The Prosodic System of Southern Bobo Madaré

by Kate Sherwood

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Linguistics)
in The University of Michigan
2020

Doctoral Committee:

Associate Professor Jelena Krivokapić, Chair
Professor Matthew Gordon, UC Santa Barbara
Professor Jeffrey Heath
Associate Professor Nicholas Henriksen

Kate I. Sherwood

ksher@umich.edu

ORCID iD: 0000-0002-2469-4698

Acknowledgements

There are many people that made the process of getting a PhD possible for me, and even more who made it the positive experience it has been.

First, I want to thank my advisors. Jelena Krivokapić and Jeffrey Heath. Jelena, thank you for introducing me to prosody and all of your support throughout this process. I had no idea what I was getting into when I started to study prosody—and that's been half of the fun. You changed my perspective on language. Jeff, thank you for the amazing opportunity to do fieldwork in Burkina. I always dreamed of doing fieldwork and you helped me make it happen. I'd also like to thank the other members of my committee, Nick Henriksen and Matt Gordon, for their comments and input during this process.

There would be no dissertation without the speakers who I worked with while in Burkina. Thank you to Salimata Sanou, Sata Sanou, Siaka Sanou, and Zakaria Sanou, my four main language consultants. It also would not have been possible without the funding I received for my fieldwork from the African Studies Center, the Rackham Graduate School, and the Linguistics Department.

People often remark that graduate school can be a lonely and isolating time, but this was not my experience. For this, I first want to thank my fellow graduate students—especially Stephen Tyndall, Sagan Blue, and Hayley Heaton, who became good friends. Thank you Stephen for always being willing to talk about cats and for taking care of mine when I was doing fieldwork, thank you Sagan for being my first friend in graduate school, and thank you Hayley for all of the bad TV. Second, thank you to my good friend outside of graduate school, Katie Starbuck., for still

wanting to hang out even though I became really, really boring for a while.

And finally, I'd like to thank my parents. Thank you for understanding me even when I wasn't very understandable. I couldn't have done this without your love and support. I don't have the words for how lucky I am.

Table of Contents

Acknowledgements ii

List of Tables viii

List of Figures x

List of Abbreviations xiv

Transcription Conventions xv

Abstract xvi

Chapter 1 Introduction 1

1.1 Overview 1

1.2 Southern Bobo Madaré 2

1.3 Prior Work on Related Languages 8

1.4 Prosody 12

1.5 Autosegmental-Metrical Theory 15

1.6 Tone and Intonation 17

1.7 Research Questions 22

Chapter 2 Methods 23

2.1 Overview 23

2.1.1 Speakers 24

2.2 Elicitation Tasks 24
2.2.1 2014 elicitation tasks 26
2.2.2 2016-2017 elicitation tasks 28
2.2.2.1 Word stress elicitation task (45 sentences, 136 tokens) 28
2.2.2.2 Phrase prosody elicitation task 1 (seven sentence triplets, 186 tokens) 29
2.2.2.3 Phrase prosody elicitation task 2 (21 sentence pairs, 203 tokens) 30
2.2.2.4 Negated sentences elicitation task (22 sentence pairs, 175 tokens) 32
2.2.3 Questionnaire on Information Structure (205 prompts) 33
2.3 Map Task 34
2.4 Other Spontaneous Speech Tasks 36
Chapter 3 Word-Level Prosody 37
3.1 Overview and Research Questions 37
3.2 Contour Tone Inventory 37
3.3 Phonetic Evidence for Lexical Stress 46
3.4 The Prosodic Foot 55
3.4.1 Distribution of /ə/ 56
3.4.2 Distribution of reduced vowels 59
3.4.3 Distribution of tone melodies 60
3.4.4 Summary of prosodic foot structure 64
3.5 Summary of Word Prosodic Structure 65

Chapter 4 Phrase-Level Prosody 67

4.1 Overview 67

4.2 Research Questions 70

4.2.1 What are the properties of intonational phrases? 70

4.2.2 How is prosody used to mark utterance type? 75

4.2.3 Can other phrase types be identified? 76

4.3 Declarative Sentences 78

4.3.1 Analysis 78

4.3.2 Boundary tone 79

4.3.3 Downtrends 87

4.3.4 Final lengthening 94

4.3.5 Non-modal vowel phonation 96

4.4 Questions 98

4.4.1 Analysis 98

4.4.2 Polar Questions 98

4.4.3 Wh-questions 106

4.4.4 Questions in the 2014 elicitation tasks 111

4.4.5 Questions in the spontaneous speech tasks 114

4.5 Negated Statements 122

4.6 Prosodic Phrase Types 127

4.7 Summary and Discussion 128

Chapter 5 Focus Marking 132

5.1 Overview and Research Questions 132

5.2 Analysis 137

5.3 Results 139

5.3.1 Contrastive focus 141

5.3.2 Narrow focus 145

5.3.3 Broad focus 148

5.3.4 Problematic cases 149

5.4 Discussion 152

Chapter 6 Conclusion 154

6.1 Main Findings 155

6.1.1 Word-level prosody 155

6.1.2 Prosodic phrase structure 156

6.1.3 Focus 158

6.2 Future Directions 158

References 161

List of Tables

Table 1. *Mean Amplitude (dB) of Vowels in Bisyllabic Words with Level Tone Melodies* 48

Table 2. *Mean Duration of Vowels (ms) in Bisyllabic Words with Level Tone Melodies* 50

Table 3. *Mean Amplitude (dB) of Vowels in Bisyllabic Words with Contour Tone Melodies*..... 52

Table 4. *Mean Duration (ms) of Vowels in Bisyllabic Words with Contour Tone Melodies* 52

Table 5. *Elision of /i, u, ɛ/* 59

Table 6. *Productive Tone Melodies Found in Monosyllabic and Bisyllabic Words* 61

Table 7. *F0 Contour in the Final Syllable of Words Ending in an L, M, or H Tone*..... 80

Table 8. *F0 Contour in the Final Syllable of Words Ending in an LM, LH, or HL Tone*..... 85

Table 9. *F0 Contour of Phrase-Final Level Tones in Spontaneous Speech* 86

Table 10. *Mean Duration of Phrase-Medial and Phrase-Final Vowel in Statements*..... 95

Table 11. *Vowel Phonation in Phrase-Final Vowels in Statements* 97

Table 12. *F0 Contour in Final Syllable of Polar Questions*..... 100

Table 13. *Mean Duration of Phrase-Medial and Phrase-Final Vowels in Polar Questions* 102

Table 14. *Duration of Final Vowel in Statements and Polar Questions (z-scores)*..... 103

Table 15. *Final Vowel Phonation in Polar Questions Ending in Level Tones* 105

Table 16. *F0 Contour in Final Syllable of* Wh-*Questions*..... 107

Table 17. *Mean Duration of Phrase-Medial and Phrase-Final Vowel in* Wh-*Questions*..... 108

Table 18. *Mean Duration of Phrase-Final Vowels in Statements, Polar Questions, and* Wh-*Questions*..... 110

Table 19. *Final Vowel Phonation in* Wh-*Questions*..... 110

Table 20. *F0 Contour of Final Syllables in Questions in 2014 Elicitation Data* 112

Table 21. *F0 Contour of Final Syllables in Questions in 2014 Elicitation Data by Speaker*..... 114

Table 22. *Mean F0 (Hz) of Non-Negated and Negated Sentences* 127

List of Figures

Figure 1. *Successive H tones in a phrase (S1, F)* 8

Figure 2. *Phrase-level prosodic structure of English* 14

Figure 3. *First map from map task* 35

Figure 4. *Second map from map task* 36

Figure 5. ma^{31} *'friend' (S3, M)* 41

Figure 6. gba^{51} *'sheep' (S3, M)* 41

Figure 7. tũ^{5} *'market' (S3, M)* 42

Figure 8. dũ^{51} *'chest' (S3, M)* 42

Figure 9. ba^{1}lo^{5} *'iron' (S3, M)* 42

Figure 10. da^{1}lo^{51} *'girl' (S3, M)* 42

Figure 11. sɔ^{5}gɔ^{5} *'panther' produced before an L tone (S2, F)* 43

Figure 12. sɔ^{5}gɔ^{5} *'panther' produced before an L tone (S1, F)* 43

Figure 13. da^{1}lo^{51} *'girl' produced before an L tone (S4, M)* 44

Figure 14. da^{1}lo^{51} *'girl' produced before an L tone (S4, M)* 45

Figure 15. *Amplitude of V1 and V2 in level-toned words (S3)* 48

Figure 16. *Amplitude of V1 and V2 in level-toned words (S4)* 49

Figure 17. *Duration of V1 and V2 in level-toned words (S3)* 50

Figure 18. *Duration of V1 and V2 in level-toned words (S4)* 51

Figure 19. *Amplitude of V1 and V2 in contour-toned words (S3)* 53

Figure 20. *Amplitude of V1 and V2 in contour-toned words (S4)* 53

Figure 21. *Duration of V1 and V2 in contour-toned words (S3)* 54

Figure 22. *Duration of V1 and V2 in contour-toned words (S4)* 54

Figure 23. *Declarative statement ending with L tone (S3, M)* 81

Figure 24. *Declarative statement ending with M tone (S3, M)* 81

Figure 25. *Declarative statement ending with H tone (S3, M)* 82

Figure 26. *Declarative statement ending with M tone with final fall in F0 (S3, M)* 82

Figure 27. *Declarative statement ending with L tone (S3, M)* 83

Figure 28. *Intonational phrase ending in M tone from an interview task (S2, F)* 86

Figure 29. *Declarative statement with lowered F0 during final H tone (S3, M)* 88

Figure 30. *Declarative statement with lowered F0 in phrase-medial and phrase-final LH and H tones (S3, M)* 89

Figure 31. *Declarative statement with lowered F0 in a phrase-medial H tone (S3, M)* 90

Figure 32. *Declarative statement with no declination in a string of M tones (S3, M)* 91

Figure 33. *Declarative statement with no declination in a string of M tones (S3, M)* 91

Figure 34. *Declarative statement with no declination in a string of H tones (S2, F)* 92

Figure 35. *Declarative statement with no declination in a string of H tones (S3, M)* 92

Figure 36. *Declarative statement showing an H tone lowered after an M tone (S3, M)* 93

Figure 37. *Declarative statement showing an H tone lowered after an M tone (S3, M)* 93

Figure 38. *Duration of phrase-medial and phrase-final vowels in statements by speaker (z-scores)* 96

Figure 39. *Polar question ending in an L tone (S3, M)* 100

Figure 40. *Polar question ending in an M tone (S3, M)* 101

Figure 41. *Polar question ending in an H tone (S3, M)* 101

Figure 42. *Duration of phrase-medial and phrase-final vowels in polar questions by speaker (z-scores)* 102

Figure 43. *Duration of phrase-final vowels in statements and polar questions by speaker (z-scores)* 104

Figure 44. *Wh-question with a final rise in F0 (S3, M)* 108

Figure 45. *Duration of phrase-medial and phrase-final vowels in wh-questions by speaker (z-scores)* 109

Figure 46. *Duration of phrase-final vowels in statements, wh-questions, and polar questions by speaker* 110

Figure 47. *Polar question with a final lexical L tone and no final fall in F0 (S1, F)* 113

Figure 48. *Polar question from an interview using the final particle* kɛ^{1} *(S3, M)* 116

Figure 49. *Final particle* kɛ^{1} *used in a statement (speaker S5, female)* 116

Figure 50. *Polar question ending in the negation particle* =ga^{5}/ŋa^{5} *(S2, F)* 117

Figure 51. *Polar question with a final lexical H tone and an L% boundary tone (S2, F)* 119

Figure 52. *Polar question with a final lexical H tone and no L% boundary tone (S2, F)* 119

Figure 53. *Polar question with a final lexical L tone and a final rise in F0 (S3, M)* 120

Figure 54. *Polar question with prolongation and breathy termination (S3, M)* 121

Figure 55. *Wh-question from an interview task ending in a final raised M tone (S1, F)* 121

Figure 56. *Non-negated sentence with an L-toned subject and TAM marker (S3, M)* 123

Figure 57. *Negated sentence with L-toned subject and TAM marker (S3, M)* 124

Figure 58. *Non-negated sentence with an L-toned subject (S3, M)* 124

Figure 59. *Negated sentence with an L-toned subject (S3, M)* 125

Figure 60. *Non-negated sentence with an M-toned subject (S3, M)* 125

Figure 61. *Negated sentence with an M-toned subject (S3, M)* 126

Figure 62. *Declarative statement with* wi¹ri¹ *'call' in contrastive focus (S3, M)* 142

Figure 63. *Declarative statement with* wi¹ri¹ *'call' not focused (S3, M)* 142

Figure 64. *Declarative statement with a contrastively focused object (S1, F)* 143

Figure 65. *Declarative statement with a contrastively focused object (S1, F)* 144

Figure 66. *Declarative statement with a contrastively focused object (S1, F)* 144

Figure 67. *Declarative statement with subject in narrow focus (S2, F)* 145

Figure 68. *Declarative all-new statement with M toned subject (S2, F)* 146

Figure 69. *Declarative statement with* Fa¹tou³ *in broad focus (S3, M)* 146

Figure 70. *Declarative statement with* Fa¹tou³ *in narrow focus (S3, M)* 147

Figure 71. *Declarative statement with* Fa¹tou³ *not focused (S3, M)* 147

Figure 72. *Declarative statement with* fan⁵ta¹ ta¹ dẽ¹mẽ¹=õ¹ *in broad focus (S3, M)* 148

Figure 73. *Declarative statement with* zio³ mẽ¹nẽ¹ *in broad contrastive focus (S3, M)*..... 149

Figure 74. *Declarative statement with* zõ¹ *in contrastive focus (S2, F)* 150

Figure 75. *Declarative statement with fẽ³ in contrastive focus (S2, F)* 150

Figure 76. *Polar question with unexpected fall in F0 on the object (S1, F)* 151

Figure 77. *Declarative statement with Unexpected fall in F0 on the object when the adverb is focused (S2, F)* 151

List of Abbreviations

AM	autosegmental-metrical
DEF	definite
DEM	demonstrative
EMPH	emphatic
IMP	imperative
LOC	locative
NEG	negative
NP	noun phrase
OBL	oblique
PAST	past tense
PART	particle
PM	predicative marker
REL	relative
SG	singular
TAM	tense, aspect, and modality
TBU	tone-bearing unit
VOT	voice onset time

Transcription Conventions

Glossed examples are presented in a phonemic IPA transcription unless stated otherwise. Lexical tone is transcribed using tone numbers in the following way:

a^{1}	L	low tone
a^{3}	M	mid tone
a^{5}	H	high tone
a^{13}	LM	low-mid tone
a^{15}	LH	low-high tone
a^{31}	ML	mid-low tone
a^{51}	HL	high-low tone

Annotations of audio files use the following conventions:

Tier 1	word-by-word phonemic segmental transcription
Tier 2	lexical tone
Tier 3	intonation (breaks and intonational tones)
Tier 4	word-by-word gloss
Tier 5	free translation
L%	low phrase-final boundary tone
H%	high phrase-final boundary tone
%	final phrase boundary without associated boundary tone

Abstract

This dissertation describes the word-level and phrase-level prosodic system of Southern Bobo Madaré (Bobo), a Mande language of Burkina Faso. I examine tonal aspects of Bobo's prosodic system and provide an extensive phonetic description of the use of non-modal phonation and final lengthening to mark utterance type. The data examined include both elicitation tasks and spontaneous speech tasks. The work is conducted within the framework of autosegmental-metrical theory (Pierrehumbert 1980).

Several aspects of the word-level prosodic system are discussed. Previous work on Bobo (Morse, 1976; Le Bris & Prost, 1981; Sanou, 1993) disagree on the inventory of contour tones and the existence of word stress. I present an analysis in support of three contour tones: High-Low, Low-High, and Low-Mid. I do not find clear phonetic evidence of word stress. Phonological analysis supports the existence of stress however: The distribution of reduced vowels supports the existence of iambic prosodic feet, which is common in Mande languages. Furthermore, the distribution of tone melodies is best explained by assuming that tone melodies are assigned to the foot rather than to the word or morpheme, similar to Leben's (2001) proposal for tonal feet in Bamana.

While both word-level and phrase-level prosody are discussed, most attention is given to phrase-level prosodic phenomena. In recent years, there has been increased interest in the phrase-level prosody of African tone languages (Downing & Rialland, 2016). However, detailed descriptions of the phrase-level prosody of Mande languages still remain extremely rare. This is the first such description of a Mande language with three tone levels. Bobo makes relatively little

use of intonational tones. Declarative statements are marked only through final lengthening and in some cases non-modal vowel phonation. Polar questions show some characteristics of the areal "lax question prosody" described by Rialland (2009): L% boundary tone, which is concatenated onto the string of lexical tones, extreme lengthening of the phrase-final segment (always a vowel in Bobo), and breathy utterance termination. This L% boundary tone is the only clear case of an intonational tone in Bobo. *Wh*-questions can (but typically do not) have an L% boundary tone and have a lesser degree of phrase-final lengthening than polar questions. Negated statements do not have special prosodic characteristics. The phrase-level prosodic hierarchy of Bobo is relatively flat, consisting of only the intonational phrase. In addition to investigating the prosodic marking of utterance type, I present an investigation into focus marking in Bobo. I examine the responses to *wh*-questions and corrections, two contexts in which focus-marking is typically found cross-linguistically. I find no evidence of morphosyntactic or prosodic focus marking in these contexts. Bobo is therefore an additional example of an African tone language without obligatory focus marking in these contexts.

The relevance of these results to our current understanding of prosodic typology is discussed throughout.

Chapter 1

Introduction

1.1 Overview

The purpose of this dissertation is to describe the prosodic system of Southern Bobo Madaré, a Mande language of Burkina Faso. Its primary focus is the phonetic and phonological properties of its prosody: what the phonological components of the prosodic system are, how they are expressed phonetically, and how they are used to encode hierarchy and prominence. Chapter 3 describes the word-level prosody of Bobo, Chapter 4 describes the phrase-level prosody (with a focus on phrase boundaries), and Chapter 5 investigates the relationship between prosody and information structure—namely, whether focus is expressed prosodically. One typologically unusual finding presented here is that Southern Bobo Madaré does not appear to use either prosodic or morphosyntactic means to mark focus, meaning that focus is not marked.

