írɔngɔ,	When she hid herself in the ceiling,	Movement toward climax	1.7
36	(0.3)			
37	chísɛ́ɛsɛ́ chiriá chígaachá.	those dogs came.	Movement toward climax	1.2
38	(0.4)			
39	Chigachaaka *sniff* *sniff*.	They started, “*sniff* *sniff*.	Movement toward climax	1.8
40	(0.1)			
41	Égento giatíɔ́kire igá igá.	Something has smelled here.	Movement toward climax	1.5
42	(0.1)			
43	*sniff*	*sniff*	Movement toward climax	0.7
44	(0.1)			
45	Égento giatíɔ́kire igá igá.” “Yáayá!	Something has smelled here.” “No!	Movement toward climax	1.4
46	(0.1)			
47	Oywó n'ómogeni.”	That’s a visitor.”	Movement toward climax	0.8
48	(0.5)			
49	Ómong’ina oriá ɛ́kɛrɔ́ achikoígwá igɔ́,	When that woman heard so,	Climax	1.6
50	(0.1)			
51	agaíká insé.	she came down.	Climax	0.8
52	(0.3)			
53	Chísɛ́ɛsɛ́ chiriá chikabwátá chikamɔtɛ́ɛ́bíá úútá ómorero.	Those dogs told her to make fire.	Climax	2.4
54	(0.5)			
55	Ómong’ina oriá akagɛnda akaúútá ómorero.	That woman went and made fire.	Climax	1.9
56	(0.4)			

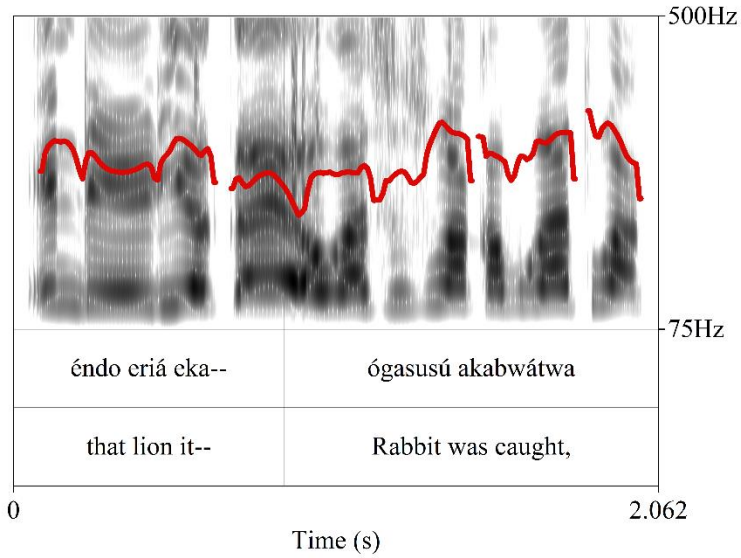
57	Ririá atúgáíma góchia kóúútá ómerero,	When she bent to make fire,	Climax	2.1
58	(0.4)			
59	chísɛɛsɛ chiriá chigaacha chikamwááka igɔ́	the dogs came from behind and hit her hard	Climax	2.3
60	(0.3)			
61	chikamoruta rííkɔ́.	and threw her into the fire.	Climax	1.2
62	(0.3)			
63	Chikamorúúsíá óó chikamotúgútá gochia ariá,	They removed her from there and threw her away,	Climax	2.1
64	(0.3)			
65	akabwáta agakwá.	and then she died.	Climax	1.1
66	(0.1)			
67	N'abo ómogano ɔ́nɛ́ oererete ígó.	That's how my story was.	Coda	1.9

There are several points to notice regarding how this speaker uses pause to structure this narrative. First, with the exception of utterances containing reported speech, the longest utterances are those in the story's climax. Whereas the utterances earlier in the story contain just a single main verb, utterances in the climax contain sequential chains (like those discussed above using *-ka-/-ga-*). By increasing the length of time between her pauses, the speaker treats broader sequences of action as a single unit, providing the perceptual effect of making it seem like more is happening at once. At the same time, the length of the pauses decreases, suggesting that the events are more closely cohesively connected and therefore happening in quick succession.

It is also interesting to note where the speaker *does not* use pause, particularly at the boundary leading into reported speech. These utterances are being structured as a single cohesive event of saying (lines 19, 29, and 39). The lack of pause at these locations parallels the tighter degree of conceptual unity of the speech frame and the reported speech (Genetti 2011b; Swerts & Geluykens 1994).

It is also rare that this speaker pauses during repairs. One example of a repair without pause is shown below.

Figure 3. Repair without pause



Repairs, restarts, and hesitations are not often new pieces of information but instead restate, reframe, or continue the idea of the first part of the utterance. One cohesive strategy that speakers can use in repaired utterances is to minimize prosodic breaks by using fillers and avoiding pauses until the entire cohesive idea they are trying to present is finished. This appears to be the predominant strategy for Ékegusií narrators. They never perform repairs via metacommentary (e.g. 'Sorry, I meant...') in any of the texts I have recorded.

It might be questioned whether speakers really are using pauses as a discourse device, or whether they simply use pause to enable breathing. To this question it can be pointed out that there are many fewer breaths than pauses in these narratives. Only approximately half of all pauses in the stories contain a breath. Thus these pause-delimited utterances cannot appropriately be thought of as 'breath groups'. In addition, since the length of the breath

itself depends in part on the length of the preceding and following utterances (Winkworth, Davis, Ellis, et al. 1994; Winkworth, Davis, Adams, et al. 1994), it would seem that breath – or at least its duration – is more a result of production factors than discourse motivations. Moreover, towards the climax of her narratives, this speaker goes longer and longer stretches without breath (partially as a result of the longer utterances mentioned above), often sacrificing her own personal comfort to maintain the flow of the narrative. In the above story, the speaker goes as long as 5.2 seconds without breathing at the point where daughter tells the mother to hide in the ceiling. It seems she did this so as not to interrupt the flow of the reported speech in those utterances. One gets the impression listening to these narratives that the narrator works her breathing in around the pacing of the narrative, rather than adjust the narrative to her breathing, so that she is often short of breath and needing to make rapid, deep inhalations.

It does not seem, then, as though pause functions to abet breathing, nor does breath appear to serve a particularly salient discourse function. Breath is at best a secondary cue to prosodic phrasing. Rather, for this speaker at least, breath seems opportunistic, something she slips into pauses before continuing. Pause, on the other hand, seems systematically motivated by the need to lend coherence to discourse by segmenting it into meaningfully cohesive segments, and using the length and presence/absence of pause to signal the degree of integration between those units. If pause were a matter of just taking a breath, we would expect to see the narrator pausing in the middle of words or clauses, especially in a several-minute long story like this. And if pause were a matter of demarcating syntactic units, we'd see more consistency in the syntactic units delimited by pauses. But as seen in the story presented above, any variety of syntactic units or multiple syntactic units can occur between

pauses. Instead, it seems speakers actively utilize pauses to separate their speech into bounded narrative events or ideas, thereby giving cohesion to their discourse.

5.2. Vowel Elision

As mentioned in §2, not all cues to prosodic constituency are suprasegmental. Ékegusií has one cue to prosodic phrasing that is segmental in nature: vowel elision within prosodic units. Ékegusií vowel elision applies to the domain of, and constitutes additional evidence for, cohesive segments of speech.

Ékegusií, like many Bantu languages (Casali 1997), elides vowels at word junctions (syntactic word junctions in particular, though for other Bantu languages elision applies to phonological word boundaries instead). When one word ends in a vowel and the next begins with one, Ékegusií resolves the potential vowel hiatus by deleting one or more vowels and resyllabifying (Bosire & Machogu 2013:xxi; Cammenga 2002:158ff). The rules regarding which vowel deletes, when compensatory lengthening occurs, and how tonal processes operate during elision are somewhat involved, and a full treatment of them cannot be given here. For present purposes it is enough to remember that the most common case is where the word-final vowel deletes, while the first vowel of the following word is resyllabified with the final syllable of the preceding word. Other processes of tonal spreading or leftward tonal shift, lengthening, and in certain contexts diphthongization may also apply.[2] It should also be noted that these processes do not always result in a CV syllable structure. Even after elision, two vowels may still be in hiatus (in cases where there were three vowels in sequence to begin with). A few illustrative examples of vowel elision from my data are provided in (1).

[2] Diphthongization is included as a type of vowel elision here because it involves resyllabification of two vowels into one, which then serves as the nucleus of the final syllable of the preceding word. The following word therefore loses its initial vowel.

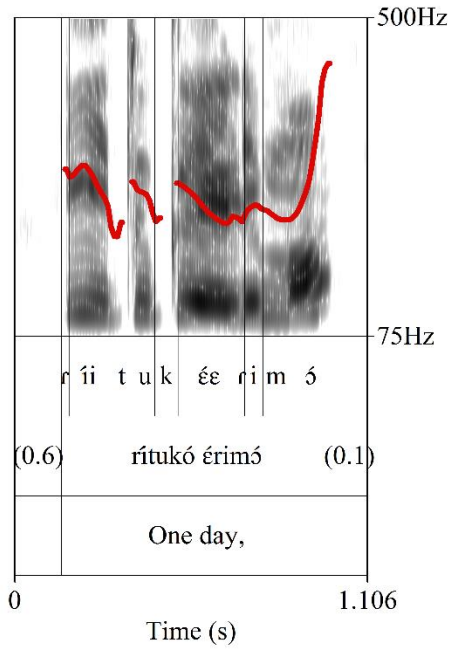
Elided words are separated by a word boundary symbol <#>, and non-elided words by a space.

(1) a. *éndo égatéébá rɛɛrɔ ígo íngóóchá kógɛnda íntégé*
lion it.said today so I.am.going to.go so.that.I.trap
[éend#éɣatéébá ɾɛɛɾɔ#ɣó#óŋgóótʃá kɔɣɛɛndá#ántéɣé]
'The lion said, "Today I'll go and lay a trap."'

b. *índóré ómonto ógoochá ória chínchúgú chíáné*
so.that.I.see person who.goes to.eat ground.nuts my
[índóɾ#ómoont#óɣootʃ#óɾí#tʃiintʃúɣú tʃíáné]
'[…] so that I see who comes to eat my ground nuts.'

The most common domain that vowel elision picks out is an utterance delimited by pauses. Figure 4 demonstrates one such case where vowel elision applies within an utterance but not at its edges.