In recent years, there has been an increased interest in the prosodic systems of languages that have not yet been investigated in detail (e.g. Jun, 2005a; Jun, 2014; Downing & Rialland, 2016a). Traditionally, phrase-level prosody has received much less attention than word-level prosody, especially in tone languages; in fact, it was once thought that lexical tone would preclude or discourage many uses of intonation (e.g. Yip, 2002). However, it is now established that the use of intonation in tone languages is widespread. The properties of intonation in tone languages have also turned out to be surprisingly variable, meaning that data from a wide variety of genetically and typologically diverse tone languages is necessary in order to establish useful typologies of

prosodic structure. There has been comparatively little detailed work on the phrase-level prosody of tonal West African languages, and almost none on languages belonging to the Mande family. This dissertation contributes a detailed description of the phrase-level prosody of a language from a region and a language family that has been underrepresented in the prosody literature.

I adopt the framework of *autosegmental-metrical (AM)* theory, which was initially developed using data from English (Pierrehumbert, 1980) but has since been applied to typologically diverse languages, including lexical tone languages. It is by far the most widespread framework for the analysis of prosody. In AM theory, surface intonational contours are analyzed as the result of an underlying sequence of discrete level tones. This is a phonological approach to prosody; under this framework, prosody is analyzed in terms of phonological elements and the phonological rules governing their expression. The primary goal of this work, therefore, is to investigate and describe the phonological properties of prosody that exist in Southern Bobo Madaré.

1.2 Southern Bobo Madaré

Southern Bobo Madaré (henceforth 'Bobo') is a Mande language spoken in southwestern Burkina Faso by approximately 181,000 people, most of whom are ethnically Bobo (Simons & Fennig, 2018). It is closely related to Bobo Konabéré, also called Northern Bobo Madaré, and Zara, also called Bobo-Dioula. The Bobo call themselves *Bɔ^{1}bɔ^{5}* (pl. *Bɔ^{1}bɔ^{5}i^{1}*), and their language *Bɔ^{1}bɔ^{5}-da^{5}*. They are also known as the Bobo Fing ('Black Bobo'), but this term has fallen out of use in favor of Bobo Madaré.[1] Both Southern and Northern Bobo Madaré are included under this term. There are considerable differences between the dialects of Southern Bobo Madaré.

[1] The Bobo-Oule ('Red Bobo') and Bobo-Gbe ('White Bobo') are speakers of Bwamu (Gur) languages.

Tiendrébéogo (1998) lists at least four varieties: Syabéré (also called Sya-dá), spoken in and around Bobo-Dioulasso; Voré, spoken to the west of Bobo-Dioulasso; Beng, to the east; and Sogokiré, to the north. This project focuses on Syabéré.

Southwestern Burkina Faso is linguistically diverse and monolingualism is uncommon. Most Bobo also speak Jula[2], another Mande language that is used as a *lingua franca* in the region (Tiendrébéogo, 1998; Showalter, 2008). Knowledge of Jula is increasing, with younger speakers, both male and female, having more competence than older speakers (Showalter, 2008). All educated Burkinabé are also proficient in French, which is the official language of Burkina Faso. Knowledge of additional languages (e.g. of close neighbors) is also common. Thus, the "typical" Bobo speaker is at least bilingual and frequently interacts with people in a language other than Bobo. This is especially true of Bobo speakers who live in towns and cities. In Bobo-Dioulasso, the influence of Jula on Bobo is noticeable as there are many Jula loanwords. Bobo is not frequently written. Literacy in Bobo is rare; it is not used in schools or by the government, and it does not have a widespread written tradition. There are some Christian religious works translated into Syabéré, but these do not appear to be widely known. The majority of the Bobo are either Muslim or follow traditional religion (see Tiendrébéogo, 1998).

There are three prior descriptive works that contain original data, all of which also concern the Syabéré dialect. The standard reference is a grammatical sketch and lexicon by Le Bris and Prost (1981). The lexicon was originally compiled by Le Bris, a priest who lived with the Bobo for twenty years, and it is therefore extensive. Prost, an experienced missionary linguist, later added the grammatical sketch and tone transcriptions. Prost describes the grammatical sketch as

[2] Written *Dioula* in Burkina Faso, following the French spelling; this is the origin of the *Dioula* in *Bobo-Dioulasso* and in *Bobo-Dioula*.

tentative, but it is nonetheless the most widely cited source on Bobo within the Mande literature. Its information on the phonetics and phonology is limited to the phonemic inventory and some of the grammatical uses of tone. For more detailed information on phonetics, he references Morse (1976), a doctoral dissertation on the phonetics, phonology, and some of the morphology of Bobo. This is the earliest available work that contains a significant amount of information on Bobo. It is in many ways an impressive piece of work, but because it has very broad scope, any given topic cannot be treated with much detail. Morse was also working without the benefit of modern recording equipment, which limited the types of details she could include. The third work is Sanou (1993), another doctoral dissertation. This dissertation is the author's original description of the grammar of Bobo, which does not appear to draw from either Morse or Le Bris and Prost.[3] There are some important differences between these three sources. All three disagree about the inventory of lexical tones, and Sanou disagrees with Morse and Le Bris and Prost about the existence of productive phonemic vowel length. Morse is the only author to address the phrase-level prosody of Bobo; her description is summarized in Chapter 4. Since there are disagreements between the prior descriptive works on Bobo and because none of these works are particularly recent, the following basic description of the language is based on my own research.

Bobo is a mostly isolating language with strict S-PM-O-V-X word order. PM stands for a special class of TAM (tense, aspect, modality) particles frequently called *predicative markers* in the Mande literature. X stands for constituents that are not core arguments of the verb, and includes postpositional phrases and adverbs of time, place, and manner. A simple transitive sentence illustrating the canonical word order is given below in example 1.

---

[3] This work is less widely referenced, probably because it is less accessible. This dissertation was almost complete before I was able to acquire a copy.

(1) *a*[1] *ma*[3]*ne*[3] *kpĩ*[51] *mɛ*[1]*nɛ*[1] *lu*[51] *ma*[1].
3.SG PAST millet.beer drink courtyard LOC
"He was drinking millet beer in the courtyard."

In *wh*-questions, the *wh*-object is often fronted, as in example 2. This is optional; the *wh*-object can also be left *in-situ*. Another case in which optional fronting can occur is in relative clauses, as is shown in example 3. Here the relativized object *ŋwĩ*[15] *sõ*[15] *be*[3] 'the man that…' is fronted.

(2) *ŋwo*[5]*nõ*[5] *fa*[1]*tu*[3] *gũ*[1]*?*
what Fatou gain
"What happened to Fatou?"

(3) *ŋwĩ*[15] *sõ*[15] *be*[15] *ma*[3] *za*[1], *a*[1] *be*[3] *pu*[1].
REL man DEM 1.SG see 3.SG 2.SG younger.sibling
"The man that I saw is your younger brother."

There is very limited segmental morphology. Most grammatical distinctions are encoded by word order or grammatical function words; some are additionally encoded through length and tone alternations. A notable exception is pluralization, which is marked through vowel ablaut or the use of plural suffixes, depending on the class of the noun.[4] Ablautive plurals preserve the tone melody of the singular. Most words are mono- or bi-syllabic, with longer words almost always being either compounds or loans. Most syllables are CV. Syllable codas are extremely rare. Onsetless syllables are also rare, although the third person singular pronoun *a*[1] is frequent in speech. Some words contain vowel sequences, which fall into two categories: sequences beginning with *i* and sequences that are the result of suffixation or encliticization (e.g. *kãẽ*[13] 'foot' from *kã*[13] 'foot'). Morse (1976) claims the vowels in a sequence belong to separate syllables, but she does not provide details. I believe most sequences beginning with *i* are phonological diphthongs, for the

[4] Noun class is not explicitly marked and does not trigger agreement patterns. It manifests only in the form of the plural.

following reasons:

- The *i* in these sequences is reduced in duration, and sometimes coalesces completely with preceding alveolar or velar consonants (e.g. *kia*^{3} > [t͡ʃa^{3}] ‘search’; *sio*^{1} > [ʃo^{1}] ‘horse’;
- In these reduced cases, *i* does not carry its own tone;
- This is not true of sequences formed by suffixation or encliticization, even when the first vowel in the sequence is an *i* (e.g. *bi*^{1}*bi*^{5} ‘baby’> *bi*^{1}*bi*^{5}*e*^{1} [bi^{1}bi^{5}e^{1}] ‘babies’, where the *i* is not reduced in duration and retains its own tone).

It is still not clear what the timing properties of diphthongal *i* are, other than that they are different from non-diphthongal *i*. If diphthongal *i* is moraic, then syllables that contain it are one of only two types of heavy syllable. The other type is syllables containing long vowels, which are extremely marginal. Long vowels are limited to some numerals (e.g. *naa*^{1} ‘four’; c.f. Jula *nàani*), some adverbials (e.g. *fee*^{5} *ga*^{5} ‘never’), some stative verb forms (e.g. *ta*^{1}*ŋaa*^{131} ‘is seated’), and a few nouns (e.g. *laa*^{1}*re*^{1} ‘butter’, a Fula loan via Jula). Sanou (1993) claims a more productive vowel length distinction but I have not found evidence of this. The only minimal pairs for vowel length that I have been able to verify are *naa*^{1} ‘four’ and *na*^{1}, a postposition. There are no minimal pairs for vowel length within the same syntactic category. Sanou’s analysis of vowel length appears to be partially motivated by the theoretical stance that contour tones should occupy two morae.

Bobo has both lexical and grammatical tone. There are three tone levels: high (H), medium (M), and low (L). These tone levels appear to be independent of vowel length and phonation. A minimal triplet illustrating the contrast between H, M, and L is given in example 4:

(4) *ta*^{1}*ba*^{1} ‘tobacco’
ta^{1}*ba*^{3} ‘support’
ta^{1}*ba*^{5} ‘knife’

Bobo also has four contour tones, which are decomposable into sequences of the level tones: low-high (LH), high-low (HL), low-mid (LM), and mid-low (ML). The prior sources agree on the level tone inventory, but not on the inventory of contour tones. Sanou (1993) proposes only two (LH and LM), Morse (1976) proposes three (LH, LM, and ML), and Le Bris and Prost (1981) propose the four listed above. More discussion of their analyses and my reasons for proposing four contour tones is provided in §3.2. Grammatical tone is extensive but not yet well described. Morse (1976) focuses on the phonetics and phonology of tone, rather than its use within the grammar. Le Bris and Prost and Sanou include some information on grammatical tone but this information is limited and is not consistent between these two sources or consistent with my own data. A particularly troublesome issue is the distinction between transitive verbs in the simple present and the past perfective: Le Bris and Prost claim that there is a tonal distinction, but these conjugations are generally homophonous for the speakers who participated in this project. Since verbs frequently appear at the end of phrases this issue affects the interpretation of intonation at phrase boundaries. Grammatical tone alternations are also common elsewhere in the grammar and include at the least:

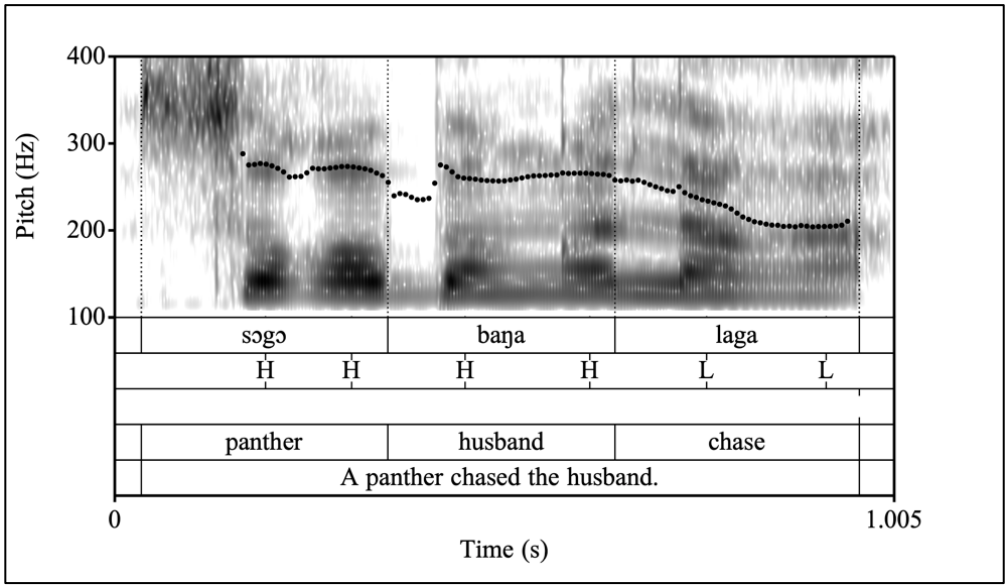
- Other verb conjugations (e.g. *ta^{1}ŋa^{1}* 'sit' > *ta^{1}ŋaa^{131}* 'is seated')
- Compounding (e.g. *ba^{1}lo^{5}* 'iron' + *sio*^{1} 'horse' > *bal^{1}-sio^{5}* 'bicycle')
- Derivation (e.g. *pe^{3}ne^{3}* 'be red', verb > *pe^{5}ne^{5}* 'red', adjective)
- Possessive constructions (e.g. *ba^{5}ŋa^{5}* 'husband' > *ne^{3} ba^{1}ŋa^{5}* 'my husband')
- Some complex verb constructions (e.g. *ma^{3} ja^{3}* 'I go' > *ma^{3} mɔ^{3}gɔ^{3} ja^{1}* 'I cry and go')

The tonal morphology is typically replacive. Instead of affixing a grammatical tone, the entire tone melody of the word is replaced, as can be seen in the examples above. Le Bris and Prost (1981) claim the existence of a few tonal suffixes, but I have not been able to verify these. Tone

sandhi is minimal or nonexistent. Within NPs it is generally precluded by replacive grammatical tone processes; the use of adjectives or possessive constructions are the primary means by which an NP can be made complex, but these involve replacement of tone melodies. There does not appear to be tone sandhi across larger syntactic boundaries, either, such as between a subject and an object or between an object and a verb. Neither does there appear to be an obligatory contour principle, as is illustrated in Figure 1. In this example, both subject and object contain successive H tones that surface as H; there is no indication of an obligatory contour tone principle affecting their realization. Tone is the most salient and productive part of the word-level prosodic system. The tone system is discussed in detail in Chapter 3**.**

Figure 1

Successive H tones in a phrase (S1, F)



1.3 Prior Work on Related Languages

Mande is a large language family. It has proposed time depth comparable to Indo-European (Vydrin, 2009) and its languages are spoken across a wide region in West Africa, including Senegal, Sierra Leone, Mali, Burkina Faso, Côte d'Ivoire, and Liberia, with scattered pockets as

far east as Nigeria (Glottolog). The division of the family into two primary branches, Western Mande and Eastern Mande, is uncontroversial, as are most low-level groupings. Mid-level groupings are more problematic (Vydrin, 2009). Bobo is typically placed somewhere in the Western branch.[5] Historically, Mande has been placed within the Niger-Congo family but due to a lack of convincing evidence of the relationship it is now frequently treated as its own primary language family (Dimmendaal, 2011; Glottolog). If Mande is Niger-Congo it is an early, divergent branch (see Dwyer, 1998 for one such proposal).

Although Mande is a large language family with a time depth of several thousand years, Mande languages do share some notable typological characteristics. The S-PM-O-V-X word order found in Bobo is typical of the family. Most Mande languages also have simple syllable structure, although some have undergone reduction of the initial vowel in bisyllabic words, such as Bamana (Green, 2010; Green et al., 2012) and Seenku (McPherson, 2020). All or almost all Mande languages are tonal (Vydrin, 2002). Apart from this fact, it is difficult to make generalizations about their prosody as a whole. Most Mande languages, such as Bamana, have two tone levels (Dwyer, 2006; Dumestre, 2003; Green, 2010), but some, such as Bobo and Seenku (McPherson, 2017; McPherson, 2020) have three or more. Productive vowel length distinctions are common, occurring in Bamana (and the closely related Jula), but as previously discussed, not in Bobo. Since the word-level prosody of Mande languages is so variable, there is no reason to expect that the phrase-level prosody would not be. There is very little work on the phrase-level prosody, however.

[5] There is little consensus on where in the Western branch Bobo belongs. Kastenholz (1997) places it within the Northwestern subgroup of Western Mande, as a sister to Soninke-Bozo. Kastenholz's classification has been adopted by Glottolog and Ethnologue. It has also been suggested that Bobo forms its own primary branch (Welmers 1971). Bobo shares relatively little core vocabulary with other Mande languages, making it difficult to reliably place it within the family.

Four noteworthy exceptions are Mountford (1983), an investigation of the declarative sentence intonation on Bamana, Vydrina (2017), a grammatical description of Kakabe that includes a detailed chapter on intonation, Konoshenko and Sigova (2018), a study of *yes-no* question intonation in Guinean Kpelle, and Vydrin and Diané (2016), a study of question intonation in Guinean Maninka.

The majority of Mountford (1983) is analysis of declination over sentences containing different tone patterns. Although the data collected was limited due to the methods available at the time, he did find consistent, steady declination, suggesting that Bamana has phonetic downdrift rather than downstep—that is, Bamana sentences show a gradual decrease in F0 over the entire course of a sentence, rather than an abrupt lowering of F0 that occurs in particular phonological environments. Mountford did not investigate other intonational phenomena in Bamana. The description of Kakabe in Vydrina (2017) is much more comprehensive. In it, Vydrina presents a phonological analysis of Kakabe intonation in a Register Tier Theory (Hyman, 1993) framework. According to her analysis, Kakabe has several phrase-final intonational morphemes: (↑)H%, HL%, L%, and ↑%, where <↑> represents register raising. In Kakabe, the TBU is the mora; the intonational morphemes that contain boundary tone also have an associated mora, which causes phrase-final lengthening. The intonational tones have the same phonetic and phonological properties as lexical tone, and they attach to the tonal tier after the string of lexical tones. They are distinguished from lexical tone only by their association with the phrase boundary, rather than any other difference in their other phonetic or phonological properties. She does not investigate whether there are distinct prosodic phrase types, or the information structural category of *focus.* However, she does find that discourse prominence is expressed with register raising.

Konoshenko and Sigova (2018) were interested in the interaction of boundary tone and

intonation in Guinean Kpelle. In this language, yes-no questions are marked by a final rise, which they analyze as an H% boundary tone. The characteristics of this tone differ from the H% in Kakabe. Rather than being inserted after the string of lexical tones, it is superposed on the final word, resulting in a slight rise in the penultimate vowel followed by a sharper rise in the final vowel. Its precise expression is dependent on the lexical tone; HL and H tone melodies are neutralized by the rise, but L and L° (an L followed by an underlying floating H) are not. They do not have an explanation for why one contrast should be neutralized and the other should not. The preserved contrast between L and L° is even more puzzling because it is L that is realized with a higher rise than L°, opposite of what one might expect. Another difference between H% in Kakabe and Guinean Kpelle is that in Guinean Kpelle, H% does not cause lengthening on the vowel; there is no evidence of an associated mora. The final vowel of yes-no questions marked by H% is the same length as in declaratives, which have no boundary tone. Vydrin and Diané (2016) find that yes-no questions are also marked by a final rise in F0. Though they do not interpret this as the result of H%, the F0 contour is similar to the contour Vydrina (2017) describes for Kakabe questions in which there is an H%. In addition, the F0 of the first lexical H tone in in a question in Guinean Maninka is raised.

All four of these languages belong to the Central-Southwestern branch of Mande. Bamana, Maninka, and Kakabe are even more closely related, belonging to the Manding cluster within the Central-Southwestern branch. Most previous work on phrase-level prosody in Mande therefore is concentrated on a small slice of the overall family. In addition, all four of these languages have only two tone levels in comparison to Bobo's three. There is no reason to expect that phrase-level prosody in Bobo will have the same (or similar) characteristics.

1.4 Prosody

The term *prosody* is used here to denote the linguistically structured suprasegmental phonetic and phonological properties of a language. Under this definition, *prosody* includes both word-level phenomena and phrase-level *intonational phonology* in the sense of Pierrehumbert (1980) and Ladd (2008). At the word level, it includes phenomena such as word stress, vowel length, and tone. At the phrase level, it includes the suprasegmental properties that encode rhythm, prominence, and hierarchical structure. Segmental phenomena can clearly be related to the prosodic structure of an utterance too, in the sense that they are conditioned by prosodic structure. Strengthening of segments in prosodically prominent or domain-initial positions has been found in a variety of languages (Keating, 2006) and may be universal. Other segmental phenomena that are conditioned by prosodic structure are language specific. For example, in Tariana (an Arawakan language), phrase-final vowels can undergo centralization and lowering under certain discourse conditions (Aikhenvald, 2003). In Oneida, phrase-final words can have segments added, devoiced, or otherwise changed (Abbott, 2006). Although these are not suprasegmental phenomena, they can encode (or be triggered by) the same prosodic structures as suprasegmental phenomena. This is relevant for Bobo, as Morse (1976) claims that major intonational phrases in Bobo are marked by a phrase-final glottal stop[6], which would be excluded by strict adherence to a definition that excludes segmental phenomena. Ladd (2008) raises the question of whether segmental morphemic markers of pragmatic meaning, such as question or focus particles, should also be considered part of intonation. If the definition of intonation is based on its communicative function, then perhaps they should be. If the definition based on phonological content, then perhaps not—but many

[6] They do not, as a general rule. See Chapter 2.

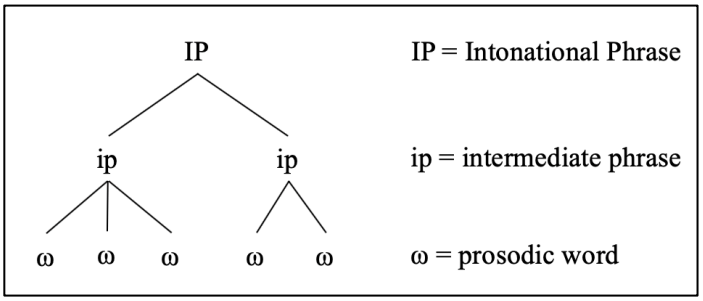
common definitions of intonation include both functional and phonological criteria, as Ladd's does (Ladd, 2008; pp. 4-6). Definitions of *intonation*, like *prosody,* also vary. I will use *intonation* to refer specifically to the pitch phenomena that are associated with the phrase, rather than the word. For an in-depth discussion of the difficulty of finding a universal definition of prosody or intonation, see Shattuck-Hufnagel and Turk (1996) and Ladd (2008, pp. 4-6).

It is generally accepted that speech is organized into phonological constituents that exist within a hierarchical structure (e.g. Nespor & Vogel, 1986; Pierrehumbert, 1980; Wagner, 2005; Selkirk, 2011; Ladd, 2008). The proposed levels of the hierarchy vary depending on the language, the theoretical perspective, and the individual analysis. In general, there is broad agreement about the constituents at the lowest levels of the hierarchy, in the domain of word prosody. There is also broad agreement on the existence of something the size of the intonational phrase, despite differences in terminology and in what are considered its defining properties. However, there is considerably less agreement about the levels of the hierarchy above the word, in the domain of prosodic phrases. Each level of the hierarchy corresponds to a phrase type. For example, the most commonly accepted AM analysis for American English includes an intermediate phrase and an intonational phrase, and it is this analysis that the ToBI transcription system encodes (Beckman et al., 2005). The intonational phrase (IP) dominates the intermediate phrase (ip), which dominates the prosodic word (ω). The prosodic hierarchy of a hypothetical English sentence is shown in Figure 2.

Identification of a prosodic constituent is based on it possessing a unique phonological property, such as the presence of a boundary tone for example, or it being the sole domain of a phonological process, such as tone sandhi or downstep—or both (e.g. Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Selkirk & Tateishi, 1991; Peng et al., 2005, and many others).

Figure 2

Phrase-level prosodic structure of English



Different levels of the hierarchy can also be encoded using finer-grained levels of detail, such as different degrees of phrase-final lengthening and different degrees of articulatory strengthening at boundaries (Fougeron & Keating, 1997). However, the relationship of these gradient differences to underlying categorical differences is still an ongoing subject of research (Keating, 2006; Krivokapić & Byrd, 2012; Wagner, 2005). The number of different analyses for English points to the difficulty of finding unequivocal evidence of smaller phrase types. The prosodic phenomena that encode phrasing can be affected by many factors, which often makes data variable and difficult to interpret. An overview of some different proposals for English is available in Shattuck-Hufnagel and Turk (1996). Chapter 4 investigates what types of prosodic phrases can be identified in Bobo and the phonological elements that encode them.

In addition to prosodic phrasing, this project also investigates prosodic prominence. *Prosodic prominence* is a relationship between elements within a prosodic constituent. At the word level, prominence refers to the pattern of strong and weak syllables, with the stronger syllables being prominent in some way—e.g. by having the property of lexical stress. At the phrase level, a word can be made more prominent than other words, frequently in order to express information structural properties like focus. Definitions of what precisely constitutes prominence vary. Wagner et al. (2015) note that most have the following general form: "<We> say that a <linguistic entity>

is prosodically prominent when it <stands out> from <its environment> (by virtue of its prosodic characteristics>).” (p. 3) This definition is problematically vague, but a more specific, universal definition is difficult because the ways in which prominence is expressed are language specific. Wagner et al. suggest that the researcher should tailor their definition of prominence to the phenomena that they are investigating. However, since it is not known what phenomena related to prominence exist in Bobo, this is not possible at this time. Two specific functions of prominence are covered in this dissertation: Chapter 3 investigates lexical stress (among other aspects of word-level prosody), and Chapter 5 investigates whether prosodic prominence encodes focus.

1.5 Autosegmental-Metrical Theory

A central assumption of any model of prosodic grammar is that prosodic properties are linguistically structured. There are significant differences of opinion about how to model this underlying structure, however. One major theoretical split is between models that posit that prosody can be derived from the syntactic structure (e.g. Wagner, 2005), and those that posit a separate phonological prosodic structure (e.g. Selkirk, 2011; Pierrehumbert, 1980). Another split is between models that treat the underlying representation of surface F0 contours as movements or trajectories (e.g. Fujisaki & Hirose, 1984), and those that treat the underlying representation as discrete phonological entities, i.e. tones. AM theory belongs to the latter category in both cases. That is, it is a theory of phonology in which intonational contours can be decomposed into a sequence of intonational tones. These intonational tones are not specified for each syllable but are instead assigned to important landmarks within the prosodic phrase, leaving the rest of the syllables unspecified for intonational tone. The surface F0 contour is the result of interpolation between the tones (e.g. Pierrehumbert, 1980). A major contribution of AM theory is that it can account for the structured variation in surface F0 contours between sentences of varying lengths and prominence

patterns (Arvaniti & Ladd, 2009; Arvaniti, 2011).

AM theory is a development of autosegmental phonology (Goldsmith, 1976). Autosegmental phonology posits that lexical tones, instead of being an intrinsic part of the syllable, are instead discrete phonological entities that exist on their own tier within a hierarchical phonological structure. Tone is linked to tone-bearing units (TBUs), usually a syllable or a mora. Phonological processes that affect a lexical tone's association with a TBU can explain phenomena such as tone spreading (linking of a single tone to multiple TBUs) and tone crowding (linking of multiple tones to a single TBU). Contour tones are the result of a sequence of underlying level tones, such as an HL tone being the result of a sequence of an H and L tone. Beginning with Pierrehumbert (1980), this approach has been widely adopted for the analysis of intonational contours as well. Under AM theory, intonational contours are the result of sequences discrete underlying tones as well, typically H and L. The three types of intonational tones are: pitch accents, which mark prominence; phrase accents, which mark the end of an intermediate phrase; and boundary tone, which marks end of an intonational phrase (Beckman & Pierrehumbert, 1986; Beckman et al., 2005). The intonational contour of an utterance is primarily determined by the sequence of intonational tones and how they are aligned with their TBUs. Under AM theory, lexical tone and intonational tone must be closely related in any language that has both; they are the same type of phonological element and both are aligned with the TBUs within the segmental string. This view provides a clear framework for analyzing the interaction between lexical and intonational tones.

The prosodic hierarchy in AM theory consists of constituents ranging in size from the mora or syllable to the major intonational phrase (IP). In an AM analysis, the intonational phrase is universal; all languages have prosodic phrasing, and the intonational phrase is simply the largest

prosodic phrase. Cross-linguistically, the IP is commonly marked by a final intonational contour, lengthening, and a pause. The levels of the hierarchy between the word and the intonational phrase vary considerably across languages. For example, Mandarin is claimed to have the tone sandhi group, identified as the domain of tone sandhi processes (Peng et al., 2005); Japanese is claimed to have an intermediate phrase (ip), which is the domain of downstep (Beckman & Pierrehumbert, 1986); Korean is claimed to have the accentual phrase (AP), identified as the domain of a regular accentual melody (Jun, 1998); and Cantonese is claimed to have no prosodic phrase smaller than the IP at all (Wong et al., 2005). The inventory of intonational tones is also variable between languages; notably, many lexical tone languages do not use pitch accent to mark prominence (e.g. Peng et al., 2005 for Mandarin; Wong et al., 2005 for Cantonese; Jannedy, 2007 for Vietnamese; Downing & Pompino-Marshall, 2013 for Chichewa; Rialland & Aborobongui, 2016 for Embosi; and many others.)

AM theory is the most widely used framework in the study of prosody, including most recent work on prosodic typology (e.g. Jun, 2005a; Jun, 2014; Downing & Rialland, 2016a). It is well-developed and widely understood. It provides a common framework for comparison of different intonational systems, while still being flexible enough to handle variation–much in the way autosegmental theory can handle variation between lexical tone systems. This dissertation focuses first on accurately describing the phonetic properties of the data and then on phonological interpretation.

1.6 Tone and Intonation

Since the primary correlate of both lexical tone and intonational tone is F0, it is possible for intonation to obscure lexical tone when they co-occur—or vice versa. It has been suggested that lexical tone languages make less use of intonation overall, and that the reason for this is the

need to preserve lexical contrast (e.g. Jun, 2005a; Ladd, 2008, p. 36). Lexical tone languages can use other strategies to mark the same functions that intonation does in non-lexical tone languages, such as using particles to mark utterance type and focus (Yip, 2002, pp. 271-272). Mambila, a Bantoid language of Nigeria and Cameroon, is an extreme example. Mambila has four tone levels; fine pitch distinctions already carry a heavy functional load. It also has very limited intonation. There is no evidence of boundary tone and utterance type and focus are encoded morphosyntactically (Connell, 2016). If lexical contrast preservation is indeed important enough to influence the intonation of a language, then in addition to finding less use of intonation overall in lexical tone languages, we should also expect that what intonation exists will often be realized in a way that does not obscure lexical tones. Mandarin seems to fit both of these predictions: It makes frequent use of grammatical particles (Yip, 2002, pp. 271), but intonation is not absent. According to Peng et al. (2005), instead of using pitch accent to mark narrow focus, Mandarin uses pitch range expansion—which, unlike pitch accent, does not obscure the lexical tones of the focused word. Additionally, they note that they only observed boundary tones on phrases ending in lexically toneless grammatical particles; this also does not obscure lexical tone, because there is none to obscure. They did not find any examples of intonational tones co-occurring with lexical tones.

Lexical tones and intonational tones do co-occur in some languages, but there is not much research into their interaction. Most existing studies on the phrase-level prosodies of tone languages is focused on establishing what the basic phonological components are, rather than providing detailed accounts of how intonation realizes in different lexical tone contexts. In the research that does exist, different outcomes have been observed. In some cases, phonetic realization of the boundary tone is adapted so that it does not obscure the boundary tone. This

occurs in Kammu, an Austroasiatic language of Laos. Northern and Western Kammu have two level tones, H and L, and contour tones that are composed of these level tones. Karlsson, House & Svantesson (2012) found that focal accent, which has an LH contour, is reduced or neutralized when it would otherwise obscure an HL lexical tone. They also investigated the realization of boundary tone. Kammu prosodic phrases are marked by a final H% boundary tone, and it is also reduced or neutralized when it co-occurs with a final HL lexical tone. They found no cases where the lexical tone is obscured by boundary tone. Rice (1987) reports that Slave, an Athabaskan language, also preserves lexical tone over boundary tone. Slave has two tones, H and L, with H being the marked tone and L being the default or unmarked tone. The H tone of the stem will spread to the syllable to its left under certain morphosyntactic conditions, causing it to be linked to two syllables. Phrase-final syllables are typically neutralized to L due to the addition of a phrase-final L% boundary tone. However, the H tone of a final syllable is only neutralized when the H tone spreading rule has applied—meaning that the underlying tone pattern of the stem is still recoverable, due the H still being present on the preceding syllable. When H tone spreading has not applied, the final H tone is not neutralized. The preservation of lexical tone over intonation is not universal, however. Ma, Ciocca and Whitehill (2006) investigated whether question intonation in Cantonese could obscure the identity of lexical tone. Unlike Mandarin, Cantonese has no toneless syllables, so intonational tones always co-occur with lexical tone. Polar question intonation involves both global F0 raising and a rising phrase-final intonational contour, regardless of the lexical tone of the final syllable. They found that Cantonese speakers are more likely to misidentify the final lexical tone in polar questions. This shows that lexical tone is not *always* preserved over intonation—and that at least in some cases, intonation can obscure lexical contrasts.

Hyman and Monaka (2008) propose three different strategies into which the interactions

between lexical and intonational tones can be divided. Mambila's minimal intonational system is an example of *avoidance*: the use of intonational tone is minimized and other methods of marking sentence-level meaning, such as particles, are used. Mandarin's use of pitch range expansion rather than pitch accent is another example of avoidance. Mandarin is also an example of another strategy, *accommodation*, in which intonational tones exist but do not co-occur with lexical tones—possible when a language has lexically toneless syllables, as Mandarin does. In both of these strategies, lexical tone contrasts are preserved. This is commonly reported in lexical tone languages. The third strategy is *submission,* in which intonational tones override lexical tone contrasts. Cantonese polar question intonation is an example of this strategy.

In some languages, the distinction between lexical tone and intonational tone is blurry. The Bantu languages typically have two lexical tone levels, L and H (Hyman, 1999). Many Bantu languages have a phenomenon of "penultimate high tone", in which the penultimate syllable of a phrase receives an H lexical tone. This cannot be interpreted as a boundary tone, since it does not occur at the boundary, but instead occurs on the syllable before it. Instead, it is clearly a case of lexical tone being sensitive to the position within a phrase. Boundary tone can also participate phonological linking and de-linking rules with lexical tone, as it does in lu-Haya and ci-Cewa (Hyman, 1999). In these languages, boundary tone is distinguished as "intonational" only because it is introduced by the phrase boundary, rather than because of its phonetic or phonological characteristics. As we have already seen, this occurs outside of Bantu languages as well; in Kakabe, the boundary tones also behave like lexical tones (Vydrina, 2017). These languages naturally lend themselves the most easily to an analysis using AM theory, as these intonational phenomena are clearly tonal in nature. Languages like Mandarin pose more of a challenge. Mandarin's lexical tone inventory consists of four tones, including one that is frequently accompanied by creaky voice

(3rd tone, as in *ma*^{214} 'horse'). The lexical tones are subject to complex rules of tone sandhi, but the intonational tones do not participate in it. For example, syllables that carry a phrase-final L% boundary tone do not trigger the same tone changes in a previous syllable as if they were carrying a lexical L. This difference in behavior between lexical tones and intonational tones requires an explanation; it does not inevitably follow from the treatment of both as tones. An AM analysis does not preclude an explanation, but does not come packaged with one, either.