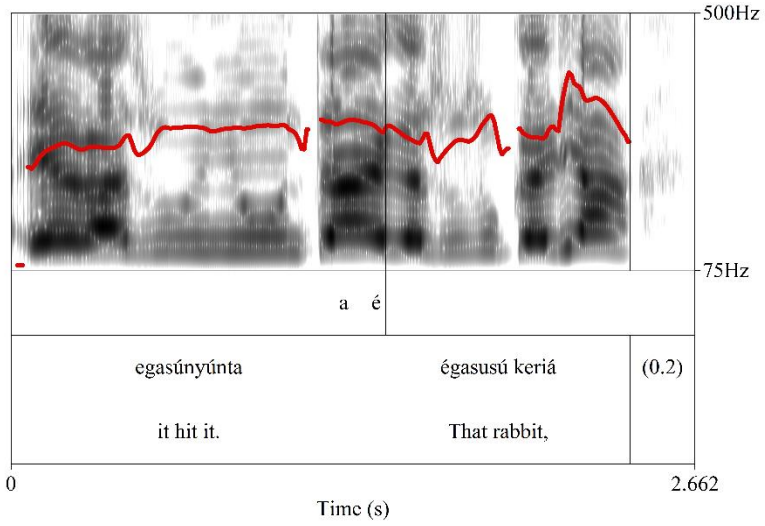
Figure 4. Canonical vowel elision with an utterance



In this example, the final /ó/ of *rítuko* is elided because it is utterance-internal, but the final /ɔ́/ of *érimɔ́* is not because it is utterance final and borders a pause.

However, there are cases where vowel elision fails to apply within an utterance, creating a prosodic break that does not align with any pause. One instance is shown in Figure 5.

Figure 5. Lack of vowel elision indicating a prosodic break



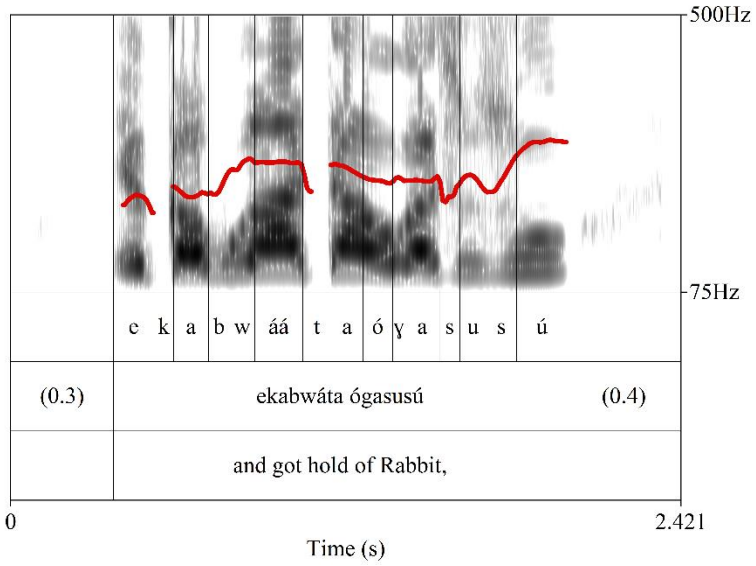
The word-final /a/ of *egasúnyúnta* and the word-initial /é/ of *ígasusú* appear in hiatus, with neither one elided, providing evidence for a prosodic break. There are other cues indicative of prosodic discontinuity at this point: the first word is pronounced on a level register throughout with no declination, and lengthened for dramatic effect. Both of these effects stop at the word *ígasusú*. Notice too that the final vowels of *keríá* are also present and not elided, since they are adjacent to a salient prosodic boundary with a pause.

What is the discourse function of these non-pausal phonological boundaries indicated by lack of vowel elision? The full context of this example, in translation, is 'It [the lion] hit it [the rabbit], it hit it. Then that rabbit? That rabbit died.' In the above case, there is a shift in discourse topic from the lion (the one doing the hitting) to the rabbit (the one being hit).

Let us examine a few instances in more detail. In one case (Figure 6), there is no vowel elision between words because the speaker is drawing the words out, speaking them slowly

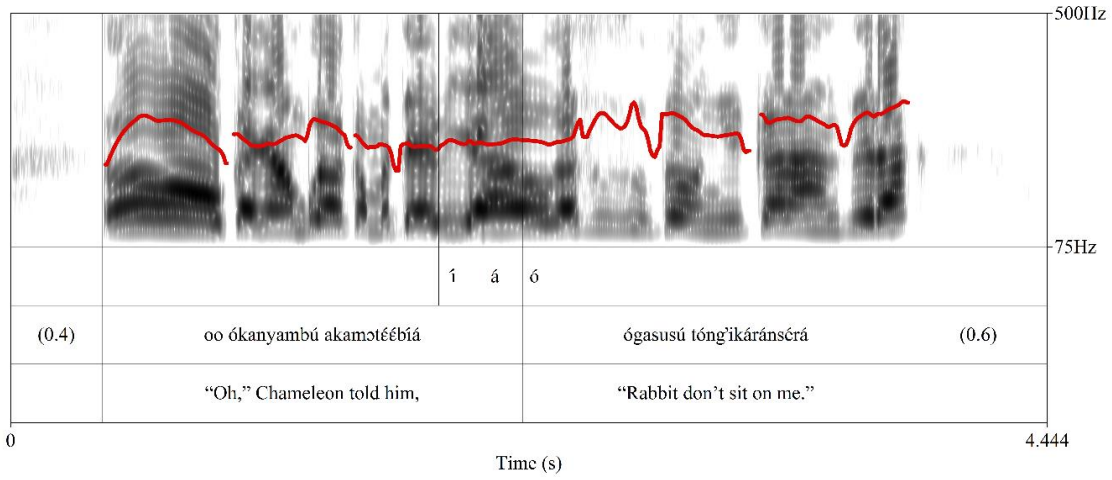
in imitation of the ominous nature of Lion's approach and Rabbit's impending death. The narrator does not pause between the words however, and so the utterance retains a degree of cohesiveness. Lack of elision slows the pacing of the narrative by creating a minor prosodic break between words.

Figure 6. Lack of vowel elision motivated by narrative pacing



Vowel elision also often fails to apply at the transition into reported speech, as in Figure 7. The reported speech at both the beginning and end of this utterance is also marked by a higher register than the speech frame, *ókanyambú akamɔtéébíá*. The lack of vowel elision here indicates a minor prosodic break signaling the transition into reported speech, while the lack of pause suggests that this event is being construed as a single cohesive event of saying.

Figure 7. Lack of vowel elision at the transition into reported speech



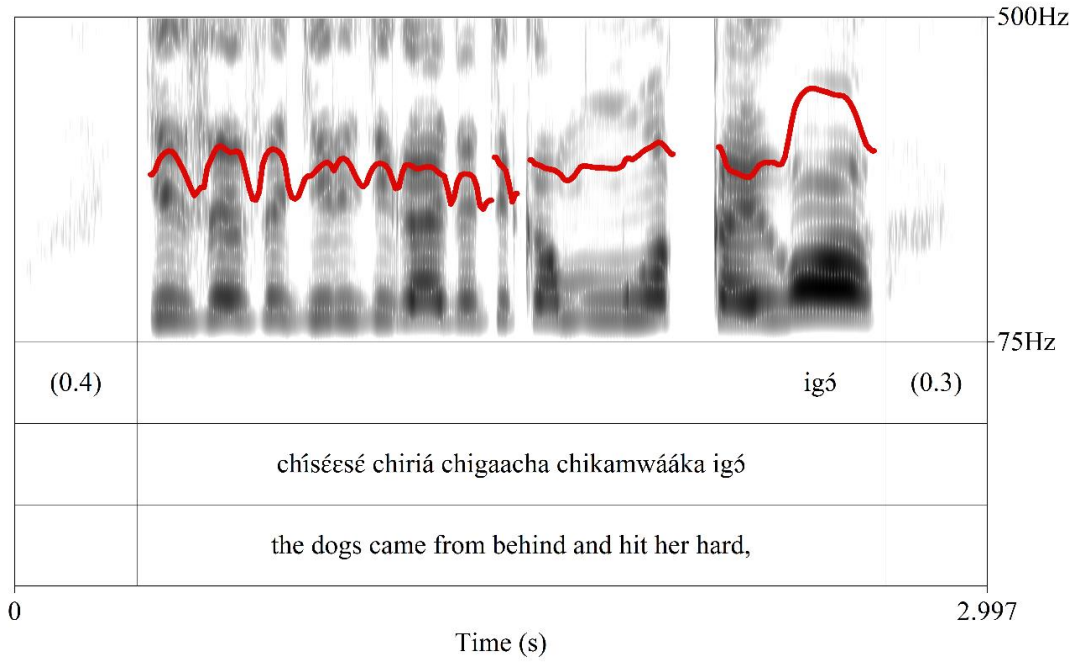
The above examples make clear that vowel elision can be manipulated to create minor prosodic breaks without sacrificing cohesion, thus yielding a layered prosodic structure. Lack of vowel elision is a minor prosodic discontinuity that reflects an equally minor disconnectedness between two segments of speech. Pause, on the other hand, creates a much stronger prosodic discontinuity between units of speech while simultaneously giving strong cohesion to the units that it separates, so that even when lack of vowel elision creates a minor prosodic break, lack of pause ensures that the utterance is still treated as a cohesive entity.

5.3. Prosodic Accent

In the definition of prosody given in §3, I stated that prosody is not just about signaling the kinds of relations that hold between segments of discourse, but also their relative prominence. One of the most important ways in which this prominence is realized is via prosodic accent, the use of phonetic-phonological cues to give some point in the discourse greater prominence in the discourse (van der Hulst 2010). The particular phonetic-phonological cues that create this prominence vary from language to language. In English, they are the bundle of cues that yield the perception of stress, while in Ékegusií the most

important cue is pitch. Take the canonical example of prosodic accent in Ékegusií shown in Figure 8.

Figure 8. Prosodically accented word in Ékegusií



The intonational contour on the final word of this phrase, *igɔ́* 'thus, that way, like that', is significantly higher than would be expected for a simple lexical H tone – indeed, it is higher in pitch than any other H tone in the utterance. This large pitch excursion is not due to a terminal contour (transitional continuity/boundary tone), because for this speaker a rising terminal contour is a sharp rise to a pitch peak at the end of the syllable, as can be seen clearly in Figure 4 above.

What the prosodic prominence is marking in this case is not a transitional boundary but rather the relative discourse prominence of *igɔ́*. Here, *igɔ́* is functioning as an intensifier, and by prosodically accenting the word, the narrator is indicating that the dogs hit the person extremely hard. This prosodic accent made the word and its semantics (intensification)

prominent enough that my translator Gladys included the adverb ‘hard’ in her translation even though there is no word meaning ‘hard’ in the Ékegusií (*igɔ́* is typically translated as ‘like that’, or not translated at all).