Some question whether an AM analysis is appropriate for all languages, citing languages like Mandarin as examples. Michaud (2008; 2015) proposes that while intonation is universal, intonational tones are not. He divides languages into two typological categories: Some languages, like lu-Haya, ci-Cewa, and Kakabe, have intonational phenomena that have demonstrably identical formal phonological characteristics as lexical tone, and these can be said to have intonational tones. All other languages do not. He provides Vietnamese as an example of the latter category. Vietnamese has complex tones that are difficult to decompose into sequences of level tones such as L, M, or H. Some speakers use a final rise in pitch to mark questions, which has been interpreted as a possible H% boundary tone (Brunelle et al., 2012). However, there is no clear lexical counterpart of this boundary tone. The interpretation of the final rules as an H% is due to the assumptions of the AM framework and not its similarity to other tonal phenomena in the language. Although it is not theoretically *necessary* that lexical tone and intonational tones be identical, differences in their behavior raises important questions: Are these indeed the same phenomena? What explains their differences? One consideration is the extent to which intonation is grammaticalized for the expression of specific functions. Vietnamese is again a case study here. Brunelle et al. (2012) found considerable inter-speaker variation in intonational contours used for the expression of certain meanings, including for the basic contrast between interrogative and

declarative sentences. Their conclusion is that intonational contours are not fully grammaticalized in Vietnamese. Vietnamese, like many other tonal languages, can use sentence particles for these functions, although it does not have to. Jannedy (2007) found similar variation for the expression of prominence.

The flexibility of an AM analysis is both one of the strengths and one of the weaknesses of the theory. It is assumed that the phonetic expression of intonational tones can vary between languages and contexts, just like the phonetic expression of any other phonological element. For example, the prototypical effect of an H% or L% is a change in the F0 trajectory, but this is not the only outcome. Downing and Rialland (2016b) propose that in a lexical tone language, a boundary tone can also result also result in raising or lowering of the F0 over the entire phrase-final syllable, instead of a change in trajectory. This flexibility means that multiple interpretations are sometimes possible and that what one researcher interprets as an intonational tone may be interpreted as a non-categorical or non-phonological phenomenon by another.

1.7 Research Questions

This dissertation is broad in scope. Its intention is to provide a basic description of the prosodic structure of Bobo, with a focus on its phonetic and phonological properties. The research questions that guided the design of the project are:

- What are the levels of the prosodic hierarchy?
- What are the phonological elements that are associated with each level of the hierarchy, and what are their characteristics?
- To what extent is prosody used to encode meanings such as utterance type and focus, and when it is used, what form does it take?

Chapter 2

Methods

2.1 Overview

This project is based on data collected in Bobo-Dioulasso, Burkina Faso between the years of 2014 and 2017. There are two general types of data: Recordings of elicited sentences and recordings of spontaneous speech. Both types of data have benefits and limitations. Elicited sentences allow the researcher to target specific prosodic phenomena by controlling for other factors that can affect the prosody of an utterance. This is especially important when investigating intonation in a lexical tone language because it can otherwise be difficult or impossible to separate the effects of lexical tone and intonation on F0. The most common method of eliciting sentences is through tasks in which the speaker reads a prompt since this allows the most control over the sentences produced. Despite concerns that read sentences are "unnatural" and uninformative, they have been shown to be consistent with more "natural" speech produced outside of an elicitation context (Lickley et al., 2012). Elicited sentences will obey the grammatical (including phonological) structure of the language unless there are more specific issues with the methodology. Nevertheless, it is true that elicited sentences can only provide a narrow view of the prosodic system. By design, the pragmatic contexts and social contexts that elicited sentences cover are limited. Spontaneous speech data, on the other hand, provides a much broader view. Its limitation is that it does not enable controlled comparisons. This limitation makes it difficult to interpret, especially before a basic phonological analysis has been conducted. By using both elicitation data and spontaneous speech it is possible to identify the basic phonological characteristics of prosodic

phrases and then to investigate how those characteristics vary in context. Since this project is the first detailed investigation of the phrase-level prosody of Bobo, the analysis focuses primarily on the elicited data with the spontaneous speech used to add additional information about context and variation. The chapter on word prosody also incorporates some conclusions based on descriptive fieldwork notes.

The following is a general description of the design and collection of the data. The analysis of the data frequently depends on the specific questions being addressed and is discussed in the relevant chapters.

2.1.1 Speakers

Data from five adult speakers, three female (S1, S2, S5) and two male (S3, S4) were collected. All five are native speakers of Bobo from the Kuinima neighborhood in Bobo-Dioulasso. They are all fluent in Jula, which they have spoken from a very young age. Three are also fluent in French (S1, S3, S4) and assisted with translations of stimuli and glossing of spontaneous speech. Two have limited French (S2, S5). S5 only participated in recordings of spontaneous speech, specifically monologues and interviews. S2 participated in both types of tasks but was always accompanied by another speaker during the elicitation tasks.

2.2 Elicitation Tasks

There are seven sets of elicitation sentences, each designed to investigate a specific question (or set of questions) about the prosodic structure of Bobo. I created the sentences in Bobo and then checked for accuracy and acceptability with the speakers. If any of the speakers objected to a sentence, it was changed. During recording, the sentences were presented in random order on a computer screen in both Bobo and French. The French translation was provided in case the meaning of the Bobo sentence was ambiguous. The recordings were made in a quiet room in a

house in a residential neighborhood using either a Shure SM-10A head-mounted microphone and a Tascam DR-40 digital recorder or a Sennheiser E845-S microphone and a Scarlett 2i2 audio interface. Although the room was not as quiet as a laboratory, background noise is minimal and does not often interfere with phonetic analysis. The speakers preferred to read the elicitation sentences before recording in order to ensure that they understood all of them, which they were allowed to do. Before recording, they were instructed to say the Bobo sentences aloud clearly and naturally. They were asked not to read the French aloud.

A quick note about the use of written Bobo in the elicitation prompts is necessary at this point. Since Bobo does not have a written tradition, speakers can find full sentences difficult to read. Bobo is only written occasionally and usually in short snippets—for example, in song titles and shop names. When speakers do write it, there is no conventional orthography. They use either French orthography, Jula orthography, or a mixture of both. The Jula orthography is much more well-suited to Bobo than the French since the phonemic inventories of Bobo and Jula are nearly identical. For this reason, I used a slightly modified version of the Jula orthography in the elicitation sentences. However, tone is not marked, which leads to significant ambiguity. Additionally, reading Bobo still does not always come naturally even when a familiar orthography is used. The four speakers who participated in the elicitation tasks had worked with me for a while and in the process had become more used to reading Bobo, although French translations were still necessary for disambiguation. Additional speakers were recruited for the elicitation tasks but had too much difficulty and their data were discarded. S2 is partially literate and was assisted by another speaker during the recordings of elicitation sentences. In the case that she had difficulty with a word in a prompt, the speaker assisting her was instructed to tell her the correct word but not to say the entire phrase. The process of recording was more time intensive for S2 and she did

not do as many repetitions as the other speakers.

Each recording was kept under five minutes long to reduce fatigue and boredom, which means that some elicitation sets were recorded across separate days. If the set that was recorded that day contained matched sentences (e.g. a control and a test), they were recorded on the same day. Second repetitions were recorded at least two days after the first repetitions when possible, to avoid repetition effects. In the case of a speaker error, the sentence was discarded from the analysis. An error is defined as the speaker not following the prompt (e.g. producing a different word or sentence type) or obvious disfluencies such as mid-word hesitation or self-correction. The speakers were asked to repeat sentences with errors if I noticed the error during the recording session.

2.2.1 2014 elicitation tasks

This data was collected in the summer of 2014 for a preliminary investigation into the phrase-level prosodic characteristics of Bobo. My analysis of this data guided the data collection in 2016-2017. The 2014 elicitation data consists of two sets of sentences:

M-toned statements (40 sentences, 151 tokens): This set of sentences consists of simple declarative sentences in which all words have an M tone melody when possible. It is easier to identify changes in F0 that are due to phrase-level prosody when lexical tones are kept constant (Himmelmann & Ladd, 2008). The sentences vary in length but only contain a single clause. One limitation of this data set is that the majority (30) of the sentences end in a transitive verb with a grammatical L tone melody, which means that an L% boundary tone, if present, would not be easily discernable. The remaining sentences (ten) end in an adverb with an M tone melody. S1, S2, and S3 participated in this task. Each recorded two repetitions, but S3's second repetition was lost due to a technical malfunction. The same malfunction claimed many miscellaneous tokens from the remaining repetitions, leaving 154 tokens out of the initial 240 recorded. Three of these were

excluded from the project entirely due to speaker error.

Questions and responses (30 sentence pairs, 48 tokens): This set consists of ten *wh*-questions and 20 polar questions with accompanying responses. The response is always a full sentence that mirrors the question, as is shown in examples 5-6.[7]

(5) *kɔ³ be³ ŋwĩ¹?*
who you bite
"Who bit you?"

ɲi³ni³ *ne³ ŋwĩ¹.*
imbecile me bite
"An imbecile bit me."

(6) *sõ³ nõ³ ma¹ra¹?*
snake child injure
"Did a snake hurt a child?"

sõ³ nõ³ ***la¹ga¹****.*
snake child chase
"A snake chased a child."

The purpose of these question-and-response pairs was to investigate both the prosodic properties of questions and how focus is marked. In all of the responses, a single word item is the answer to the *wh*-question (and therefore in narrow focus) or a correction to one of the words in the polar question (and therefore in contrastive focus). The focused word was shown in bold in the sentence prompt. The tone melody of the focused word varies between L, M, and H, while other words have an M tone melody when possible. S1 and S2 recorded two repetitions of this task. In the first repetition, S1 posed the question and S2 responded, and in the second repetition the roles were reversed. Six tokens were excluded because of a recording error. An additional six tokens were excluded from the project due to speaker error, leaving 48 question-and-answer pairs

[7] Examples here are given in phonemic transcription rather than the orthography used in the prompts to show the structure of the elicitation sentences. The major difference is the tone marking, which was not present in the prompts.

included.

The 2014 elicitation sentences are discussed when relevant, but the majority of the analysis presented in the following chapters is based on data collected during 2016-2017. This later data contains more varied elicitation sentences as well as additional tasks.

2.2.2 2016-2017 elicitation tasks

2.2.2.1 Word stress elicitation task (45 sentences, 136 tokens)

This is the first of the elicitation tasks collected during fieldwork in 2016-2017. The purpose of this elicitation task is to investigate the existence of lexical stress. Bobo does not have contrastive lexical stress, but Morse (1976) claims that it does have automatic stress that is determined by the tone melody of the word. The details of her analysis are discussed in §3.3. This elicitation task consists of 48 sentences, each of which contains a target word in phrase-medial position. The target word is bisyllabic and contains identical vowels to avoid inherent differences in duration or amplitude affecting the results.[8] The target word is underlined in example 7 below (but was not underlined in the prompts):

(7) *da^{1}la^{51}* *<u>to^{1}ro^{1}</u>* *kɔ^{3}* *mi^{3}ri^{3}* *ɲa^{3}ŋa^{3}mũ^{1}.*
girls soumbala with rice mix
"The girls mixed the soumbala with the rice."

The tone melodies of the target words include both contour tone and level tone melodies. When possible, there are three target words for each tone melody, but the requirement that the vowels be identical limited the number of potential target words. The level tone melodies (L-L,

[8] Words with identical vowels were used because it would be difficult to ensure equal numbers of the same vowels in each position otherwise. There is a very strong tendency for the second vowel in a bisyllabic word to be more sonorous than the first; for example, CuCa is a common vowel pattern, but CaCu is unattested. This elicitation set also included monosyllabic and trisyllabic target words because it was also used to investigate tone realization. The monosyllabic and trisyllabic tokens were not analyzed for this project and the numbers presented here only include the bisyllabic tokens.

M-M, H-H) and L-M are represented by three words each, and the tone melodies L-H and H-L are represented by two words each. The are no target words with an ML tone melody that met the criteria, so ML tone was not included. Thus, there is a total of 15 target words. Each target word is placed in three different carrier phrases to reduce potential confounded introduced by the phrase context. S3 and S4 recorded two repetitions each. S2 recorded one repetition but her data was not analyzed due to time constraints. Eight tokens were excluded due to a speaker error and one was excluded because the speaker placed a phrase boundary after the target word. 23 tokens were excluded due to a recording error. 26 tokens were excluded because the target segments could not be segmented. A total of 136 tokens were analyzed.

2.2.2.2 Phrase prosody elicitation task 1 (seven sentence triplets, 186 tokens)

The purpose of this elicitation set is to investigate the prosodic characteristics of basic declarative statements, polar questions, and *wh*-questions. It consists of nine sets of matched sentences containing one declarative statement, one segmentally identical polar question, and one *wh*-question that is also segmentally identical except for the *wh*-word. Each sentence contains a phrase-final target word and a phrase-medial target word, which are underlined in example 8 below (they were not underlined in the prompt):

(8) *nɛ³mɛ³ ɲi⁵ kia¹ kɔ³ bi⁵.*
boy salt search with you
"The boy searched for the salt with you."

nɛ³mɛ³ ɲi⁵ kia¹ kɔ³ bi⁵?
boy salt search with you
"Did the boy search for the salt with you?"

we¹⁵ nɛ³mɛ³ ɲi⁵ kia¹ kɔ³ bi⁵?
which boy salt search with you
"Which boy searched for the salt with you?"

The target word has either an L, M, or H tone melody and is between one and three syllables

long. The variation in tone melody is intended to make pitch phenomena associated with the end of the phrase observable. For example, an L% boundary tone should be observable when it follows an H tone, an H% boundary tone should be observable when it follows an L tone, and both L% and H% boundary tones should be observable when they follow an M tone. Varying the tone melody also allows an analysis of whether (and how) lexical tone interacts with phrase-final prosodic phenomena. The phrase-medial target word allows the phonetic properties of the phrase-final target word to be compared to an equivalent in phrase-medial position within the same utterance. The control word always has the same final vowel and final lexical tone as the target word. To reduce the effects of the tone context, it is always preceded by the same tone as the final syllable of the target word (in this example it is an M tone). There was a total of nine target words.

Each set of sentences was presented in both the simple present and past perfective. Le Bris and Prost (1981) claim that the difference between these two TAM (tense, aspect, modality) categories is marked by tonal alternations on the verb. Although this turned out not to be true for the speakers who participated, for whom simple present and past perfective are homophonous, both TAM categories were included in the analysis because they provided extra repetitions. S3 and S4 recorded two repetitions each, and S2 recorded one repetition in the past perfective only. 30 tokens were excluded because the target word had an unexpected tone melody and three were excluded due to speaker error, leaving seven target words and 186 tokens from this elicitation set included in the project.

2.2.2.3 Phrase prosody elicitation task 2 (21 sentence pairs, 203 tokens)

The purpose of this elicitation task is to provide more data on the prosodic characteristics of declarative statements and polar questions. The previous elicitation task does not contain much data. The requirement that each phrase-final target word be matched with a phrase-medial target

meant that only target words with level tone melodies could be used; there are not enough words with contour tone melodies that meet the criteria. The sentences in this elicitation task do not have a phrase-medial target word, which allows for more varied phrase-final target words. In addition to L, M, and H tone melodies, there are phrase-final target words with LH, HL, and LM tone melodies. (ML is not included due to a lack of unambiguously ML-toned words that can be used phrase-finally.) There are 29 target words in 29 sentence pairs. Each sentence pair consists of a declarative statement and a segmentally identical polar question, as is shown in example 9 below:

(9) *a*^{1} *ja*^{3} *tũ*^{5}*=õ*^{1} *kɔ*^{3} *ne*^{3} *pu*^{1}.
3.SG go market=LOC with 1.SG.OBL younger.sibling
"She went to the market with my younger sister."

a^{1} *ja*^{3} *tũ*^{5}*=õ*^{1} *kɔ*^{3} *ne*^{3} *pu*^{1}*?*
3.SG go market=LOC with 1.SG.OBL younger.sibling
"Did she go to the market with my younger sister?"

Since the polar questions are identical to their corresponding statements, speakers rely on their prosodic characteristics to signal that they are questions. *Wh*-questions are not included in this elicitation task because a preliminary analysis of the data from the previous elicitation task showed that their prosodic characteristics are more variable than polar questions, but do not include any phenomena not found in polar questions. S3 and S4 recorded two repetitions and S1 recorded one repetition. A total of 290 tokens were recorded. 80 of these tokens were excluded because the target word had an unexpected tone melody, leaving 21 sentence pairs included.[9] Seven tokens were excluded due to speaker error. 203 tokens were included in the project.

[9] This is an unfortunately large amount of data to exclude. There were two problems in the choice of target words: (1) Some words had a different tone melody than previous work indicated or their tone melody varied, and (2) some words were unexpectedly affected by grammatical tone. I excluded them entirely to avoid any uncertainty about their tone melodies. Since the target words were repeated across multiple tokens, the problem was magnified.

2.2.2.4 Negated sentences elicitation task (22 sentence pairs, 175 tokens)

The purpose of this set of elicitation sentences is to investigate the prosody of negated utterances in Bobo. The prior literature disagrees about how negation is marked. All agree that there is a clause final negation particle, *=ga⁵/ŋa⁵*. Le Bris and Prost (1981) claim that in a negated sentence, an H tone is also added to the subject. They do not make any claims about the intonation of the utterance. Morse (1976) claims that an H tone is added to the subject or the TAM marker, and also that the overall pitch of the utterance is raised. If the latter part of her claim is true, then it would be an example of an intonational phenomenon that cannot be analyzed in terms of a discrete intonational tone that is simply inserted into the tonal string. The structure of a negated sentence is illustrated below in example 10 below. According to Le Bris and Prost (1981), the subject *ma³* has a suffixed H tone. According to Morse, the suffixed H tone occurs on either the subject or the TAM marker *ma³ane³*, and the entire utterance also has raised pitch.

(10) *ma³* *ma³ne³* *bɛ³re³* *kɔ³* *a¹we¹=ga⁵.*
1.SG PAST talk with 3.SG.OBL=NEG
"I wasn't talking with him."

This elicitation set consists of 22 pairs of sentences containing matched negated and non-negated sentences, as shown in example 11:

(11) *ku³* *ti¹* *po³ro³=o¹=ga⁵*
turtle be bag=inside=NEG
"The turtle is not in the sack."

ku³ *ti¹* *po³ro³=o¹*
turtle be bag=inside
"The turtle is in the sack."

The lexical tone melody of the subject was varied between L, M and H to make the addition of a tone easy to identify, regardless of whether the tone is itself L, M, or H. There is only one TAM particle that occurs in this position, *m³ane³*; six of the sentences contain this particle. The

full set of subject pronouns was also included in the set of target subjects, since pronouns are sometimes affected by grammatical tone processes in a different manner than nouns.

S2 recorded one repetition and S3 and S4 recorded two repetitions. 218 tokens were recorded, but 39 were excluded due to an error in the recording and four were excluded due to speaker error. 175 were included in the project.

2.2.3 Questionnaire on Information Structure (205 prompts)

The Questionnaire on Information Structure (QUIS) (Skopeteas et al., 2006) is a collection of tasks designed to elicit ways of marking information structure in languages whose information structure has not been investigated extensively before. The tasks are designed to elicit spontaneous or semi-spontaneous speech with minimal use of an intermediary language. I used the elicitation questionnaire and the map task. The map task is described in §2.3.

The elicitation questionnaire is designed to elicit the encoding of information structure, especially broad, narrow, and contrastive focus. The majority of the prompts are interactive and involve two speakers. One speaker poses a question or makes a statement, and the other responds with either the answer to the question or a correction. Instead of using pre-designed sentences for the response, a minimal prompt is given. The prompt includes as little information as necessary to elicit the desired type of response. Unlike elicitations that use pre-designed sentences, this allows the speaker answering the question to use morphosyntactic focus-marking strategies that the researcher might not have anticipated. This addresses one of the limitations of the 2014 data collection, which would not have captured any optional morphosyntactic focus marking. There are 196 question-and-response prompts or statement-and-response pairs in the questionnaire. There are also nine prompts that elicit different information structures without the interactive context, for a total of 205 prompts. In many of the prompts, the same lexical item is repeated in different focus

contexts in order to allow comparisons between them. However, because the responses are free form, with the exception of what is contained in the prompt itself, the phonetic and phonological context of the focused item is not controlled. This means that the data collected from this questionnaire is not suited for quantitative analysis. Example 12 below shows one question and response prompt before and after being translated into Bobo.

(12) Question prompt: **What did the woman eat?**
Response prompt: **[beans]**

Question prompt: **Ya ye ŋwonon n zon?**
Answer prompt: **[demen]**

The questionnaire was published in English. It was translated into French by me, and then translated into Bobo with the assistance of S3 and S4, who are the speakers who participated. They were given a print-out that included the prompt in French and English. The questionnaire was recorded once. S4 posed the question and S3 responded. He was asked to respond how he would naturally in a conversation, with the caveat that the responses should be full sentences rather than only the item included in the prompt. S3 and S4 then assisted with the transcription and glossing of any difficult sentences.

2.3 Map Task

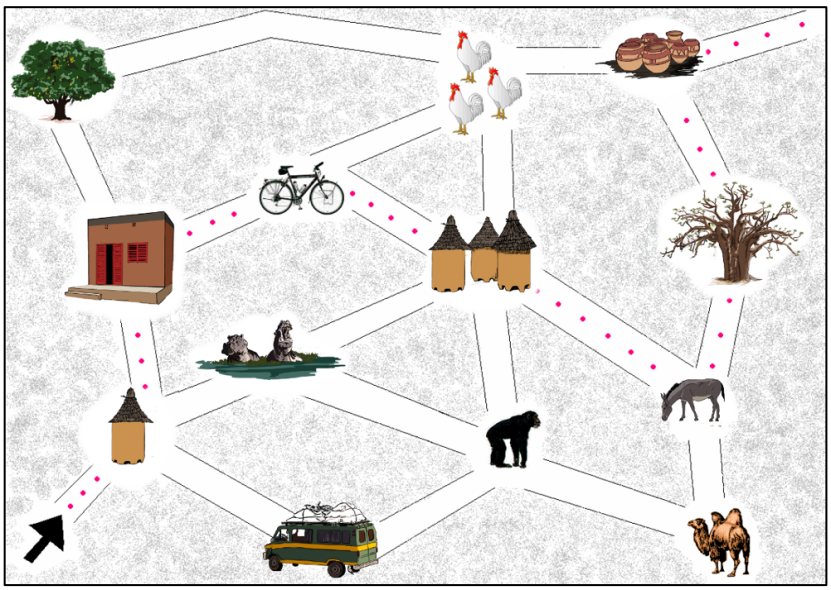
The map task is a cooperative task designed to prompt interaction between two speakers. One speaker is given a treasure map with a route marked. The second speaker is given a similar map without a route marked. The second speaker is then asked to draw the route that is on the first speaker's map by asking questions of the first speaker. Neither can see the other's map. This encourages the speakers to ask questions and provided responses in which certain items on the map are focused. The landmarks can be chosen by the researcher to elicit specific words. Some of the landmarks on the second speaker's map are changed. Since the speakers do not know which

items are changed and cannot see each other's maps, they must navigate these discrepancies by discussing them, which encourages more interaction and the use of contrastive focus. Figure 3 and Figure 4 show one of the pairs of maps used during this task.

Data from two pairs of maps were collected. The maps were based on the maps given in the Questionnaire on Information Structure (Skopeteas et al., 2006), but some of the landmarks were changed in order to target different lexical tones. S2 and S3 participated in the task. They did not see the maps prior to the task. They were instructed that this was a map from a starting point (the arrow) to the village exit (top right). To contextualize the task, they were asked to imagine that S2 was at the starting point and had called S3 to ask for directions to the exit. They then switched roles for two more interactions. Although S3 had seen the routes during the first interactions, he was able to pretend that he had not. The recordings were then transcribed and glossed with the help of S3 and S4.

Figure 3

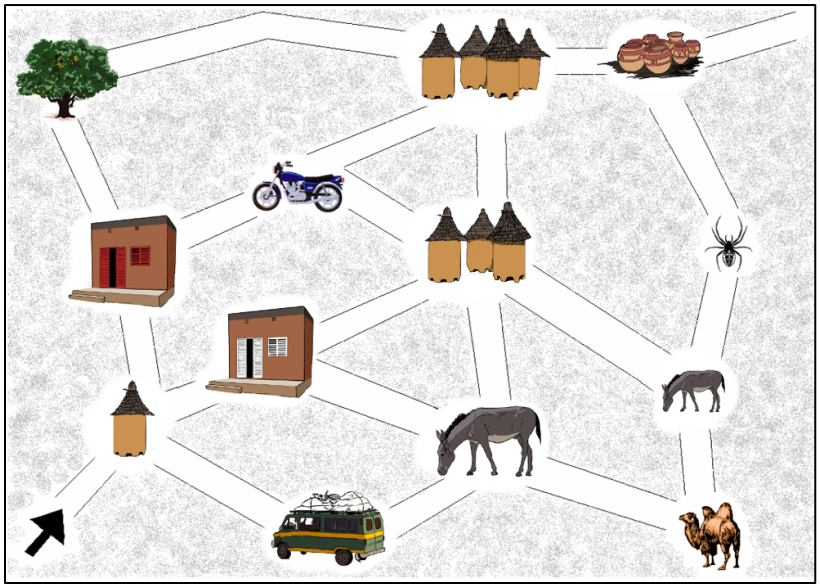
First map from map task



Note. The speaker giving directions is given the map with the route marked.

Figure 4

Second map from map task



Note. The speaker seeking directions is given a similar map but with no route marked and some landmarks changed.

2.4 Other Spontaneous Speech Tasks

This data was collected between 2014 and 2017. It was not specifically collected for the investigation of prosody but is used during the analysis. All five speakers participated in some of the spontaneous speech tasks. They consist of: (a) narratives, (b) other monologues in which the speaker explains how to perform a task or shares their opinion on a topic, (c) conversations in which two speakers discuss a topic, (d) interviews in which one speaker asks another questions about a topic. For each of these tasks, the speakers were asked to choose a story or topic that they preferred. The recordings were then transcribed and glossed with the help of S1, S3, and S4. A subset of the spontaneous speech data was analyzed more closely; 659 phrases, defined by the presence of a pause, were coded according to utterance type, their final lexical tone, and surface F0 contour. Examples from the rest of the spontaneous speech data are sometimes used when relevant.

Chapter 3

Word-Level Prosody

3.1 Overview and Research Questions

This project is primarily concerned with the phrase-level prosody of Bobo. However, word-level and phrase-level prosody cannot be treated as separate systems and neither can be understood without the other. Word-level prosody in Bobo has been treated more extensively than phrase-level prosody in previous works (Morse, 1976; Le Bris & Prost, 1981; Sanou, 1993), but a comprehensive description does not yet exist. In this chapter, I address some aspects of word-level prosody in Bobo that either have not been well-described yet, or about which previous researchers disagree: the inventory of contour tones (specifically the existence of ML and HL), lexical stress, and the prosodic foot. This chapter draws on descriptive fieldwork as well as the recorded data described in Chapter 2.

3.2 Contour Tone Inventory

Previous descriptions (Morse, 1976; Le Bris & Prost, 1981; Sanou, 1993) disagree on the contour tone inventory of Bobo. They do agree on the following: There are three tone levels, L, M and H. There are at least two contour tones, LM and LH. Tone melodies, or the sequences of tones within a morpheme or word, play an important role in the grammar. The tone melodies of certain word classes are restricted, and grammatical tone alternations often involve the replacement of an entire tone melody rather than the deletion or affixation of a tone. At his point, the analyses depart from one another.

The first departure is what is considered to be the tone bearing unit (TBU). Sanou (1993) claims that the TBU is the mora and links each individual component of a contour tone to its own mora; thus, according to his analysis, contour tones can only occur on long (bimoraic) vowels. Additionally, for Sanou vowel length is a productive distinction, although he provides limited evidence for phonemic differences in vowel length that are independent of tone.[10] Sanou's analysis is similar to Boone's (2016) analysis of Northern Bobo Madaré, where contour tones are similarly restricted to long vowels. Neither Morse (1976) nor Le Bris and Prost (1981) agree, however. According to their analyses, the vowel length distinction is not productive, and contour tones do occur regularly on short vowels. They claim that the TBU is the syllable. As discussed in §1.2 my own analysis of vowel length is consistent with Morse and Le Bris and Prost, which is that Bobo does not have a productive vowel length distinction. This means that whether the TBU is the mora or syllable is moot. Both analyses make the same prediction, i.e. that the individual components of a contour tone can be linked to the same TBU.

The second departure is the inventory of contour tones. Le Bris and Prost (1981) argue for the largest inventory. In addition to the two rising tones, LM and LH, they claim that Bobo has two falling tones, ML and HL. Morse (1976) agrees that Bobo has an ML tone but disagrees about HL; her analysis is that HL only occurs on a single word and is therefore not a productive part of the inventory. Finally, Sanou (1993) does not include any falling tones in his inventory. He only includes the two rising tones, LM and LH. The disagreement about the inventory of contour tones

[10] Sanou (1993) provides only a few true minimal pairs. Three involve adjectival, adverbial or stative word forms, which is where the majority of long vowels occur: *kùː* 'volumineux' vs. *kù* 'credit', *toːro* 'fumée' vs. *toro* 'aller à quatre pattes', *pɛː* 'complètement (idiot)' vs. *pɛ* 'faire (froid, chaud'). The three remaining pairs are *suːro* 'soigner' vs. *suro* 's'affaisser', *wɔːrɔ* 'des os' vs *wɔrɔ* 'servir (pâte)', and *yɛːro* 'ombre' vs *yɛro* 'se câcher'. I was not able to specifically verify these minimal pairs because I did not acquire his dissertation until after my fieldwork was over.

is not surprising since the identification of contour tones can be difficult. In a phrase-medial context, contour tones can be affected by grammatical tone processes (e.g. *pu¹³* 'younger sibling' > *ne⁵ pu¹* 'my younger sibling'). They can also be affected by undershoot, especially in faster speech. Contour tones are produced more reliably in isolation, but then phrase-final phenomena are a potential concern, as any phrase-final F0 phenomena that interact with lexical tone can obscure the underlying tone. Adding to the difficulty of reliably eliciting contour tones is the lack of full minimal sets for all of the tone contrasts. Some lexical tone contrasts can be demonstrated through minimal sets, while others cannot. The contrast between the three tone levels L, M, and H can be demonstrated through minimal triplets:

(13)	*ku¹*	'armpit'	*ta¹ba¹*	'tobacco'
	ku³	'turtle, tortoise'	*ta¹ba³*	'support'
	ku⁵	'debt'	*ta¹ba⁵*	'knife'

The two rising tones, LM and LH, can also be demonstrated through minimal pairs. They are distinct from each other, as shown in 14, and also from H and M, as shown in 15 and 16:

(14)	*sõ¹³*	'sing (intransitive)'	*kõ¹³*	'house'
	sõ¹⁵	'man, person'	*kõ¹⁵*	'voice, noise'

(15)	*ŋwẽ¹⁵*	'razor'	*sẽ¹⁵*	'mouse'
	ŋwẽ⁵	'there'	*sẽ⁵*	'tomorrow'

(16)	*sõ¹³*	'sing (intransitive)'	*su¹³*	'pot'
	sõ³	'cultivate'	*su³*	'remedy'

The ML and HL tones, about which previous sources disagree, are more difficult to demonstrate, but a partial minimal set exists for the ML tone, as is shown in example 17. This demonstrates that ML is distinct from both L and M.

(17)	*ma¹*	'in, at'	*ku¹*	'debt'
	ma³	'I'	*ku³*	'turtle, tortoise'
	ma³¹	'friend'	*ku³¹*	'back'

However, for Sanou (1993) the distinction between these sets appears to be length and not

tone; for example, he transcribes 'friend' with a long vowel and a low tone, meaning that it is no longer a minimal pair for tone. Since I was not able to establish a productive length distinction (as discussed in §1.2), I follow Morse (1976) and Le Bris and Prost (1981) in analyzing these words as having an ML tone.

There are no minimal sets that can demonstrate the existence of the HL tone, and it is also the tone about which previous works contain the most disagreement. Morse (1976) claims that an HL tone only occurs on a single word: *na*^{51} 'mother.' This is specifically a vocative; the non-vocative word is *sie*^{5}. If HL tone does occur only on this word, then it is an extremely marginal contour, as not only would it be limited to a single word, that word could have special prosodic characteristics due to its being vocative. Le Bris and Prost (1981), on the other hand, include the HL tone the inventory without any special comment and use it more frequently in transcription, including on nouns, adjectives, and adverbs (e.g. *gba*^{51} 'sheep,' *bĩ*^{51} 'green, unripe, uncooked,' *ba*^{3}*siɛ*^{51} 'now')[11]. One reason for their differing analyses appears to be Morse's analysis of the H tone melody. She states: "For example, morphophonemically there are two classes of morphemes with a High-High lexical tone pattern. One remains High-High and the other becomes High-Low when followed by either a low tone or a pause." (p.84) Morse does not provide any further detail or examples, but where Le Bris and Prost transcribe a word as HL, she transcribes it as H. It appears that when there is a high falling pitch contour in a monosyllabic word, Morse analyzes the word as having an underlying H tone, while Le Bris and Prost analyze it as either H or HL. Finally, Sanou (1993) does not include HL in his inventory at all. He does not discuss previous analyses so the reason that his analysis differs is not clear.

[11] Original transcriptions and glosses: *gbâ* 'mouton'; *bîn* 'vert, pas cuit, pas mûr; *bāsiê* 'maintenant'

My own analysis is that there is an HL tone, although its realization is dependent on the phrase context. To demonstrate that the HL tone exists, it is necessary to establish the contrast between HL and ML, and HL and H, as these are the most phonetically similar tones. In some contexts, ML and H are indistinguishable from HL; however, in other contexts they are phonetically distinct. For ML and HL, the distinction is the height of the F0 peak, which is illustrated in Figure 5 and Figure 6. When produced in isolation, the ML-toned word *ma³¹* 'friend' has a lower F0 peak than the HL-toned word *gba⁵¹* 'sheep'. In running speech, the height of the F0 peak for an HL tone is not always as pronounced. The distinction is clear in careful speech, however. Importantly, ML is never produced with a similar F0 peak to HL, which is consistent with M tones in general. The F0 peak in Figure 6 can thus only be the result of an underlying H tone.

Figure 5

ma³¹ *'friend' (S3, M)*

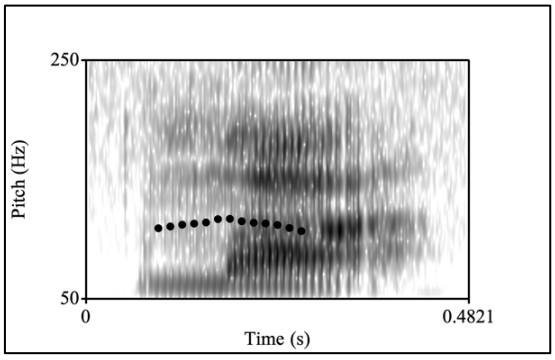
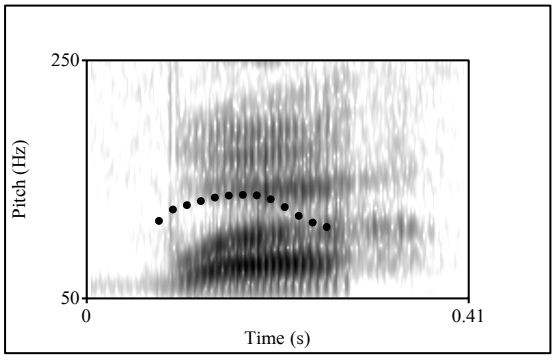


Figure 6

gba⁵¹ *'sheep' (S3, M)*



The distinction between H and HL is the presence of a fall in F0 for HL. This is shown in Figures 7 through 10, which show two H-toned and HL-toned words produced in isolation. The H tone (shown on the left, Figure 7 and Figure 9) is not produced with a fall in F0, whereas for the HL tone (shown on the right, Figure 8 and Figure 10), the fall in F0 can be quite pronounced.