Notice that it would be impossible to interpret the prosodic accent on *igɔ́* as accent if it appeared in isolation. There would be no accompanying prosodic context against which to judge whether *igɔ́* were more or less prominent. Since cohesion occurs wherever the interpretation of one element in the discourse is dependent on that of another (Halliday & Hasan 1976:4), even these cases of prosodic accent function to add cohesion to the discourse. They create a contrast between what is being foregrounded and what is being backgrounded, creating a cohesive tie between the two.

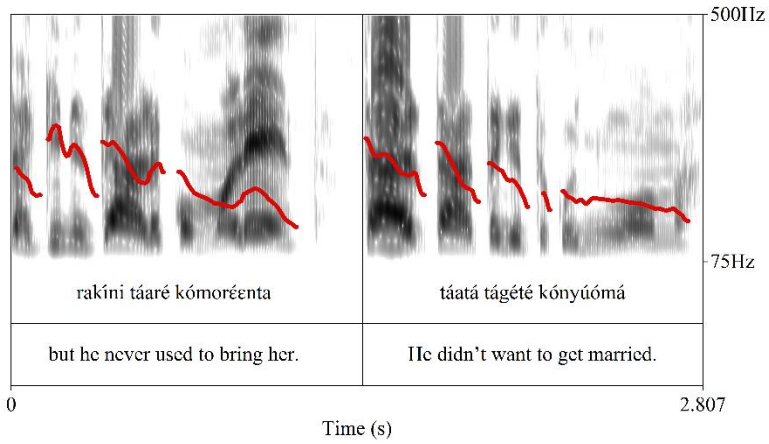
5.4. Pitch Reset

While prosody research has paid a great deal of attention to the ends of prosodic units, there has been much less attention to their beginnings (though see Couper-Kuhlen (2004)). The concept of boundary tones, for example, is investigated almost exclusively with respect to the intonational contours near the ends of phrases rather than at their beginnings. The exception to this is research on pitch reset (Amir, Silber-Varod & Izre’el 2004; Ladd 1988; Oliveira 2003; Scheutze-Coburn, Shapley & Weber 1991), the difference in pitch between two adjacent units. In this section, I discuss some difficulties in defining and operationalizing pitch reset for naturalistic discourse, using my Ékegusií data for exemplification. I suggest that these difficulties can be at least partially overcome if we understand pitch reset less as a mechanistic resetting of pitch due primarily to production considerations, and more as a functional means that speakers have of signaling the conceptual relatedness between units of speech.

Pitch reset is generally agreed to be a key feature defining the boundaries between prosodic units (Gussenhoven 2004:113; Himmelman & Ladd 2008:252; Ladd 1986; 1988; Oliveira 2003). However, the way pitch reset is defined, and the domain it applies to, varies from one researcher to another. Many researchers leave it undefined, appealing to common knowledge of what admittedly seems like a rather intuitive concept. Some researchers focus on the difference in pitch between two adjacent units (Genetti & Slater 2004; Kong 2004; Lin 2009; Oliveira 2003; Swerts 1997; Yufang & Bei 2002), while others focus on a change of the overall slope or trendline of the pitch declination over the phrase (Gussenhoven 2004; Ladd 1988; Matsumoto 2003; Scheutze-Coburn, Shapley & Weber 1991); Connell (2001) defines reset as when the pitch returns to its original height. Some researchers focus on pitch reset as a property of prosodic phrases, others as a property of prosodic sentences (elsewhere *declination units* (Scheutze-Coburn, Shapley & Weber 1991) or *utterances*), although there is general acknowledgement that pitch reset occurs at the boundaries of both types of units, and there have been several studies showing that the degree of reset corresponds to the level of that boundary in the prosodic hierarchy (Ladd 1988; Oliveira 2003; Swerts 1997). Still, there is no consensus on what actually constitutes pitch reset. Moreover, few studies examine pitch reset in a tonal language in a way that makes clear how the reset is being operationalized (though see Yufang & Bei 2002).

The canonical realization of pitch reset involves a gradual declination in pitch over the course of an utterance. Then, when the next phrase starts, the pitch jumps sharply back up, allowing the declination process to repeat. A few – and only a few – such clear cases occur in the Ékegusií narratives. One case is shown in Figure 9.

Figure 9. A canonical case of pitch reset in Ékegusií

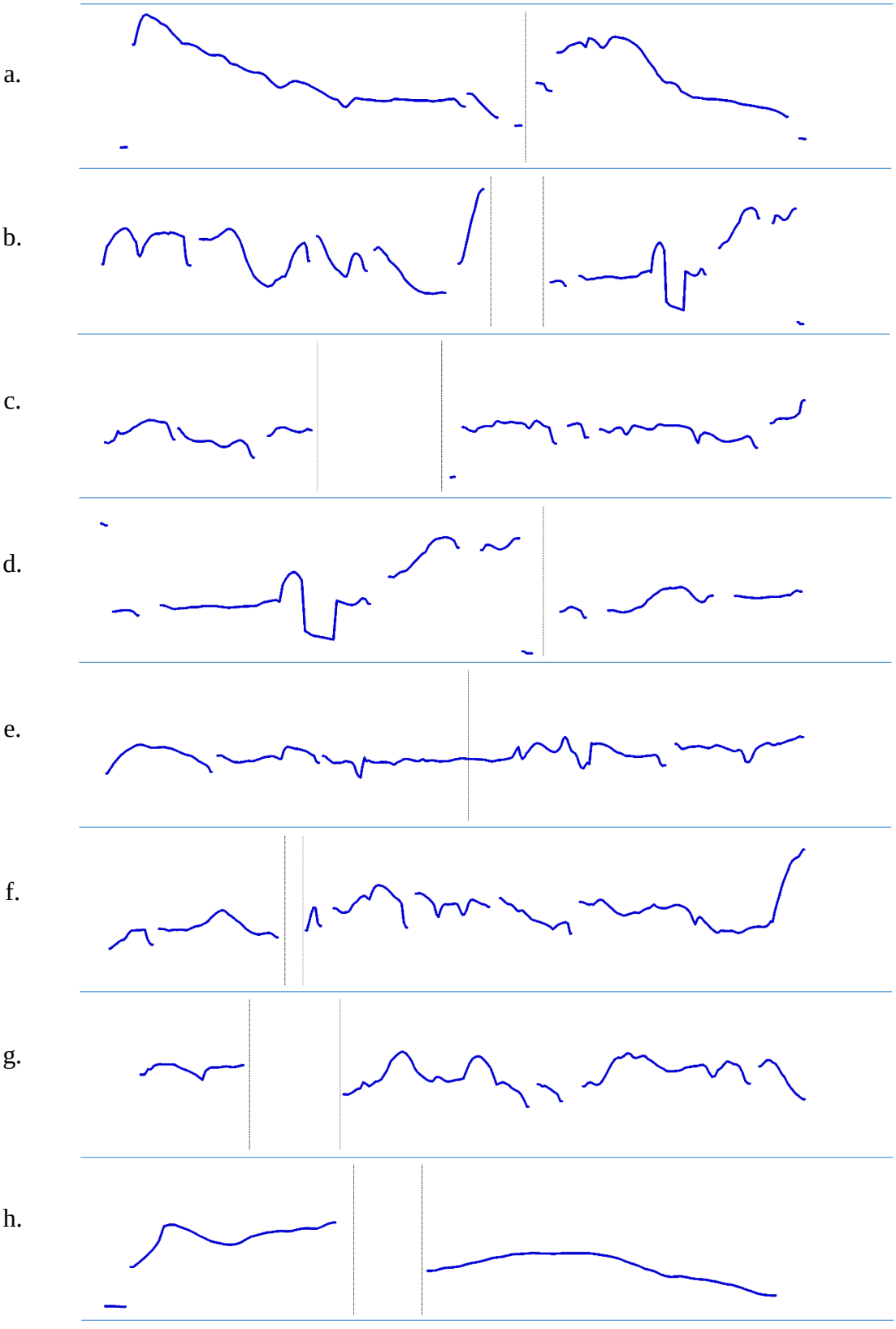


Despite a mix of both high and low tones, the overall declination trend in each phrase is quite clear. But what, precisely, are we picking out when we perceive this pattern of reset? Technically, there are *seven* cases of pitch reset in this example rather than one. Why does only the one in the center count as pitch reset? In this example, the reset happens to co-occur with pause, but this is not a necessary feature of pitch reset generally, as Figure 10a, where there is no pause accompanying the pitch reset, shows. (All the examples in Figure 10 show just the pitch traces and phrase boundaries, and not the transcriptions, for reasons of space. The pitch range of each example is 75 Hz – 500 Hz.)

The reasons why clear cases of pitch reset occur so infrequently in the data are manifold. For starters, the presence of a terminal rising contour may interrupt any declination trend, as in (10b). There may simply be no significant change in pitch from the end of one utterance to the start of another, as in (10c). There may even be a pitch reset down rather than up (10d). Longer utterances are often characterized by a flat contour or even gradual rise rather than a declination (10e). It is also the case that reported speech shows a higher pitch register than the surrounding speech, as in the second utterance in (10f) (Genetti & Hieber 2015). In some

cases, an utterance may simply be too short to establish a trendline, like the first utterance in (10g). Finally, the presence of a prosodically accented syllable can also obscure any pattern of trendline or reset, as in the first utterance in (10h).

Figure 10. Types of phrases that prevent identification of pitch reset in Ékegusií



Given all these problematic cases, the actual number of cases of canonical pitch reset that can be identified in the Ékegusií narratives is quite low. If pitch reset is supposed to be an important cue for prosodic structure (Gussenhoven 2004:113), but it can only be reliably identified in a small proportion of phrase boundaries, and it occurs with high frequency *within* phrases (as seen in Figure 9 above), can we really say that pitch reset tells us anything useful about prosodic phrasing?

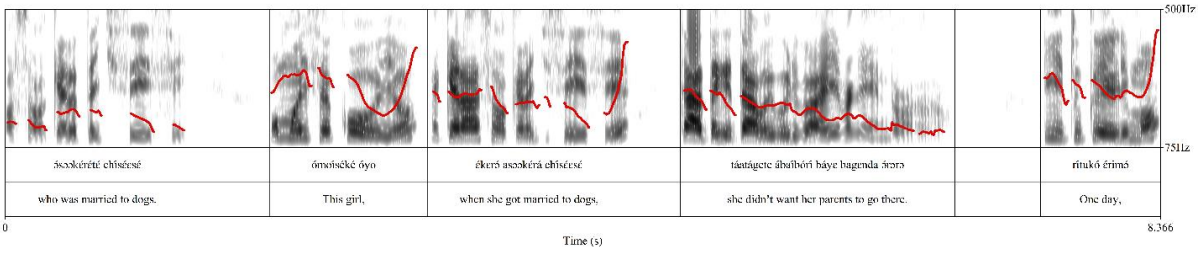
I do believe we can still talk usefully about pitch reset, and how it contributes to prosodic phrasing, if we understand it as a more holistic property of how hearers perceive phrases, abstracting away from the parts of the pitch contour that are irrelevant to the overall trend, such as the terminal contour, modulations due to lexical pitch or prosodic accent, and changes in register. What counts as a declination in one type of intonational contour may be very different from what counts as a declination in another type. If pitch reset has a role to play in the identification of prosodic boundaries, it has that role only by virtue of being mediated through numerous other phenomena affecting the overall pitch contour of the phrase.