Figure 7

tũ^{5} *'market' (S3, M)*

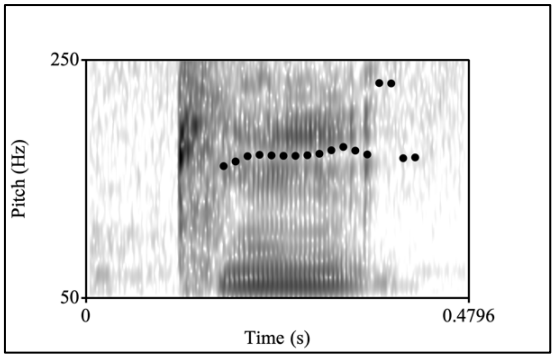


Figure 8

dũ^{51} *'chest' (S3, M)*

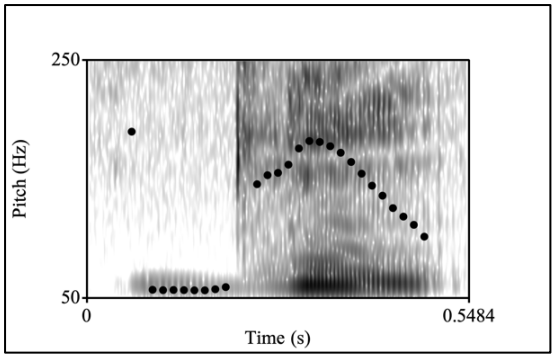


Figure 9

ba^{1}lo^{5} *'iron' (S3, M)*

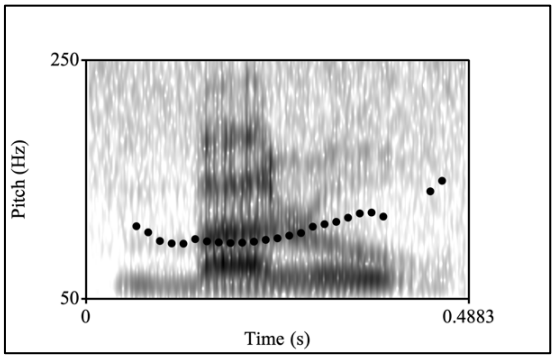
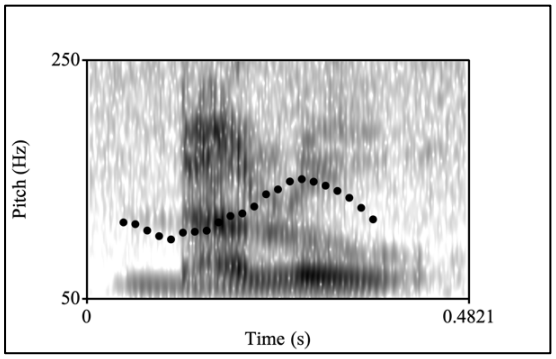


Figure 10

da^{1}lo^{51} *'girl' (S3, M)*



The distinction between H and HL is clear in a phrase-final context when there is no final boundary tone. In a non-phrase-final context, the distinction is less clear. The pronunciation of H and HL tones can be affected by the context, by undershoot, and by other factors that are not yet well understood. In the case of H tone, it is common for F0 to fall during the H-toned syllable when it is immediately followed by an L tones. Figure 11 and Figure 12 show an example for the H-H-toned word *sɔ^{5}gɔ^{5}* 'panther'; in Figure 11 there is a fall in F0 before a following L tone, while in Figure 12 there is not. Since the fall in F0 in phrase-medial H tones is not reliably triggered by a following L tone, it is probably caused by coarticulation with the following L tone rather than by a regular morphophonological process, in which case we would expect it more regularly.

Before an M tone, an H tone can have a fall in F0, but it will not fall below the M-tone target. Before another H tone there is rarely a fall in F0.

Figure 11

sɔ^{5}gɔ^{5} *'panther' produced before an L tone (S2, F)*

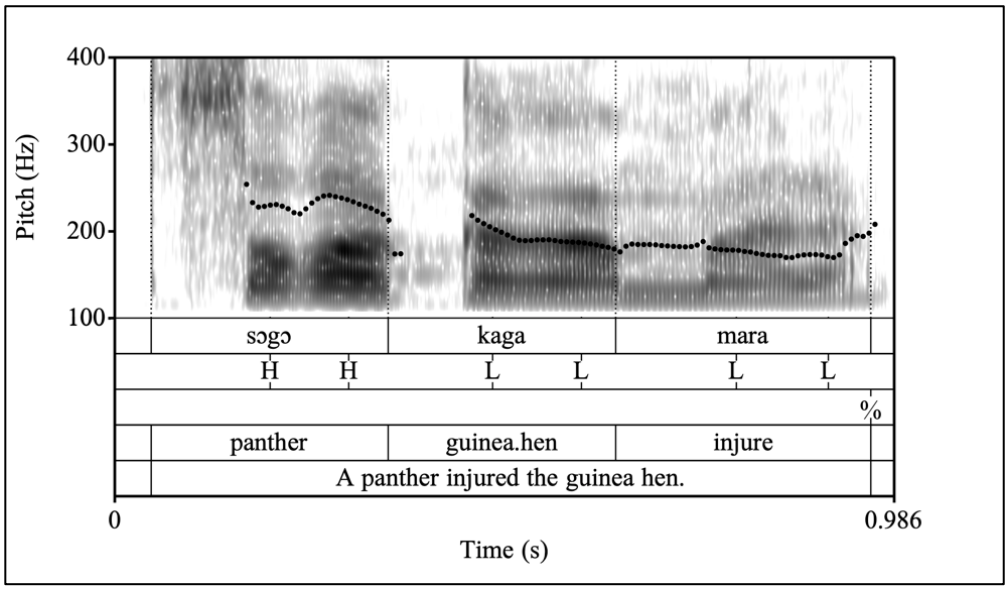
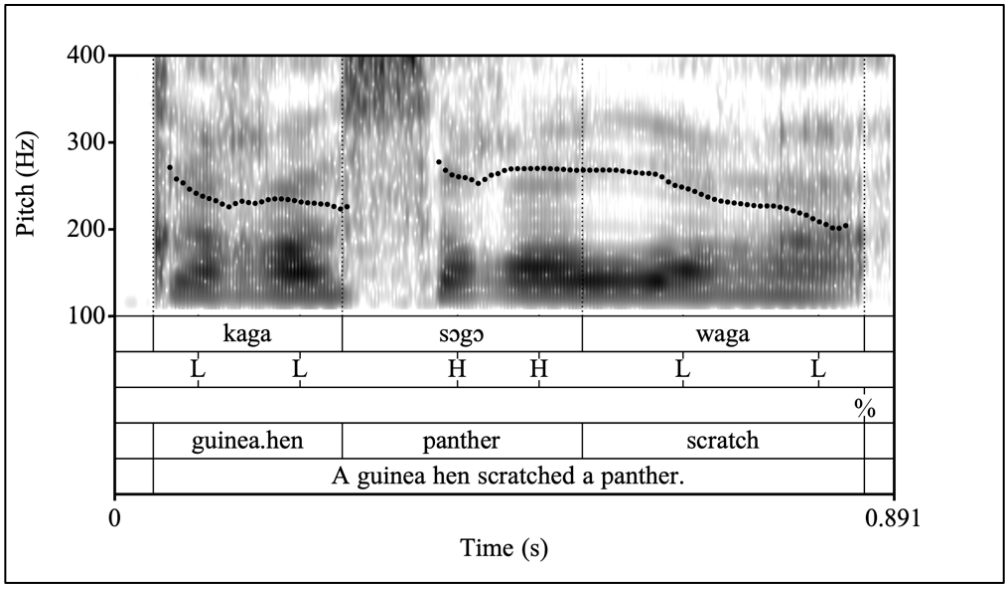


Figure 12

sɔ^{5}gɔ^{5} *'panther' produced before an L tone (S1, F)*

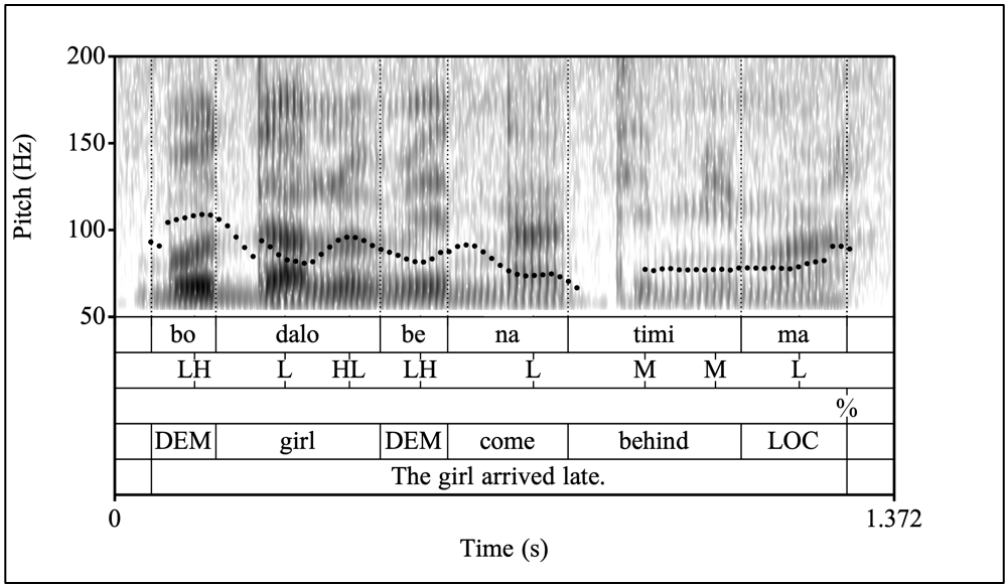


Like H tone, an HL tone will sometimes, but not always, be produced with a fall in pitch

before a following L tone. An example is shown in Figure 13 and Figure 14. In Figure 13, the HL-toned syllable in *da¹lo⁵¹* 'girl' has a fall in F0, while in Figure 14 it does not, even though it is followed by an L tone in both phrases. There are fewer examples of non-phrase-final HL-toned syllables in the recorded data, which makes it difficult to draw conclusions about the causes of the variation.[12]

Figure 13

da¹lo⁵¹ *'girl' produced before an L tone (S4, M)*

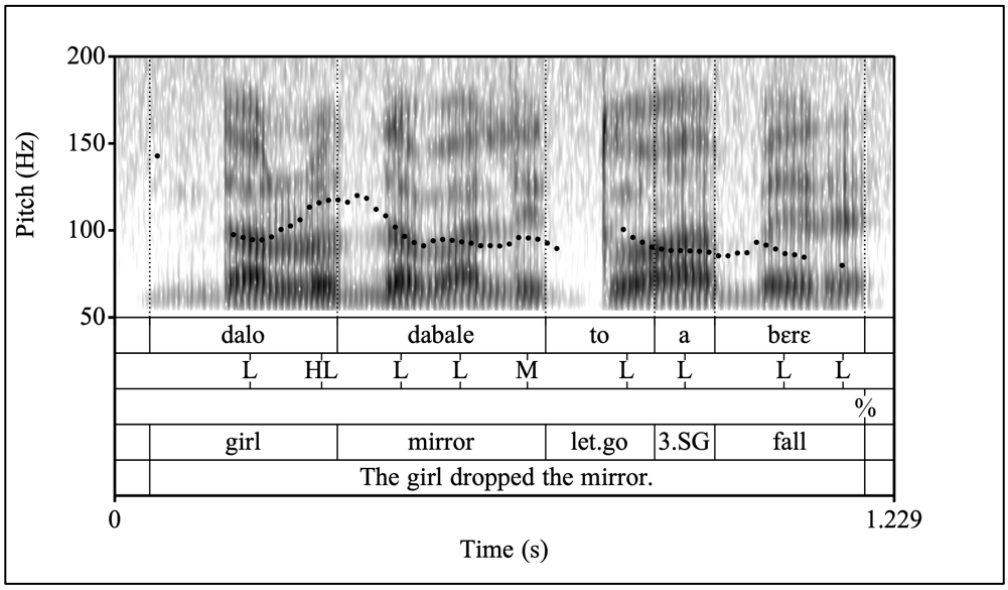


Thus, in a non-phrase-final context, an H-toned syllable can be realized similarly to an HL-toned syllable and an HL-toned syllable can be realized similarly to an H-toned syllable. Nevertheless, H and HL are distinct in phrase-final position and this needs to be accounted for in the underlying representation of these words in some way. This data is not wholly consistent with

[12] In particular, examples of HL tone in different tone contexts are needed. All of the known examples of HL-toned syllables in the spontaneous speech data are followed by an L tone. This is due to HL being a relatively rare tone and L being the most common tone. My impression from non-recorded elicitation sessions is that HL tone frequently surfaces as H when in non-phrase-final context, regardless of the following tone, but how frequently needs to be investigated with systematically collected recorded data.

Figure 14

da[1]lo[51] *'girl' produced before an L tone (S4, M)*



either the analysis by Morse (1976) or the analysis by Le Bris and Prost (1981). Le Bris and Prost do not discuss the production of H and HL tone and their criteria for identifying a tone as either H or HL are unclear. Morse's analysis, which is that a subset of words with an H-H tone melody become H-L before a pause or an L tone, does capture some of the data. However, there are two issues with this analysis. The first issue is that the unit to which her analysis applies is the tone melody assigned to the word (H-H), rather than the individual tone that is assigned to the final syllable (H). She claims that words with an H-H tone melody belong to two morphophonemic classes that behave differently in a phrase context. This does not account for the existence of HL tone in monosyllabic words such as *gba*^{51} 'sheep' or in words that do not have an initial H tone such as *da*^{1}*lo*^{51} 'girl.' Instead, the data indicates that a surface HL tone is not limited to words with a specific tone melody. There is also no other evidence that it is associated with a particular morphophonemic class. The second (but smaller) inconsistency is the environment that triggers the alternation between the H and HL tone. Morse claims that the H-HL form occurs before pauses

and L tones, while the H-H toned form occurs elsewhere. However, this data shows that at the very least, there is more variation in the surface form than this. H and HL forms both occur before L tones.

HL tone is unique among the contour tones in that it only reliably surfaces as HL in phrase-final position. This raises the question of whether this its underlying representation is different in some way, e.g. if the L is a floating tone. It is not possible to answer this question with the existing data. Regardless, it is clear that HL contrasts with other lexical tones and therefore it is part of the lexical tone inventory. This results in the following tone inventory for Bobo: L, M, H, LM, LH, ML, and HL.

3.3 Phonetic Evidence for Lexical Stress

There is no contrastive lexical stress in Bobo and no indication that stress plays a role in the grammar, e.g. in word derivation or inflection. In this section, I examine whether there is phonetic evidence for non-contrastive stress. Here, I define stress as the syntagmatic relationship between prominent and non-prominent syllables within a domain (Ladd, 2008; Féry, 2017): Stressed syllables are strong, and unstressed syllables are weak. Under this definition, stress is an abstract and categorical phonological property that can be expressed in a variety of ways. Cross-linguistically, stressed syllables often have greater amplitude and duration than unstressed syllables (Ladefoged, 2003; Ladd, 2008; Féry, 2017).[13] Phonological alternations, such as the reduction of stressed vowels, can also indicate stress; these will be addressed in §3.4.

Since Bobo does not have contrastive lexical stress, I assume that any stress will be

[13] Most research on the relationship between stress and pitch has focused on languages with pitch accent, which Bobo does not have. Other effects, such as the expansion of pitch range, are difficult to investigate because of the effect of neighboring lexical tones. Therefore, pitch is not addressed here.

predictable from the phonological context. Since Bobo has a simple CV syllable structure and syllable weight cannot be a factor, any stress will likely be conditioned by either position or by lexical tone. Morse (1976) claims that Bobo has automatic primary and secondary stresses, which are determined by the tone melody of the word. In words containing identical tones, the first syllable has primary stress. Otherwise, primary stress occurs on the first H tone in the phonological word, or on the first M tone if there is no H tone present. She also claims that primary stress occurs "on the first element of tonal glides (either High or Low) occurring on single vowels in morphemes with a CV pattern" (p. 110), where "tonal glide" is a contour tone. This analysis treats stress as a property of the tone, rather than of the syllable, since stressing the first element of a contour tone would result in a stressed and unstressed tone within the same syllable. Morse does not define stress or provide any information on how stress is realized.

To investigate the existence of automatic stress, I examined the duration and amplitude of the vowels of bisyllabic words that were placed in a phrase-medial context. The recordings were collected as part of the elicitation task described in §2.2.2.1 (word stress elicitation task). The vowels were segmented in Praat and their durations and amplitudes measured using a Praat script. A statistical analysis is not possible because the results vary by speaker and by tone melody, meaning that tokens could not be pooled across these categories, and there were too few tokens to conduct statistical analyses for speakers separately.

There is no meaningful difference in amplitude between the first (V1) and second (V2) vowel. Table 1 shows the results for the level tone melodies for each speaker. For both speakers, the mean amplitude of the first and second vowel is nearly identical across the three tone melodies. Figure 15 and Figure 16 display the data. If there is lexical stress in level-toned words, it is not reflected in the relative amplitude of stressed and unstressed vowels.

Table 1

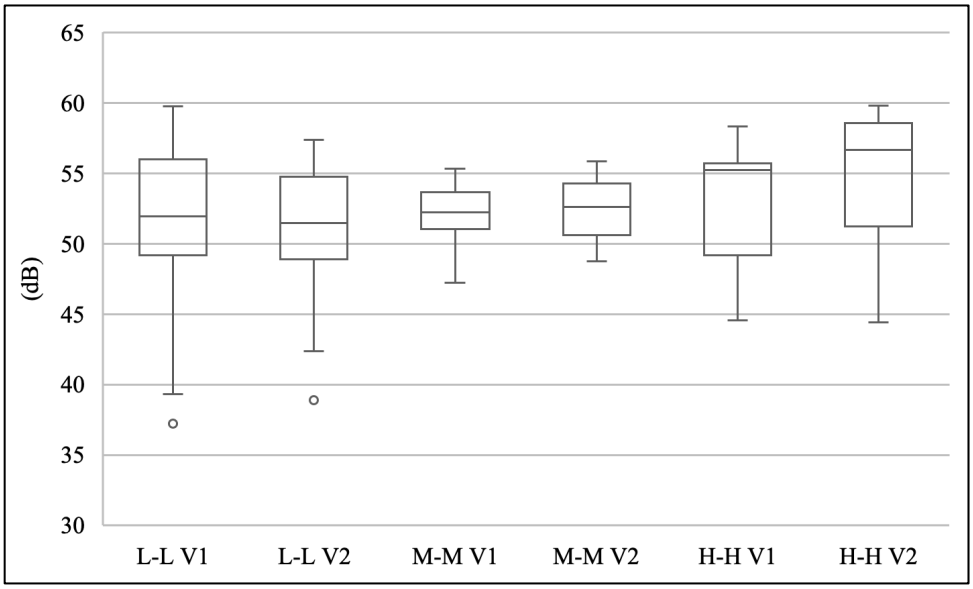
Mean Amplitude (dB) of Vowels in Bisyllabic Words with Level Tone Melodies

	V1 Amplitude		V2 Amplitude		V2-V1 Amplitude		# of Sentences
	M	*SD*	*M*	*SD*	*M*	*SD*	
S3							
L-L	51.3	6	51.3	4.9	-0.1	2.2	16
M-M	52.2	2	52.5	2.1	0.4	1.9	12
H-H	52.6	4.3	54.8	4.7	2.28	2.4	9
S4							
L-L	45.8	3.1	43.1	3.3	-2.7	1.3	15
M-M	43.9	1.8	43.5	3	-0.3	1.9	12
H-H	44.4	2.2	45.5	3.7	1.1.	1.9	6

Note. M=mean, SD=standard deviation. Means and standard deviations are calculated using raw data.