Pitch reset is often viewed as a rather mechanistic resetting to a speaker's 'baseline' or default pitch (Cole 2015:15; Myers 1996). While terminal contour is thought to play an important role in signaling the type of phrase and its relationship to the ones around it, pitch reset and other initial cues are not viewed as having a particularly strong functional basis. Perhaps this is because the beginnings of phrases are assumed to be rather obvious – listeners know when an utterance has begun because they hear segmental content. But some researchers have provided strong evidence that more is happening at utterance beginnings, functionally speaking, than is typically recognized. Some earlier work in prosody showed

that the degree of pitch reset signaled the boundaries of smaller or larger units of discourse (Ladd 1986; 1988). More recently, Couper-Kuhlen (2004) shows that the extent of pitch reset signals how tightly conceptually connected the new unit of speech is to the old one. Couper-Kuhlen examines "those crucial moments in conversation when sequences have reached recognizable closure and a next turn could begin something new – or not" and demonstrates that prosody is a key means by which speakers indicate which of the two paths they are following (Couper-Kuhlen 2004:335). Wichmann (2000) likewise introduces the notion of *onset depression*, wherein some pitch resets are made less drastic than others. Wichmann shows that onset depression is motivated by the fact that speakers need to indicate a new discourse topic, thus creating a prosodic break by means of a pitch reset, but simultaneously need to show the degree of relatedness between that topic and the previous one, thus making the pitch reset smaller than it would otherwise be.

We see exactly this pattern in the Ékegusií narratives, where little or no pitch reset signals a close cohesive relation between segments of speech, and larger-sized resets indicate greater discontinuities in the narrative. Consider the relatively straightforward example below.

Figure 11. Greater and lesser degrees of pitch reset indicating greater and lesser degrees of breaks in the discourse



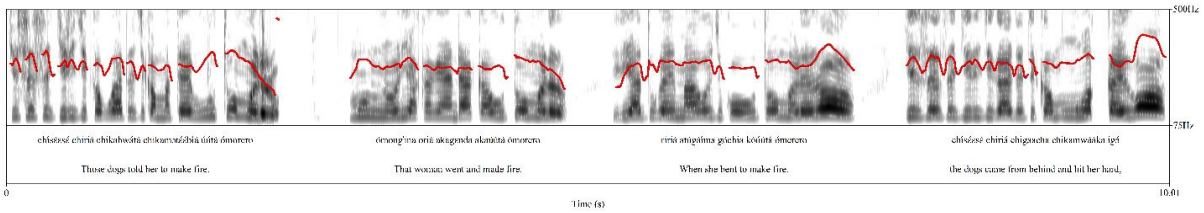
The first feature to notice in this excerpt is that there are resets both down and up. The pitch resets up occur at the two larger pauses in the selection, while the pitch resets down occur

with much shorter pauses (only 0.05s for the second pause). Or, if one wants to ignore the terminal contour for the purpose of examining reset, one could say that the second and third pauses in this selection actually show a reset up. Even so, the pitch resets at those pauses are significantly smaller than those at the first and last pauses, so that the degree of pitch reset correlates with the length of the pause. Why the larger prosodic break at the first and last pauses here? The first pause signals the end of the introduction of the characters, and the beginning of the complicating action. Likewise the last pause finishes setting up the background against which the complicating action is set. The girl does not want her parents to come visit, and in the next utterance we're told that her mother sets out to visit. So these two prosodic break signal accordingly large transition points in the narrative.

The minor pitch resets also have important functions. The first minor pitch reset above (at the beginning of the third utterance) signals that the topic phrase (*ómoiséké óyo* 'this girl') is complete and a new idea is starting. The second minor pitch reset (at the beginning of the fourth utterance) signals that the adverbial 'when' clause is complete, and that that clause is the background against which the main clause is situated.

Lack of pitch reset has the opposite function, signaling a greater degree of narrative continuity and relatedness. Consider the extended excerpt taken from the climax of 'A Girl Who Got Married to Dogs' (lines 53–59 from example 0 above):

Figure 12. Lack of pitch reset indicating narrative continuity

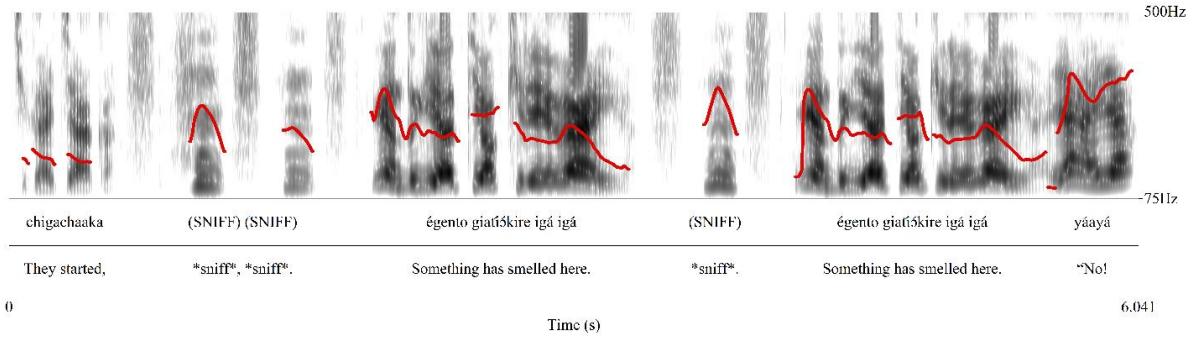


All of these utterances are pronounced on a high register, with a brief terminal fall on the first two phrases, a terminal rise on the third, and the prosodically-accented *igɔ́* (discussed above) obscuring any terminal contour at the end of the fourth. Overall, the terminal contours of these utterances are not nearly as exaggerated as is usually the case for this speaker, with the result that there is little to no pitch reset between utterances. This neatly aligns with the climactic nature of this portion of the text, where the narrator is construing these events as a chain of actions happening in rapid succession, building in speed and anticipation until the final point of the climax (the next utterance, not pictured here, where the woman dies). Thus the intonational continuity of these units signals their cohesiveness as part of a chain of connected actions, and it is left to pause to signal the slightly less significant breaks between each event in the sequence. Just as it was shown that vowel elision creates a more minor prosodic break than pause, here we see that pauses likewise constitute more minor prosodic breaks than major pitch resets. These and other prosodic features come together to create a hierarchical prosodic structure that in turn reflects the hierarchical structure of events and subevents, discourse topics and subtopics in the narrative. It is this prosodic connectedness and structure that is partially but significantly responsible for the cohesive nature of texts. This example also nicely illustrates that the units picked out by prosody do not necessarily align with syntactic units. Here, there is no significant prosodic difference between utterances that are entire sentences and ones that are just clauses. This supports the understanding of prosody put forward here, where its major function is to give structure to discourse, not syntax per se.

5.5. Isotony

Isotony, also called *tonal parallelism* (Wichmann 2000:85–93) (though here I use *intonational parallelism*), is defined by Du Bois (2014:119) as 'the realization of the same tune on successive prosodic phrases'. Isotony is perhaps the best exemplar of how prosody creates cohesive ties in text. By repeating the same intonational contour across two different stretches of speech, the speaker explicitly highlights some similarity between them, whether in function, form, or semantic content. At the same time, each repeated segment of speech constitutes a cohesive unit, by virtue of the fact that its parts are repeated together. Likewise, the edges of any repeated units in isotony create prosodic breaks or 'transition zones' at that those points. In total, I identified 155 instances of isotony in the data, or approximately 20% of utterances. Isotony therefore appears to be a robust narrative technique in these stories.

Figure 13. Basic isotony in Ékegusií

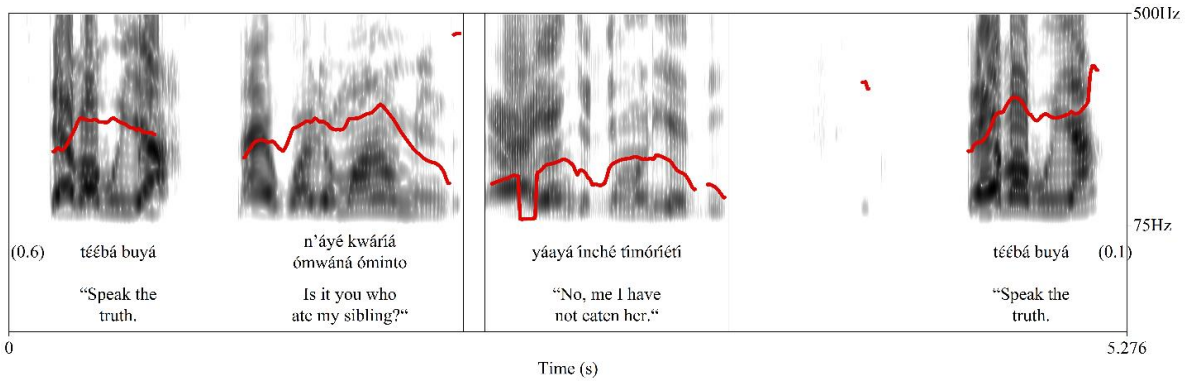


A simple example of isotony is given in Figure 13. Here, the lexical content of the third utterance is repeated along with its intonational contour. Notice that the isotony here parallels the way the entire set of actions is repeated in the narrative. They (the dogs) sniff around, then say that they smell something. They sniff around again, and again say that they smell something. The fact that the intonational contour is repeated here is crucial to the interpretation, and not just a mere byproduct of the lexical repetition. Despite the single

repetition, the narrative effect created by this intonational parallelism is one of a continuous, ongoing action. The dogs are going about sniffing things, not just sniffing twice. The phrase is not being repeated just for emphasis – it is not on a higher register and it does not exhibit any prosodic accent. In addition, the denial that there is anything to smell (*yáayá* 'No!') comes after the repeated phrase, not before, so the dogs are not insistently repeating the phrase in response to the denial. Instead, the repetition functions as a kind of tonal reduplication, parallel to certain types of morphological reduplication where the doubling function as a progressive rather than an intensifier. If the intonational contour of the second phrase were different, its narrative function would be different as well.

This fact becomes salient if we look at a second example of lexical repetition, this time without intonational parallelism:

Figure 14. Lexical repetition without isotony

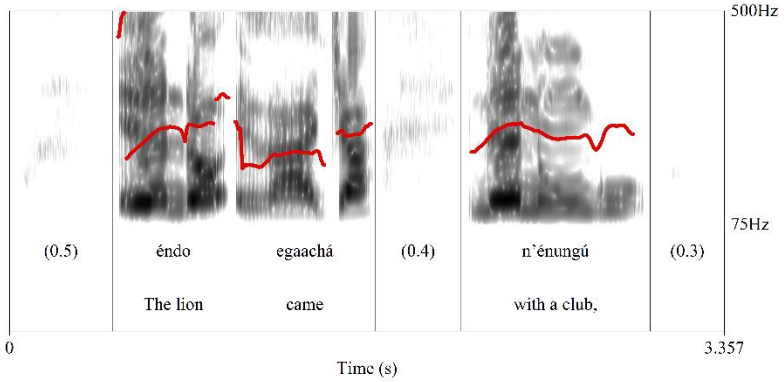


In this example, the first character repeats the phrase *téébа buyá* 'speak the truth', but with a different intonational contour. In the first instance, the phrase shows a high register and terminal fall indicative of commands. The second instance is spoken with a gradual rise, as a warning. The fact that these two phrases are not uttered using intonational parallelism therefore significantly changes their interpretation. Mere repetition of lexical content, then, is

not in itself enough to create narrative parallelism – prosodic parallelism is equally as important.

While isotony can be a property of prosodic phrases whose segmental content is repeated verbatim, it is certainly possible for a speaker to repeat just the intonational contour while changing the lexical content (i.e. using different words but keeping the same intonational contour across phrases). One such example is shown in Figure 15, where each word is pronounced under a separate intonation contour with a gradual rise. This is not the sharp rise of a continuing terminal contour, but rather a stylistic effect meant to emphasize the ominous nature of Lion stalking Rabbit with a club.

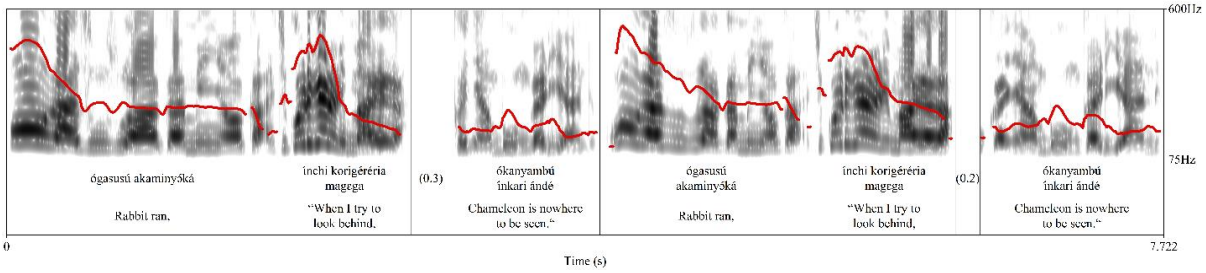
Figure 15. Isotony without lexical repetition



Interestingly, there are even cases of isotony that stretch across more than one prosodic phrase, as Figure 16 illustrates. In this example, the prosody of the entire initial three phrases repeats. While the isotonic units also happen to repeat their lexical content, the fact that the speaker chose to use the same lexical content *as well as* the same prosodic structure highlights the parallel structure and discourse functions of these phrases. This example also provides additional evidence for the hierarchical nature of prosodic structures. While pitch reset and pause create five clear prosodic breaks in this example, isotony suggests an even

stronger break between the repeated sections, thus adding another layer of prosodic structure to this selection.

Figure 16. Isotony across multiple prosodic phrases



The above examples make clear the strongly cohesion-building nature of isotony. This kind of intonational parallelism creates a concomitant parallelism in the structure of the discourse, thus making explicit the connections between different segments of the text.

5.6. Intonational Contour

Intonation has long occupied a privileged place in research on prosody (cf. §2). It is generally agreed to have sentence-level or discourse-level functions, and to demarcate prosodic units that stretch over extended segments of speech. The prosodic units picked out by intonation go by many names: *intonation group* (Cruttendon 1986), *intonation unit* (Chafe 1988), *intonational phrase* (Nespor & Vogel 1986; 2007; Selkirk 1986), *prosodic phrase* (Swerts 1997), *prosodic unit* (Genetti & Slater 2004), *tone group* (Halliday 1967), or *tone unit* (Crystal 1969). What allows listeners to identify these segments of speech as cohesive intonational units is the fact that certain intonational contours are recurring and have conventionalized meanings (Cruttendon 1986:9; Fox 2000:269; Wennerstrom 2001:17). The fact that listeners can recognize these intonational patterns across speakers with different pitch ranges, or the same speaker at different levels of excitement, or in phrases both with

and without prosodically accented words, suggests that listeners perceive these intonational contours in a holistic, gestalt fashion, abstracting away from numerous phonetic details. As Wichmann (2000:9–10) states, "we, as knowers of a language, learn to hear some differences and ignore others. It is part of our knowledge of a language which causes us to distinguish between linguistically significant and not significant acoustic information."

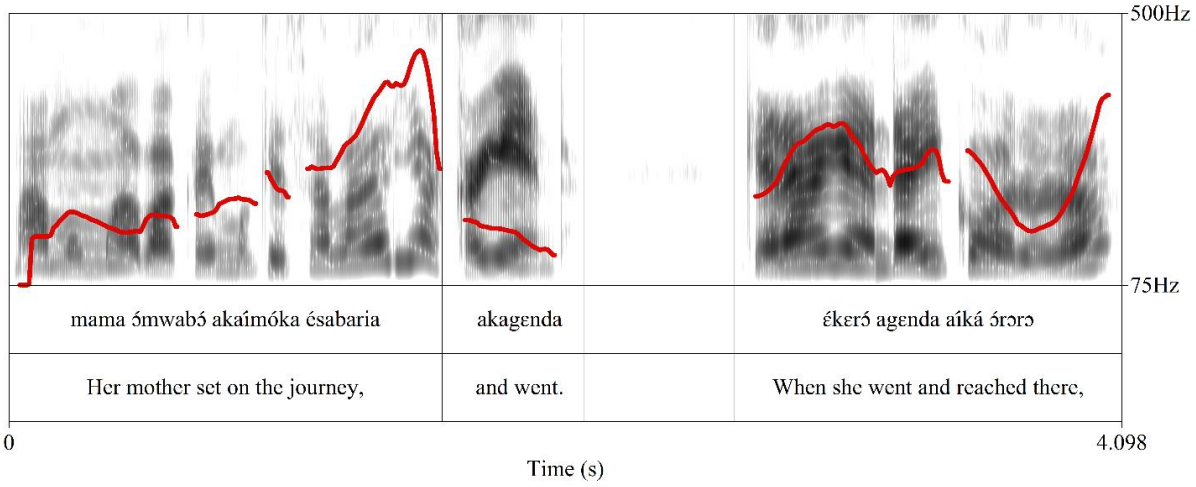
In this section, I wish to exemplify how it is that intonation units, defined as "a sequence of words combined under a single, coherent intonation contour" (Chafe 1988:22), are used to create cohesion in discourse. But first, a few initial caveats are in order: First, while many researchers consider additional prosodic cues like pause and pitch reset when demarcating the boundaries of intonation units, my focus in this section will be more narrowly aimed at the coherent intonational contours mentioned by Chafe. How do these lend cohesion to the discourse, independent of other prosodic features? A separate question is then how intonational contours work in tandem with other prosodic features to create a hierarchy of prosodic cohesion. We have already seen some examples of prosodic features interacting to create hierarchical structure in previous sections; this section will also continue to elaborate on this theme.

Much prosody research, and especially that stemming from the Autosegmental framework (Pierrehumbert 1980; Nespor & Vogel 2007), treats intonational contours as compositional. In this approach, the intonational contour of an utterance is viewed as the phonetic implementation of a small number of phonologically-salient features, such as final boundary tone or pitch-accent, plus variation due to 'extralinguistic' features of the utterance like speaker attitude or emotional state, as well as production considerations like pitch

declination across the utterance. Especially important are the phrase-final H or L pitch target (or a HL or LH combination) known as 'boundary tones' (Pierrehumbert 1980).

One difficulty with a compositional approach to intonation contours, however, is that a H or L description of terminal contours is often insufficient to capture their complexity (Szczepek Reed 2011), or the great variety of conventionalized intonational patterns that may occur earlier in the utterance. Consider the difference between the 'boundary tone' of the first and last utterance of Figure 17, for example.

Figure 17. Two different types of terminal contours ('boundary tones') in Ékegusií

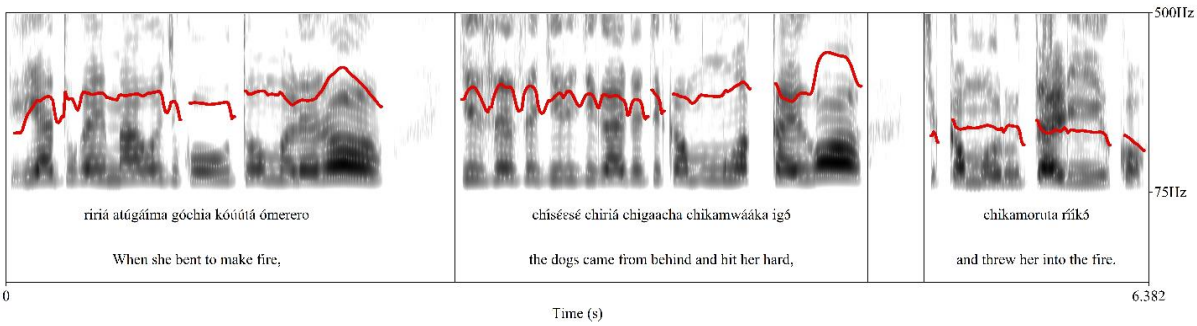


The first utterance exhibits a gradual rise in pitch across the phrase, while the last phrase exhibits a gradual fall until the final syllable, at which point there is a sharp rise. While both utterances end in a H pitch, their overall intonational patterns are quite different, indeed exact opposites. A compositional approach can of course analyze these cases as a H boundary tone and a LH boundary tone respectively. But an alternative approach, and the one I will adopt here, is to say that what speakers and listeners attend to is not specific pitch targets at different points in the phrase, but rather the overall pattern of the intonation contour as a

whole, and that these different contour patterns have conventionalized, each with their own discourse function.