Figure 15

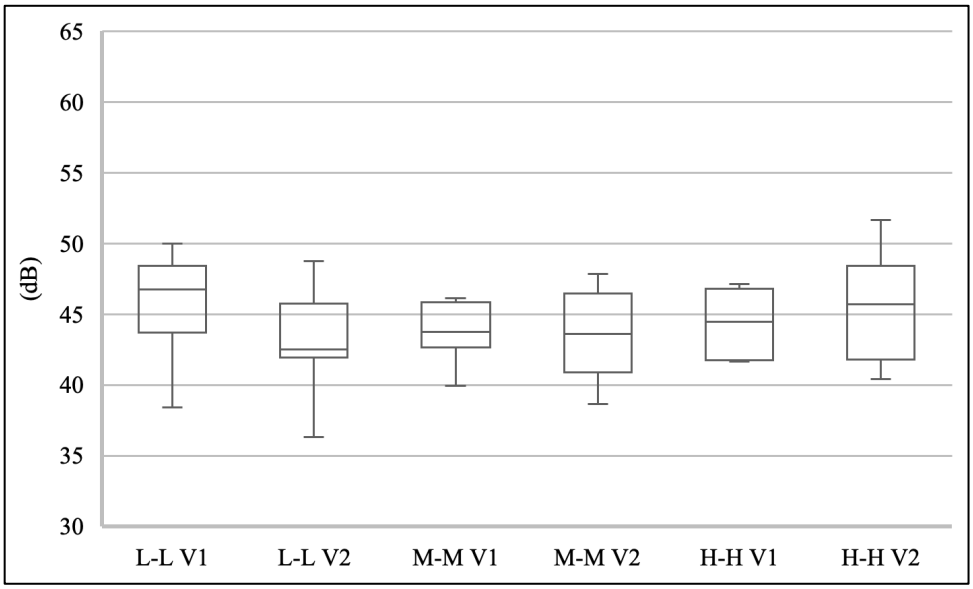
Amplitude of V1 and V2 in level-toned words (S3)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

Figure 16

Amplitude of V1 and V2 in level-toned words (S4)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

Duration is more complicated. For the level-toned words, the pattern differs by both tone melody and speaker. Their individual data is displayed in Table 2. For M-M toned words, V2 is slightly longer than V1 for both speakers, although given the variability of the data this might not be a meaningful difference. The difference in duration between V1 and V2 in L-L toned words is nonexistent for S4. For S3 it is similar to the one observed in M-M tones in magnitude but going in the *opposite* direction, namely V1 is longer than V2. In H-H toned words, the speakers diverge clearly: For speaker S3, V2 is longer than V1 by an average of 25 milliseconds, while there is no such difference for speaker S4. Thus, for S3, L-L has longer V1 than V2, and longer V2 than V1 for H-H and marginally for M-M, while S4 has longer V2 than V1 for M-M, and no durational difference (that goes beyond measurement error) for L-L and H-H. There is no systematic pattern noticeable in the data and the large variability in the data further supports this conclusion. Figure 17 and Figure 18 provide a visual representation of the vowel durations.

Table 2

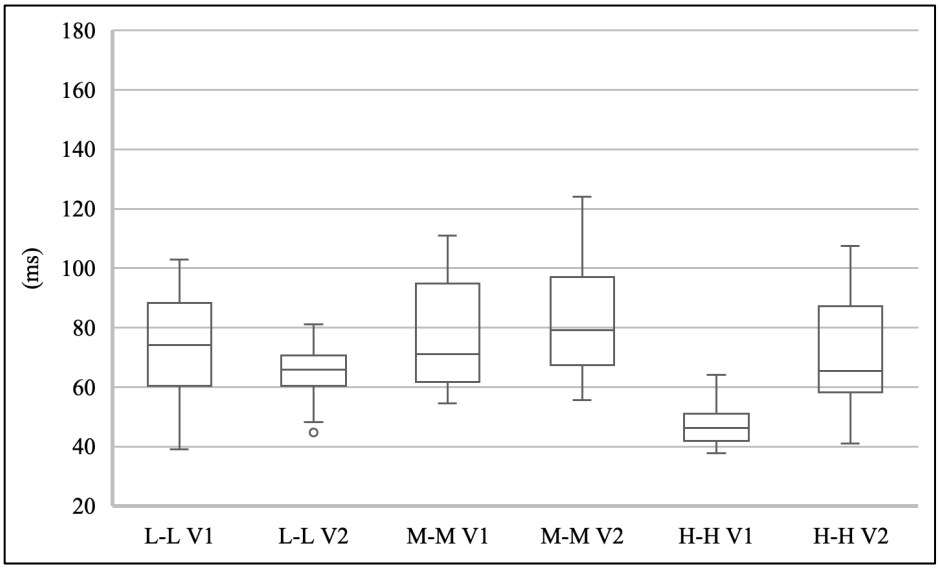
Mean Duration of Vowels (ms) in Bisyllabic Words with Level Tone Melodies

	V1 Duration		V2 Duration		V2-V1 Duration		# of Sentences
	M	*SD*	*M*	*SD*	*M*	*SD*	
S3							
L-L	73	17	65	9	-9	17	16
M-M	77	17	83	19	7	27	12
H-H	47	7	71	19	25	24	9
S4							
L-L	64	13	63	12	0	10	15
M-M	86	12	99	29	13	28	12
H-H	64	11	61	25	-2	16	6

Note. M=mean, SD=standard deviation. Means and standard deviations are calculated using raw values.

Figure 17

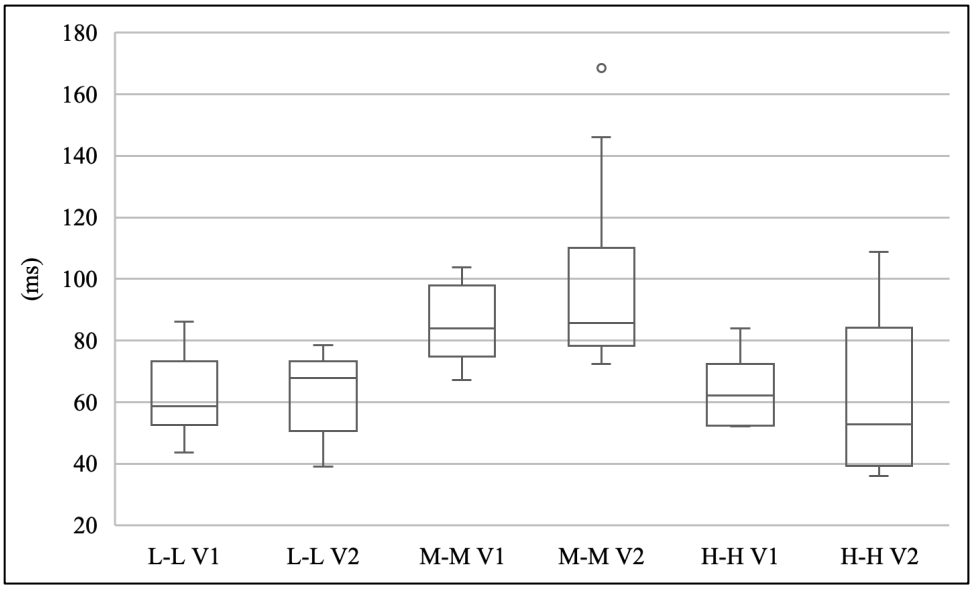
Duration of V1 and V2 in level-toned words (S3)



Note. Based on raw values. Whiskers display minimum and maximum values.

Figure 18

Duration of V1 and V2 in level-toned words (S4)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

To summarize, for the level-toned melodies, the duration of vowels for these two speakers does not provide any clear evidence of stress. Neither tone melody nor position appears to be a predictor of vowel duration in level-toned words.

The data on contour tone melodies show more agreement between speakers. There is still no difference in amplitude, as can be seen in Table 3. Figure 19 and Figure 20 show the amplitude data by speaker. For all three contour tone melodies included in this data set, the mean duration of the second vowel is longer than the mean duration of the first vowel. These data are shown in Table 4. For the L-M tone melody the difference is negligible for S3, and only 7ms for S4, with large variability in the data. For the L-H and H-L tone melodies, however, the mean difference is much larger: at least 32 milliseconds for the L-H tone melody and at least 16 milliseconds for the H-L tone melody. This difference is large enough that it is probably not random. Figure 21 and Figure 22 provide a visual representation of each speaker's data.

Table 3

Mean Amplitude (dB) of Vowels in Bisyllabic Words with Contour Tone Melodies

	V1 Amplitude		**V2 Amplitude**		**V2-V1 Amplitude**		**# of Sentences**
	M	*SD*	*M*	*SD*	*M*	*SD*	
S3							
L-H	49.4	6.1	49.7	6	0.2	2.5	11
H-L	54	2.9	53.6	2.5	-0.5	1.5	8
L-M	53	5.6	52.7	5.1	-0.2	2	16
S4							
L-H	45.8	4.5	48	4.4	2.2	1.9	11
H-L	45	4	43.9	3.5	-1.1	1.8	9
L-M	45.2	4.2	43.3	3.3	-1.9	2.4	11

Note. M=mean, SD=standard deviation. Means and standard deviations are calculated using raw values.

Table 4

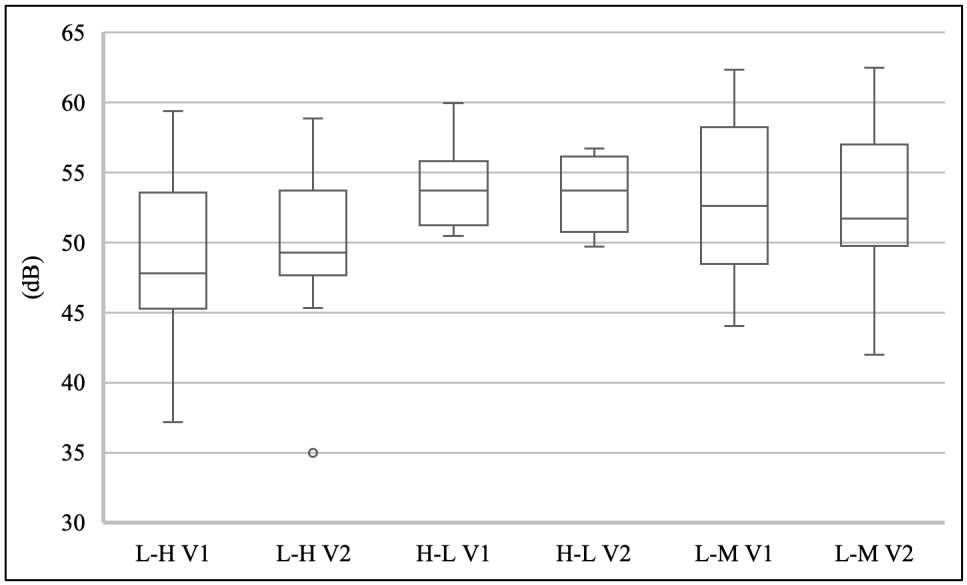
Mean Duration of Vowels (ms) in Bisyllabic Words with Contour Tone Melodies

	V1 Duration		**V2 Duration**		**V2-V1 Duration**		**# of Sentences**
	M	*SD*	*M*	*SD*	*M*	*SD*	
S3							
L-H	55	11	87	14	32	21	11
H-L	56	8	72	13	16	9	8
L-M	72	17	74	11	2	22	16
S4							
L-H	59	13	95	14	36	18	11
H-L	62	9	88	16	26	15	9
L-M	82	24	89	21	7	23	11

Note. M=mean, SD=standard deviation. Means and standard deviations are calculated using raw values.

Figure 19

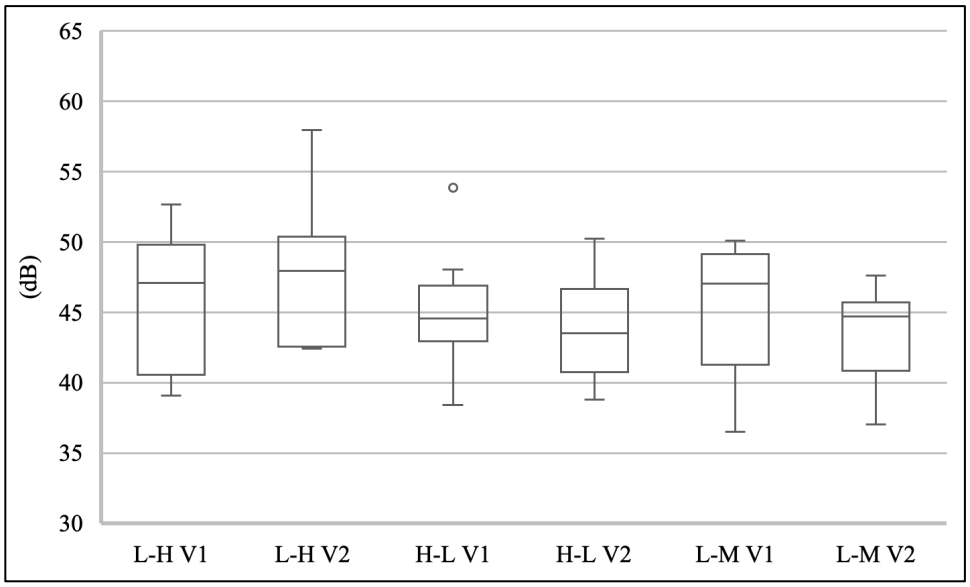
Amplitude of V1 and V2 in contour-toned words (S3)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

Figure 20

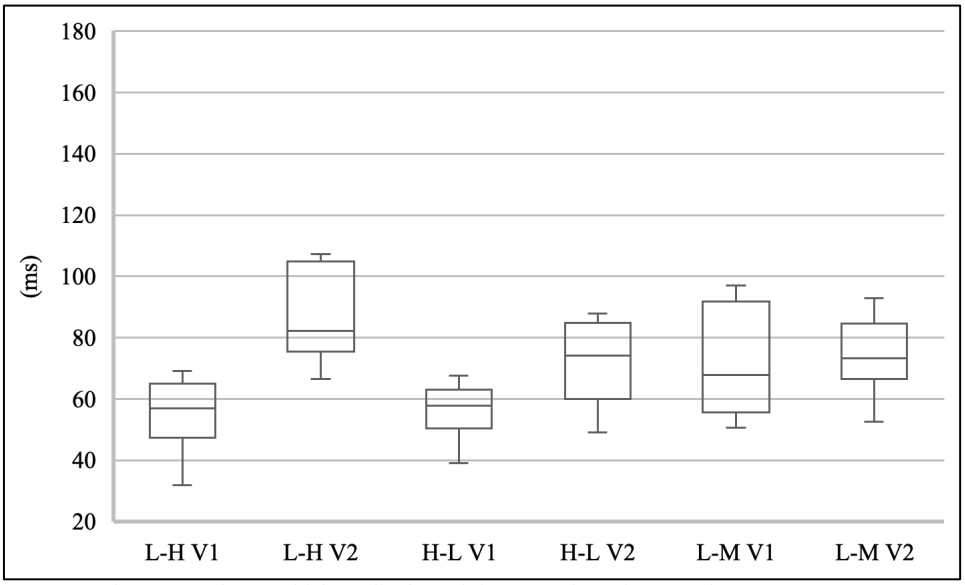
Amplitude of V1 and V2 in contour-toned words (S4)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

Figure 21

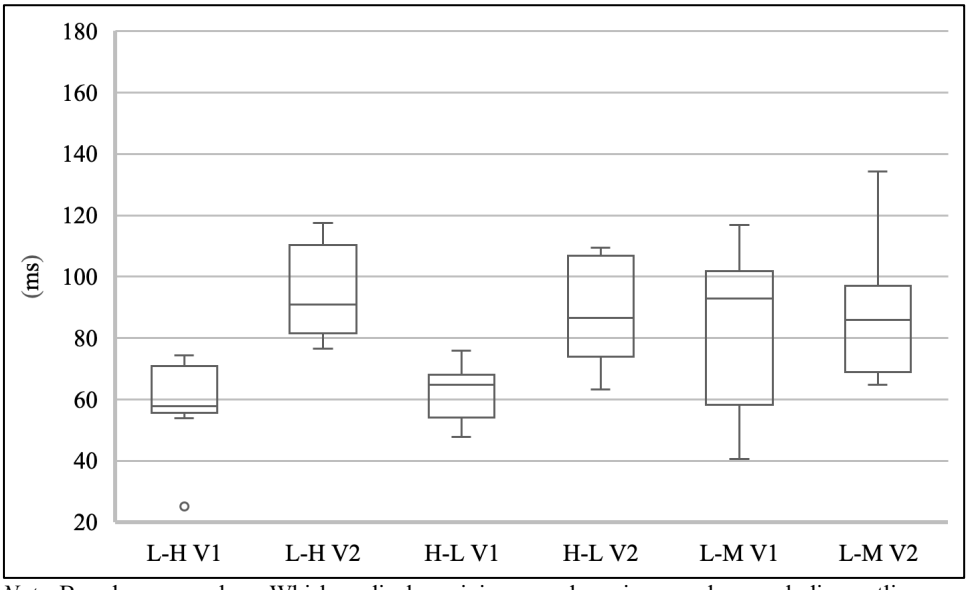
Duration of V1 and V2 in contour-toned words (S3)



Note. Based on raw values. Whiskers display minimum and maximum values.