An additional motivation for adopting this approach is the important discourse function of register. For example, the short utterance in the middle of Figure 17, *akagɛnda*, is spoken on a low register and pronounced rapidly, a technique this speaker uses to conclude dramatic buildups in the narrative. Consider again part of the excerpt from Figure 12, this time with concluding event included:

Figure 18. Use of high and low register in Ékegusií narrative

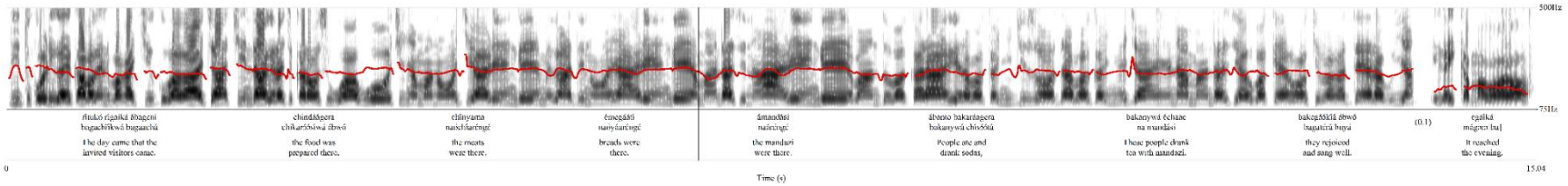


The narrator drops to a low register for the final phrase, the concluding event following five other intonation units that lead up to it. Those preceding intonation units are themselves pronounced on a high, level register with little to no pitch declination, a technique this narrator uses for sequences of events (not just those leading up to a climax). Like the two utterances in Figure 17, the first utterance in Figure 18 also ends in a H 'boundary tone'. Each of these intonational patterns – a gradual rise, a gradual fall with a sharp final rise, and a high level register with a final H – all end in a final H, but otherwise differ significantly in their intonational pattern. Each pattern has a conventionalized, meaningful discourse function, and should therefore be considered phonological rather than merely phonetic. Thus rather than focus on just boundary phenomena when segmenting discourse into intonation

units, in this section I utilize a more holistic approach that looks at the intonation contour of the entire intonation unit, and attempts to correlate that contour type with its discourse function.

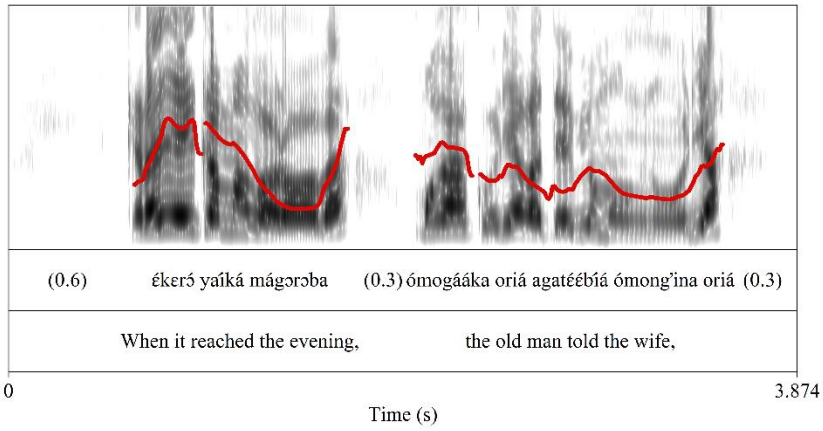
One type of intonational contour that is fairly common in the texts is the high level register with a final rise just mentioned. Below is another (rather remarkable) example of this contour used in a sequence. Close inspection of the intonation contours in these phrases reveals a rise in pitch – higher than other H tones in the utterance – on the last one or two syllables of each phrase. It is clear from the mundane nature of the items in the sequence (food items for a festive gathering) that this high register has nothing to do with a climactic turn of events in the narrative. Instead, this contour type appears to indicate sequentiality, and in this way contributes cohesion to the text by signaling the relation that the present intonation unit has to those around it. Plus, the entire chain of utterances linked together by high level contours forms a cohesive unit in itself, dividing the narrative into larger episodes. As before, this larger chain of events is concluded on a low register, creating a marked prosodic contrast to the preceding phrases and signaling that the sequence is complete. Once again we see how prosody works to create both prosodic and discourse hierarchy: individual intonation units are chained together into a larger prosodic structure ending on a low register, and each intonation unit represents a single idea or action in a larger, narratively-related chain of events. It is data like these that have led many prosody researchers to recognize the existence of larger-level prosodic units above the intonation unit, which are variously termed paratones (Yule 1980), declination units (Scheutze-Coburn, Shapley & Weber 1991), or prosodic sentences (Genetti & Slater 2004).

Figure 19. Chain of intonation units with a high level register



By far the most common intonation contour in the stories is the gradual fall ending in a sharp rise on the final syllable. We have already seen several examples of this contour type in use (Figures 4, 11, 17), and the observant reader may have noticed that this contour type occurs with temporal adverbial clauses like ‘one day’ or ‘when …’. This is its primary function in these texts. The events denoted by the phrase on which this contour type is overlaid are generally either temporally prior to the following intonation unit, or mark the point of departure for the following action, rather than being simultaneous with it. A canonical example is shown in Figure 20.

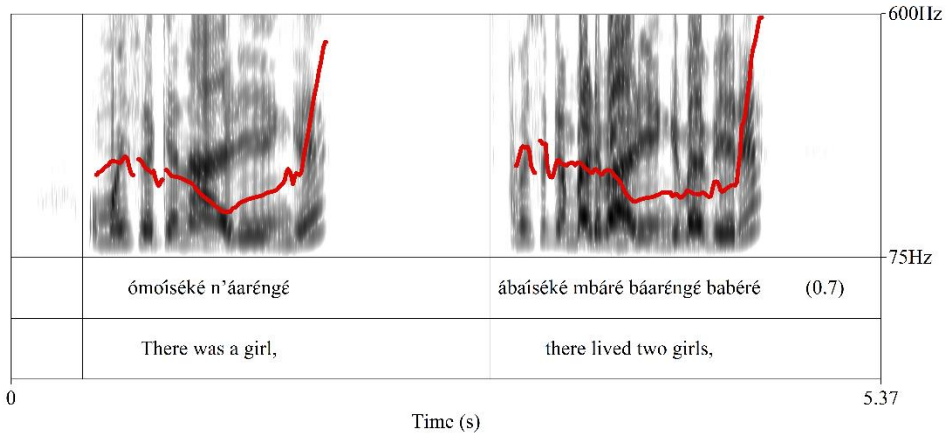
Figure 20. The terminal sharp rise contour type, used with adverbial ‘when’ clauses



This contour type is also used to introduce brand new topics or characters into the discourse, or when there is a major shift in discourse topic in the narrative. This fact is no coincidence. Haiman (1978) demonstrates the logical and diachronic connections between conditionals (both if- and when-conditionals) and topics, in that both become presuppositions of the discourse to follow. They become, in the framework used here, the background information against which the forthcoming new information is set. Especially given that

conditionals and topics frequently share morphological markers, it should not be surprising to discover that certain intonational patterns also exhibit this syncretism in functions. Figure 21 shows how this contour type is used to introduce characters at the beginning of these stories. Another excellent example, showing both the topic marking and temporal conditional functions side by side, is Figure 11 above.

Figure 21. The sharp rising contour type as a topic marker



The sharp rising contour is thus the perfect illustration of cohesion at work, because its primary function is to mark that one stretch of speech should be considered the background for the next, just as was described in §3. In doing so, it creates a cataphoric (forward-looking) cohesive tie to the next unit of speech. It both signals the transition from one unit of discourse to the next and simultaneously indicates the type of meaning relation that holds between them.

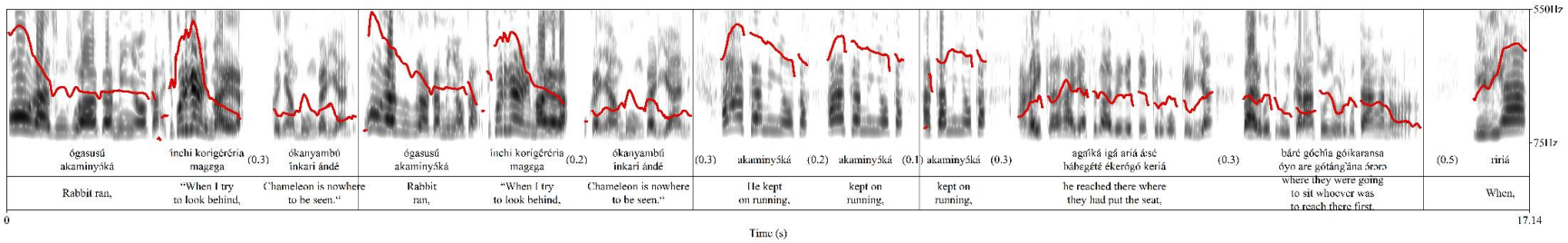
6. Discussion & Conclusion

In her excellent study of the discourse functions of prosody, Wennerstrom argues that "there is an English-specific system of intonation built upon a rhythmic foundation that

functions as a “grammar of cohesion.” This means that prosody, particularly intonation, contributes information about connections among constituents in discourse, conveying meaning beyond what is provided through lexical and syntactic systems” (Wennerstrom 2001:7). This thesis has attempted to expand on this insight, taking cohesion to be the defining characteristic of prosody, and showing how each of the features generally considered prosodic have important roles to play in the structuring of discourse.

It was seen that this cohesive structuring can be accomplished in a variety of ways. Consider the final example in , which exemplifies all six of these prosodic features

Figure 22. Examples of the cohesive functions of prosody in Ékegusií narrative



We saw above how pauses segment the discourse into cohesive units, and the length of those pauses indicates the degree of narrative discontinuity between one unit and the next. In , we likewise see that the longest pause comes after the chain of statements about running, and the character has reached the finish line (the seat). The shortest pauses, on the other hand, are between the instances of *akaminyɔ́ká* 'he kept on running' repeated in sequence. The repetition of this verb does not represent a sequence of three distinct actions, but rather the ongoing and extended nature of the action. The shorter pauses between them index their greater degree of cohesiveness as a single, ongoing action rather than three separate ones.

Next, vowel elision (or its absence) was shown to create minor prosodic breaks which the speaker used to manipulate the perceived connectedness of points in the narrative, thereby adjust its pacing. This is straightforwardly demonstrated in by the fact that each utterance-final word ends in a full vowel rather than an elided one, and utterance-medial words elide their final vowels. The exception is the final vowel of the first instance of *akaminyɔ́ká*, because this occurs at the transition into reported speech, creating a minor prosodic break.

Then, we saw how prosodic accent functions to indicate the relative prominence of different points in the text, creating cohesive ties between prominent and non-prominent information. In , the initial syllable the three *akaminyɔ́ká*'s in sequence, and the initial syllable of *ógasusú* 'rabbit' is prosodically accented with a significantly higher pitch and lengthened vowel, and this accent serves an intensifying function for the events in the phrase ('Rabbit *really* ran'). Without the neighboring prosodic context for juxtaposition, these prosodic accents would not be interpretable as such.

I next showed that the identification of pitch reset is not as straightforward as is often conveyed, and that it is best viewed as a more holistic property of entire utterances. I then

demonstrated how the degree of pitch reset correlates with the degree of the narrative break at that point, in line with previous research by Couper-Kuhlen (2004). In , for instance, we see the relatively small to nonexistent pitch change between the repeated *akaminyɔ́ká*'s, but a very large pitch excursion when the narrative shifts from reported speech to a sequence of statements about the race and running (after *ókanyambú ínkari ándé*).

Next, I illustrated the way that isotony, or intonational parallelism, creates explicit relations of similarity between units of discourse in a way that lexical repetition alone cannot accomplish. It is the intonational parallelism with the repeated *akaminyɔ́ká*'s in that allows the listener to interpret them as a cohesive sequence. Moreover, this same intonational pattern occurs on the first phrase, *ógasusú akaminyɔ́ká*, creating parallelism even when the lexical content of the utterances differs.

Finally, the discussion of intonational contours showed how prosody fulfills the pivotal need that speakers have of backgrounding sections of discourse and signaling the transitions from one discourse topic to the next. Moreover, the fact that a stretch of speech falls under a recognizable contour pattern lends cohesion to that stretch of speech as a unit. In , we once again see the pattern of a sequence of utterances on a high level register (the three *akaminyɔ́ká*'s), followed by a rapidly-spoken, low register conclusion, indicating to the listener that the entire chain of utterances should be considered a cohesive narrative episode or discourse topic – what some have called a *prosodic sentence* (Genetti & Slater 2004) or *paratone* (Yule 1980).

Several themes emerged in the discussions of these features. Many of the features create transition zones ('boundaries' or 'cesuras'; Barth-Weingarten (2013)) of different strengths that signal different degrees of relatedness between portions of the text (Barth-Weingarten

2013; Swerts 1997). Stronger breaks in the discourse constitute the boundaries between larger episodes in the narrative (prosodic sentences), while less strong but still salient breaks constitute the transition from one discourse topic to the next (prosodic phrases or intonation units; (Chafe 1987)), with the transition points between narratives being of a gradient rather than binary nature (Ladd 1988). In , for example, there is a strong transition zone at the second pause, bolstered by a large pitch reset, that signals not just the edge of two intonation units, but two prosodic sentences. Thus it was shown how different prosodic features interact to create a hierarchy of prosodic structure, which correlate to more or less significant discontinuities in the narrative.

The approach to prosody advanced here has several advantages. First, this approach has the capacity to provide a synthesis between research focusing on the phonological nature of prosody and research emphasizing its interactional nature. When prosody is viewed as purely a matter of phonological structuring or suprasegmentals, it is difficult to understand why there should be a systematic connection between prosodic features and social action. But when the function of prosody is understood as providing coherence to discourse and signaling the relations of meaning between units of speech, those signals become interactionally-relevant (Couper-Kuhlen & Selting; Selting 2010:5). Consider Barth-Weingarten's (2013) method of examining the prosodic cues that appear at points in the conversation where listeners inserted themselves into the discourse: the reason listeners are attentive to the prosodic cues they are is because those cues signal when the speaker is transitioning from one discourse topic to the next (cf. §5.6). This is a natural place for the listener to come in and respond to the topic that has just been finished.

A second advantage of the present approach is its functional-typological nature. In place of list-definition or suprasegmental approaches, I have offered a functional definition of prosody that focuses on the phonetic-phonological features that speakers use – and listeners rely on – to help structure their discourse. That discourse structure can be independently established in a number of ways, whether through the morphosyntactic marking of topic shifts (Clancy & Downing 1987) and relations between clauses (Mithun 2008), or the interactional relevance of different breaks in the discourse (Barth-Weingarten 2013). The definition is flexible enough to allow for the diversity of ways that prosody is realized crosslinguistically, while still providing a means of operationalizing which features should be considered prosodic. By providing a functional, feature-independent and language-independent definition of prosody, I hope to have provided a framework in which future examinations of prosody in a wide variety of languages can be based, and their results usefully compared.

References

Amir, Noam, Vered Silber-Varod & Shlomo Izre'el. 2004. Characteristics of intonation unit boundaries in spontaneous spoken Hebrew: Perception and acoustic correlates. *Speech Prosody 2004, Nara, Japan, March 23-26, 2004: Proceedings*. 677–680.

Barth-Weingarten, Dagmar. 2013. From "intonation units" to cesuring: An alternative approach to the prosodic-phonetic structuring of talk-in-interaction. In Beatrice Szczepek Reed & Geoffrey Raymond (eds.), *Units of talk, units of action*, 91–124. (Studies in Language & Social Interaction 25). Amsterdam: John Benjamins.

Barth-Weingarten, Dagmar & Elisabeth Reber (eds.). 2010. *Prosody in interaction*. (Studies in Discourse & Grammar 23). Amsterdam: John Benjamins.

Bois, John W. Du. 2016. *Representing discourse*. Manuscript.

Bois, John W. Du, Susanna Cumming, Stephan Schuetze-Coburn & Danae Paolino (eds.). 1992. *Discourse transcription*. (Santa Barbara Working Papers in Linguistics 4). Santa Barbara: University of California, Santa Barbara.

Bosire, Kennedy M. & Gladys K. Machogu. 2013. *Authoritative Ekegusii dictionary / Endabaro endabasia y'Ekegusii*. Kenya: Ekegusii Encyclopedia Project.

Brazil, David. 1985. *The communicative value of intonation in English*. Birmingham: English Language Research, University of Birmingham.

Brierley, Claire & Eric Atwell. 2005. An approach for detecting prosodic phrase boundaries in spoken English. *Crossroads*. doi:10.1145/1349332.1349337. http://dx.doi.org/10.1145/1349332.1349337.

Brown, Gillian. 1977. *Listening to spoken English*. London: Longman.

Cammenga, Jelle. 2002. *Phonology and morphology of Ekegusii: A Bantu language of Kenya*. (East African Languages & Dialects 12). Cologne, Germany: Rüdiger Köppe.

Casali, Roderic F. 1997. Vowel elision in hiatus contexts: Which vowel goes? *Language* 73(3). 493–533.

Chafe, Wallace L. 1979. The flow of thought and the flow of language. In Talmy Givón (ed.), *Discourse and syntax*, 159–181. New York: Academic Press.

Chafe, Wallace L. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 21–52. (Typological Studies in Language 11). Amsterdam: John Benjamins.

Chafe, Wallace L. 1988. Linking intonation units in spoken English. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar & discourse*, 1–28. (Typological Studies in Language 18). Amsterdam: John Benjamins.

Chafe, Wallace L. 1993. Prosodic and functional units of language. In Jane A. Edwards & Martin D. Lampert (eds.), *Talking data: Transcription and coding in discourse research*, 33–43. Hillsdale, NJ: Lawrence Erlbaum.

Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Clancy, Patricia M. & Pamela Downing. 1987. The use of wa as a cohesion marker in Japanese oral narratives. In John Hinds, Senko K. Maynard & Shoichi Iwasaki (eds.), *Perspectives on topicalization: The case of Japanese "wa."*, 3–56. (Typological Studies in Language 14). Amsterdam: John Benjamins.

Cole, Jennifer. 2015. Prosody in context: A review. *Language, Cognition & Neuroscience* 3798(January). 37–41. doi:10.1080/23273798.2014.963130.

Connell, Bruce. 2001. Downdrift, downstep, and declination. Paper presented at *Typology of African Prosodic Systems Workshop*, May 18–20, Bielefeld University, Germany.

Couper-Kuhlen, Elizabeth. 2004. Prosody and sequence organization in English conversation: The case of new beginnings. In Elizabeth Couper-Kuhlen & Cecilia E. Ford (eds.), *Sound patterns in interaction*, 335–376. (Typological Studies in Language 62). Amsterdam: John Benjamins.

Couper-Kuhlen, Elizabeth. 2015. Intonation and discourse. In Deborah Tannen, Heidi E. Hamilton & Deborah Schiffrin (eds.), *The handbook of discourse analysis*, 82–104. 2nd ed. (Blackwell Handbooks in Linguistics). Malden, MA: Wiley-Blackwell.

Couper-Kuhlen, Elizabeth & Cecilia E. Ford (eds.). 2004. *Sound patterns in interaction: Cross-linguistic studies from conversation*. (Typological Studies in Language 62).

Amsterdam: John Benjamins.

Couper-Kuhlen, Elizabeth & Margaret Selting. 1996. Towards an interactional perspective on prosody and a prosodic perspective on interaction. In Elizabeth Couper-Kuhlen & Margaret Selting (eds.), *Prosody in conversation: Interactional studies*, 11–56. (Studies in Interactional Sociolinguistics 12). Cambridge, MA: Cambridge University Press.

Couper-Kuhlen, Elizabeth & Margret Selting (eds.). 1996. *Prosody in conversation*. (Studies in Interactional Sociolinguistics 12). Cambridge: Cambridge University Press. doi:10.1016/S0378-2166(97)81938-8. https://books.google.com/books?id=C77MlvXZL0wC&pgis=1.

Croft, William. 1995. Intonation units and grammatical structure. *Linguistics* 33(5). 839–882.

Cruttendon, Alan. 1986. *Intonation*. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Crystal, David. 1969. *Prosodic systems and intonation in English*. (Cambridge Studies in Linguistics 1). Cambridge: Cambridge University Press.

Dilley, L., S. Shattuck-Hufnagel & M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4). 423–444. doi:10.1006/jpho.1996.0023.

Donlay, Chris. 2015. A functional grammar of Khatso. University of California, Santa Barbara.

Downing, Laura J. 2010. An edge-based approach to the alignment of syntactic phases and prosodic phrases. *Transactions of the Philological Society* 108(3). 352–369. doi:10.1111/j.1467-968X.2010.01245.x.

Downing, Laura J. 2013. Issues in the phonology-syntax interface in African languages. In Ọlanikẹ Ọla Orie & Karen W. Sanders (eds.), *ACAL 43*, 26–38. Somerville, MA: Cascadilla Proceedings Project.

Downing, Laura J., Al Mtenje & Bernd Pompino-Marschall. 2004. Prosody and information structure in Chichewa. *ZAS Papers in Linguistics* 37. 167–186.

Fónagy, Ivan & Eva Bérard. 2006. Functions of intonation. In Yuji Kawaguchi, Ivan Fónagy & Tsunekazu Moriguchi (eds.), *Prosody & Syntax*, 19–46. (Usage-Based Linguistic Informatics 3). Amsterdam: John Benjamins.

Ford, Cecilia E. & Elizabeth Couper-Kuhlen. 2004. Conversation in phonetics. In Elizabeth Couper-Kuhlen & Cecilia E. Ford (eds.), *Sound patterns in interaction: Cross-linguistic studies from conversation*, 3–4. (Typological Studies in Language 62). Amsterdam: John Benjamins.

Fox, Anthony. 2000. *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press. doi:10.1353/lan.2002.0109.

Genetti, Carol. 2011a. The tapestry of Dolakha Newar: Chaining, embedding, and the complexity of sentences. *Linguistic Typology* 15(1). 5–24. doi:10.1515/lity.2011.002. http://www.reference-global.com/doi/abs/10.1515/LITY.2011.002.

Genetti, Carol. 2011b. Direct speech reports and the cline of prosodic integration in Dolakha

Newar. *Himalayan Linguistics* 10(1). 55–76.

Genetti, Carol & Daniel W. Hieber. 2015. The start and the end of it: Prosodic marking of speech report boundaries in Dolakha Newar. Paper presented at *South Asian Languages Analysis Roundtable (SALA) 31*, May 14–16, Lancaster, UK.

Genetti, Carol & Keith Slater. 2004. An analysis of syntax and prosody interactions in a Dolakhā Newar rendition of The Mahābhārata. *Himalayan Linguistics* 3. 1–91.

Gordon, Matthew. 2016. *Phonological typology: The cross-linguistic study of sound systems*. Manuscript.

Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Gumperz, John. 1992. Contextualization and understanding. In Alessandro Duranti & C. Goodwin (eds.), *Rethinking context*, 229–252. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. (Research Surveys in Linguistics). Cambridge: Cambridge University Press.

Haiman, John. 1978. Conditionals are topics. *Language* 54(3). 564–589.

Halliday, M. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton.

Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. (English Language Series 9). London: Longman.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. doi:10.1353/lan.2010.0021.

Himmelman, Nikolaus P. & D. Robert Ladd. 2008. Prosodic description: An introduction for fieldworkers. *Language Documentation & Conservation* 2(2). 244–274. http://scholarspace.manoa.hawaii.edu/handle/10125/4345.

Hulst, Harry van der. 2010. Word accent: Terms, typologies and theories. In Harry van der Hulst, Rob Goedemans & Ellen van Zanten (eds.), *A Survey of word cccentual patterns in the languages of the world*, vol. 1, 3–53. Berlin: Mouton de Gruyter. doi:10.1515/9783110198966.

Inoue, Fumio. 2006. Sociolinguistic characteristics of intonation. In Yuji Kawaguchi, Ivan Fónagy & Tsunekazu Moriguchi (eds.), *Prosody & Syntax*, 197–222. (Usage-Based Linguistic Informatics 3). Amsterdam: John Benjamins.

Jakobson, R., G. Fant & Morris Halle. 1951. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Jun, Sun-Ah (ed.). 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Karlsson, Anastasia M., Jan-Olaf Svantesson & David House. 2014. Prosodic boundaries and discourse structure in Kammu. *Proceedings from FONETIK 2014*, 71–76. Stockholm University.

Kong, Eunjong. 2004. The role of pitch range variation in the discourse structure and intonation structure of Korean. *Proceedings of the International Conference on Spoken Language Processing (INTERSPEECH)*, 3017–3020. http://www.ling.ohio-

state.edu/~ejkong/ICSLP_EunjongKong .pdf.

Kratochvil, Paul. 1998. Intonation in Beijing Chinese. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 421–436. Cambridge: Cambridge University Press.

Kumpf, Lorraine E. 1987. The use of pitch phenomena in the structuring of stories. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 189–216. (Typological Studies in Language 11). Amsterdam: John Benjamins.

Labov, William & J. Waletsky. 1967. Narrative analysis: Oral versions of personal experience. *Essays on the verbal and visual arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*, 12–14. Seattle: University of Washington Press.

Ladd, D. Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology* 3(1986). 311. doi:10.1017/S0952675700000671.

Ladd, D. Robert. 1988. Declination "'reset'" and the hierarchical organization of utterances. *The Journal of the Acoustical Society of America* 84(2). 530. doi:10.1121/1.396830.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd ed. (Cambridge Studies in Linguistics 119). Cambridge: Cambridge University Press. doi:10.1016/S0024-3841(98)00007-2.

Laver, John. 1994. *Principles of phonetics*. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Lin, You-Jing. 2009. *Units in Zhuokeji rGyalrong discourse: Prosody and grammar*. PhD dissertation. Santa Barbara, CA: University of California, Santa Barbara.

Lovick, Olga & Siri G. Tuttle. 2012. The prosody of Dena'ina narrative discourse. *International Journal of American Linguistics* 78(3). 293–334.

Matsumoto, Kazuko. 2003. *Intonation units in Japanese conversation: Syntactic, informational, and functional structures*. (Studies in Language Companion Series 65). Amsterdam: John Benjamins.

Max Planck Institute for Psycholinguistics. ELAN. Nijmegan, The Netherlands: The Language Archive. http://tla.mpi.nl/tools/tla-tools/elan/.

Mithun, Marianne. 1996. Prosodic cues to accessibility. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 223–234. (Pragmatics & Beyond 38). Amsterdam: John Benjamins.

Mithun, Marianne. 2008. The extension of dependency beyond the sentence. *Language* 84(1). 69–119.

Myers, Scott. 1996. Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African Linguistics* 25(1). 29–60.

Nash, Carlos M. 2011. *Tone in Ekegusii: A description of nominal and verbal tonology*. PhD dissertation. Santa Barbara, CA: University of California, Santa Barbara.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. (Studies in Generative Grammar 28). Dordrecht: Foris.

Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology: With a new foreword*. (Studies in Generative Grammar 28). Mouton de Gruyter.

Nurse, Derek & Gérard Philippson (eds.). 2003. *The Bantu languages*. (Routledge Language Family Descriptions). London.

Ogden, Richard. 2002. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31(1). 139–152.

Oliveira, Miguel Jr. 2000. *Prosodic features in spontaneous narratives*. PhD dissertation. Simon Fraser University. doi:10.13140/RG.2.1.1272.9042.

Oliveira, Miguel Jr. 2003. Pitch reset as a cue for narrative segmentation. *Proceedings of IP2003, Prosodic Interfaces* (1994). www.iltec.pt/pdf/wpapers/2003-prodip-pitch.pdf.

Park, Joseph Sung-Yul. 2002. Cognitive and interactional motivations for the intonation unit. *Studies in Language* 26(3). 637–680. doi:10.1075/sl.26.3.07par.

Pierrehumbert, Janet. 1980. *The phonetics and phonology of English intonation*. PhD dissertation. Cambridge, MA: MIT.

Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*. (System Development Foundation Benchmark Series). Cambridge, MA: MIT Press.

Scheutze-Coburn, Stephan, Marian Shapley & Elizabeth G. Weber. 1991. Units of intonation in discourse: A comparison of acoustic and auditory analyses. *Language & Speech* 34 (Pt 3). 207–234. doi:10.1177/002383099103400301.

Selkirk, Elisabeth O. 1986. *Phonology and syntax: The relationship between sound and structure*. Cambridge, MA: MIT Press.

Selting, Margaret. 2010. Prosody in interaction: State of the art. In Dagmar Barth-Weingarten, Elisabeth Reber & Margaret Selting (eds.), *Prosody in interaction*, 3–40. (Studies in Grammar & Discourse 23). Amsterdam: John Benjamins.

Swadesh, Morris. 1946. Chitimacha. In Harry Hoijer (ed.), *Linguistic structures of Native America*, 312–336. (Publications in Linguistics 6). New York: Viking Fund.

Swerts, Marc. 1994. *Prosodic features of discourse units*. PhD dissertation. Technische Universiteit Eindhoven. doi:10.6100/IR411593.

Swerts, Marc. 1997. Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America* 101(1). 514–521. doi:10.1121/1.418114.

Swerts, Marc & Ronald Geluykens. 1994. Prosody as a marker of information flow in spoken discourse. *Language & Speech* 37(1). 21–43. doi:10.1177/002383099403700102.

Szczepek Reed, Beatrice. 2011. *Analysing conversation: An introduction to prosody*. Hampshire: Palgrave Macmillan.

Szczepek Reed, Beatrice & Geoffrey Raymond (eds.). 2013. *Units of talk, units of action*. (Studies in Language & Social Interaction 25). Amsterdam: John Benjamins.

Tao, Hongyin. 1996. *Units in Mandarin conversation: Prosody, discourse, and grammar*.

(Studies in Discourse & Grammar 5). Amsterdam: John Benjamins.

Vaissière, J. 1983. Language-independent prosodic features. In Anne Cutler & D. Robert Ladd (eds.), *Prosody: Models & measurements*, 53–66. (Springer Series in Language & Communication 14). Berlin: Springer-Verlag.

Wennerstrom, Ann. 2001. *The music of everyday speech: Prosody and discourse analysis*. Oxford: Oxford University Press.

Wichmann, Anne. 2000. *Intonation in text and discourse: Beginnings, middles and ends*. (Studies in Language & Linguistics). Harlow, England: Longman.

Winkworth, Alison L., Pamela J. Davis, Roger D. Adams & Elizabeth Ellis. 1994. Breathing patterns during spontaneous speech. *Journal of Speech, Language, & Hearing Research* 38. 124–144.

Winkworth, Alison L., Pamela J. Davis, Elizabeth Ellis & Roger D. Adams. 1994. Variability and consistency in speech breathing during reading. *Journal of Speech, Language, & Hearing Research* 37. 535-556.

Yabin, Liu & Li Aijun. 2003. Cues of prosodic boundaries in Chinese spontaneous speech. *International Conference of the Phonetic Sciences 15*, 1269–1272. Barcelona.

Yufang, Yang & Wang Bei. 2002. Acoustic correlates of hierarchical prosodic boundaries in Mandarin. *Speech Prosody 2002*. 1–4.

Yule, George. 1980. Speakers' topics and major paratones. *Lingua* 52. 33–47.