Figure 22

Duration of V1 and V2 in contour-toned words (S4)



Note. Based on raw values. Whiskers display minimum and maximum values excluding outliers.

An issue with interpreting this data is that since the words had to be controlled for both lexical tone melody and vowel, the number of possible target words was limited. This means that

this data includes multiple repetitions of a limited number of target words, rather than a variety of words. Additionally, it was not possible to completely control onset consonants. This means that a possible cause of the durational differences between tone melodies is not tone melody, but an unrelated property of the specific words used. There are not enough target words of each tone melody to balance out any such effect. One piece of evidence in favor of this interpretation is that the vowel reduction processes discussed in the following sections do not depend on the tone melody of the word. Thus, the large difference in duration between the first and second vowels in the L-H and H-L toned words cannot be interpreted as evidence of stress.

One thing that is clear is that there is no phonetic evidence that stress is associated with tone or tone melody in the manner described by Morse (1976). According to Morse, the first H or M tone of a word is stressed. While the words with L-H tone melody appear to follow this prediction, since the mean duration of their second vowel is indeed longer, the words with H-L tone melody contradict it; their second vowel is also longer. Speaker S3 also produces M-M and H-H words with a longer second vowel.

If there were a tendency for the second syllable to have longer duration than the first in bisyllabic words, this would be consistent with phonological evidence for stress on the second syllable, discussed below. However, only L-H, and H-L toned words showed a clear difference in duration for both speakers.

3.4 The Prosodic Foot

The foot is the prosodic constituent that falls between the syllable and the word on the prosodic hierarchy. The foot is frequently the domain of lexical prosodic phenomena, which in turn can interact with phrase-level prosodic phenomena—meaning that it is an important part of a prosodic description. Evidence for the foot is generally that it is the domain of some phonological

phenomenon, such as stress assignment. In English and other stress languages, stress is determined by foot structure: how syllables are parsed into feet and the relationship between strong and weak syllables in a foot (Liberman, 1975; Nespor & Vogel, 1986; Hayes, 1995; Féry, 2017). By far, the majority of literature on the foot concerns stress, but the foot is can be the domain of other phenomena as well, such as phonotactic constraints (Bennett, 2012). In some tone languages, foot structure can condition tone processes such as tone sandhi (e.g. Duanmu, 1999 on Mandarin and Shanghainese). Neither Morse (1976), Le Bris and Prost (1981), nor Sanou (1993) address the prosodic foot in Bobo. Evidence for the foot has been found in other Mande languages, however. For example, in colloquial Bamana the prosodic foot has been proposed as the domain of high vowel and velar consonant elision (Green, 2010; Green et al., 2012) and as the domain of lexical tone melodies (Leben, 2001; 2003). Leben calls feet that are the domain of lexical tone melodies "tonal feet," but this does not mean that they have no metrical properties. In colloquial Bamana, iambic feet are avoided; elision is blocked when it would give rise to an iambic foot (Green, 2010; Green et al., 2012).

In this section I present three types of evidence for the prosodic foot in Bobo: the distribution of the vowel /ə/, the distribution of reduced vowels, and the distribution of lexical tone melodies. Each of these provides evidence of a phonological domain the size of a (preferentially) binary foot in which the second syllable is more prominent.

3.4.1 Distribution of /ə/

Bobo has eight oral vowels, /a, i, u, e, o, ɛ, ɔ, ə/. The vowel /ə/ has properties that set it apart from the other vowels. It has an extremely short duration and is frequently completely elided. It cannot be phonemically nasalized, but it can carry tone. Importantly, it only appears in a specific environment: in the initial syllable of a bisyllabic or trisyllabic morpheme. Some examples are

shown below:

(18) *pə³la³* ‘two’
tə¹ma³ ‘to get up’
mə¹la³ ‘to desire’
sə⁵la⁵lo⁵ ‘small’

It is questionable whether /ə/ is in fact a separate vowel phoneme. Morse (1976) raises the question of whether [ə] is a reduced form of one of /i, ɛ, u/, since these vowels can be reduced or elided in similar contexts (discussed in the next section). The evidence points in different directions. When asked to produce a word with an initial [ə] slowly, speakers prefer a monosyllabic pronunciation, suggesting [ə] is excrescent, or vary between [i] and [u], suggesting that it is a reduced form of /i/ or /u/. On the other hand, there is also evidence that /ə/ is distinct from /i/ and /u/: [ə] cannot be derived from /i, u/ by phonological rule. If one posits that [ə] is a weakened form of /i, u/, its distribution cannot be predicted and it also cannot explained by free variation. While a posited /ə/ phoneme can sometimes be pronounced as [i] or [u], there are some words in which reduced /i/ and /u/ cannot be pronounced as [ə]. For example, */si³bɛ³/* ‘be able to’ can be pronounced as [sĭ³bɛ³] or [sbe³] but not *[sə³be³], and */bu³na³/* ‘grave’, which can be pronounced as [bŭ³na³] or [bna³] but not *[bə³na³]. These words cannot be distinguished from those in which [ə] is licit by their segmental or tonal content. Phonetically distinguishing [ə] from a reduced /ɛ/ is difficult in natural speech, but if speakers are asked to sound out a word with a reduced initial vowel slowly, they will reconstruct reduced /ɛ/ distinctly as [ɛ].

Whether [ə] is a non-phonemic vowel that is inserted to break up consonant clusters is another thorny issue. The distributional evidence points to it being phonologically distinct from the other vowel phonemes, but it could be distinct because there is simply no phonological vowel there. Instances of [ə] are due to the reduction of the first vowel; the question is whether the historical vowel has been completely lost. Speakers do not consistently interpret words containing

[ə] as either monosyllabic or bisyllabic, suggesting that the correct analysis is somewhere in-between (e.g. that the system is currently evolving or that these words are sesquisyllabic). I have chosen an analysis in which there is an underlying vowel because /ə/ can carry tone, but this aspect of the phonology needs additional research.

Regardless of its underlying representation, the distribution of /ə/ cannot be explained with reference to the segmental context or syllable structure. Instead, any explanation must make reference to the initial position within a phonological domain that is larger than a syllable.[14] Two possibilities for this domain are the prosodic foot and the prosodic word. However, compound words containing /ə/ word-medially provide evidence against the prosodic word as the domain. An example is *ko¹-pə³ra³* 'seven' (from *koo¹* 'five' + *pə³la³* 'two'). If /ə/ is restricted to the initial syllable of the prosodic word, this should not be a possible form. An analysis in which /ə/ is restricted to the initial syllable of a prosodic foot is more promising. A plausible footing of this word is *(ko¹)(pə³ra³)*, which places /ə/ in the initial syllable of the second foot. This footing does require the additional proposal that prosodic feet do not cross morpheme boundaries, but this is a cross-linguistically common constraint (see e.g. McCarthy & Prince, 1993 for discussion).

A binary foot is the most likely domain that explains the restricted distribution of /ə/. Unary feet do not make any explanatory contribution here; in that case /ə/ could appear in any position of the word in that case—or in no position, if /ə/ is only licensed in a weak syllable. There is no data on the distribution of /ə/ that rules out a trinary foot, but a trinary feet do not account for the distribution of tone melodies (discussed in §3.4.3).

[14] If the analysis is that all of these words have become fully monosyllabic, then this prosodic domain is of course no longer larger than a syllable. In that case it would illustrate a historical tendency toward iambic feet in Bobo.

3.4.2 Distribution of reduced vowels

The vowels /i, u, ɛ/ are often reduced or altogether elided. As with /ə/, the reduction or elision of these vowels is limited to the initial syllable of a bisyllabic or trisyllabic morpheme.[15] Some examples are given in Table 5. Note that while /u/ can be elided in *su⁵ba⁵* 'okra', it cannot be elided in *ka³su³ba³* 'sesame'. This is a clear example of elision being conditioned by position, rather than by the segmental context. Examples like *ka³su³ba³,* in which a word-medial /i, u ɛ/ does not elide, are also evidence that words are footed from the left; otherwise the footing would be (*ka³)(su³ba³),* and elision would be permitted.

Table 5

Elision of /i, u, ɛ/

si³bɛ³	[sbɛ³]	'be able'	*[sib³]
fɛ¹rɛ¹	[frɛ¹]	'things'	* [fɛr¹]
su⁵ba⁵	[sba⁵]	'okra'	* [sub⁵]
ka³su³ba³	[kasu³ba³]	'sesame'	* [ka³sba³]

The vowel /ə/ and the reduced vowels share an important property, which is that they are in some sense "weak": They are shorter and more prone to elision than other vowels. Thus, while the phonetic data on duration and amplitude discussed in section §3.3 did not demonstrate stress, there are phonological processes that appear to be conditioned by stress.

It is also worth noting that Bamana and Bobo are not the only Mande languages with vowel reduction processes that operate on the first syllable of a binary foot. Seenku, a Mande language that neighbors Bobo, is another example. Vowel reduction in this language is more extreme. Most

[15] The context is not quite the same. Vowel elision does not occur in bisyllabic words that contain identical high vowels (i.e. in CuCu and CiCi). It is also possible, but not certain, that the consonantal context also differs. All examples of /ə/ are followed by a sonorant, while /i, ɛ, u/ can also be reduced in the context s_bV. Sonorants are more common than obstruents in root-medial position, meaning that the lack of examples of /ə/ before medial obstruents could easily be due to chance. A full analysis of the segmental contexts that license a reduced vowel has not yet been done.

of its bisyllabic vocabulary has been restructured into forms that are either monosyllabic or sesquisyllabic (i.e. consisting of a minor syllable followed by a full syllable). The nucleus of the minor syllable is always a schwa. Whether the schwa is part of the underlying representation or is an excrescent vowel that breaks up certain consonant clusters is unclear (McPherson, 2017; 2020). Bobo still has a large amount of clearly bisyllabic vocabulary, like Bamana, so its vowel reduction processes have not progressed nearly as far as in Seenku, but both are a part of a trend within some Mande languages towards an iambic prosodic foot structure.

3.4.3 Distribution of tone melodies

Earlier, I defined a tone melody as the sequences of tones that belong to a morpheme or word. In this section, I propose that the domain of the tone melodies is actually the prosodic foot and that, like Bamana, the prosodic foot in Bobo has both metrical and tonal properties. Under this analysis, the tone melody is a property of the foot, while the individual tones that make up the melody are assigned to the syllables within the foot. First, some facts about tone melodies follow. As mentioned previously, tone melodies often behave as a unit within the grammar. The tone melodies of some word classes are restricted. For example, verbs can only have an L, M, or LM tone melody. Additionally, grammatical tone processes often involve the replacement of the entire tone melody. For example, transitive verbs with an M tone melody have an L tone melody in the simple present and past perfective:

(19)	*ma³ra³*	*sɔ⁵gɔ⁵*	*ka¹ga¹*	*ma¹ra¹.*
	‘to injure’	panther	guinea.hen	injure
		“A panther injured a guinea hen.”		

(20)	*wɛ³bɛ³*	*ma³*	*su¹³*	*wɛ¹bɛ¹.*
	‘to break smth.’	1.sg	pot	break
		“I broke the pot.”		

There is a clear relationship between the inventory of contour tones and the inventory of

productive tone melodies in bisyllabic words. Table 6 provides an inventory of productive tone melodies in monosyllabic and bisyllabic words with examples. Note that with the exception of L-HL, the inventory of productive tone melodies is identical to the inventory of contour tones (LM, LH, ML, HL). There is no HM or MH contour tone, and also no H-M or M-H tone melody in bisyllabic words.

Table 6

Productive Tone Melodies Found in Monosyllabic and Bisyllabic Words

	Monosyllabic		**Bisyllabic**	
L	*ku*^{1}	'debt'	*dẽ*^{1}*mẽ*^{1}	'beans'
M	*ku*^{3}	'turtle'	*dẽ*^{3}*mẽ*^{3}	'bat'
H	*ku*^{5}	'armpit'	*dẽ*^{5}*mẽ*^{5}	'lice'
LM	*sõ*^{13}	'sing'	*ni*^{1}*mi*^{3}	'dance'
LH	*sõ*^{15}	'man'	*ba*^{1}*lo*^{5}	'iron'
ML	*ma*^{31}	'friend'	*fɔ*^{3}*rɔ*^{1}	'good'
HL	*dũ*^{51}	'bridge'	*ba*^{5}*se*^{1}	'mead'
LHL	--		*da*^{1}*lo*^{51}	'girl'

A restricted set of tone melodies in bisyllabic words implies that tone melodies are assigned to a domain larger than the syllable; otherwise, all logically possible combinations of tones should be possible. Additionally, some monosyllabic words provide evidence for this larger domain. Some monosyllabic words with a contour tone have bisyllabic plurals. In the plural form the tone melody of the singular is preserved, as in as in *kã*^{13} 'foot' and *kã*^{1}*ẽ*^{3} 'feet', pu^{13} 'younger sibling' and *pi*^{1}*a*^{3}*-re*^{1} 'younger sibling'[16] or *mũ*^{1} 'rabbit' and *ma*^{1}*ga*^{1} 'rabbits'. This indicates that while the entire tone melody surface on a single syllable in the singular form, it is because this is the only syllable available. As Morse (1976) notes: "If tone occurs on a higher level than the syllable (such

[16] The *-re*^{1} is a redundant plural suffix. Since contour tones are the least common type of tone and monosyllabic words with monomorphemic, bisyllabic plural forms are rare, there are not many examples that illustrate how the individual tones in a contour tone are redistributed in bisyllabic forms. In *pi*^{1}*a*^{3}*-re*^{1}, the redistribution is clear because the plural suffix *-re*^{1} is L-toned, not ML-toned (that is, the M tone must be part of the stem).

as the morpheme or the phonological word), the pattern or contour itself could be the significant unit, making a glide on a one-syllable word like Low-High (dǎ) equivalent to Low-High on a two-syllable word (dà-ló).” (p.87) When a second syllable is available, the second tone within a contour tone melody spreads to the second syllable, indicating the existence of a larger domain.

The remaining question is what this domain or higher level is. Morse (1976) suggests that the domain is the morpheme or phonological (prosodic) word. The prosodic word as the domain does not account for the distributions of the tone melodies in compounds. The components of a compound either retain their original tone melody, as in ko^{1}-pə^{3}ra^{3} ‘seven’ (< *ko.^{1}* ‘five’ + *pə^{3}la^{3}* ‘two’), or, frequently, have their original tone melody overwritten so that the second element has a higher tone than the first, as in *ba^{1}o^{1}-sio^{5}* ‘bicycle’ (< *ba^{1}lo^{5}* ‘iron’ + sio^{1} ‘horse’). In the latter case, the original tone melody of each root is replaced, rather than the compound tone pattern being assigned to the compound word in a way that is agnostic towards its individual elements. For example, there are no forms like **ba^{1}o^{5}-sio^{5}* ‘bicycle’, where the first high tone occurs in the first element. This can also be seen in compound words in which the elements have different numbers of syllables, e.g. *ne^{1}-fɔ^{3}gɔ^{3}* ‘néré powder’ (<*ne^{3}* ‘néré + *fɔ^{3}gɔ^{3}* ‘powder, flour’). The melody of a compound word is composed of melodies that are assigned to its individual components, which should not be the case if the tone melody is a property of the prosodic word.

The question of whether the domain of tone melodies is the morpheme or the foot is more difficult to answer, because the vast majority of morphemes in Bobo are either mono- or bi-syllabic. If one posits that foot boundaries do not cross morpheme boundaries, then in the vast majority of cases whether the domain is the foot or the morpheme cannot be determined because the analyses make the same prediction. However, there are some trisyllabic words that contain only one morpheme, and for these words whether the domain is the foot or the morpheme does matter.

These words are often loans, but have nonetheless been adapted into Bobo's phonological system, e.g. *si³bi³ri⁵* 'Saturday' (< Jula *sibiri*), *ta³ba³le⁵* 'table' (< French *table)* and *fɛ¹nɛ¹ti⁵ri⁵* 'window' (< French *fenêtre*). There are also some words that might not have been monomorphemic historically, such as *sa¹mo³ro⁵¹ 'foreigner,'* but which are treated as monomorphemic now.[17] Although there are not many examples of monomorphemic trisyllabic words in Bobo, an analysis that takes them into account is preferable to one that does not. There are two facts about these words that suggests the domain of the tone melody is a binary foot, and not the morpheme:

First, there is a tendency for the first two tones in a monomorphemic trisyllabic word to be identical, as in the first three examples cited above and as in additional *va¹ga¹ka⁵¹* 'corn', ba⁵ra⁵ka¹ 'thank you,' *da¹ba¹le³* 'mirror,' *si⁵mi⁵ma¹* 'dry season,' and others.[18] This is what is expected if tone melodies are assigned to the foot rather than to the morpheme. The level tone melodies are the most frequent; Morse (1976) found that in her corpus, 54% of bisyllabic words had a level tone melody. Thus, if the first two syllables of a trisyllabic word form a binary foot, identical tones in the first two syllables of a binary foot should be strongly represented.

The second fact is that monomorphemic trisyllabic words can have more varied tone patterns than bisyllabic monomorphemic words. If tone melody was assigned to the morpheme,

---

[17] Compare *sa¹mo³oro⁵¹ 'foreigner'* (pl. *sa¹ma³ra⁵¹*) with *da¹lo⁵¹* 'girl' (pl. *da¹la⁵¹*) and *ya¹lo⁵¹* 'bird, guinea hen' (pl. *ya¹la⁵¹*). An alternation between /ɾ/ and /l/ can be seen elsewhere in the lexicon, as with *pə³la³* 'two' and *ko¹pə³ra³* 'seven'. Neither *sa¹mo³* nor *ya¹* is recognized as having an independent meaning, although dã¹ is found as an element in some compounds. New words cannot be derived through the addition of *-ro⁵¹/ lo⁵¹*. The final syllable of *sə⁵la⁵lo⁵* 'small' is likely related to this suffix.

[18] The question of how loanwords are adapted and how much of the tone pattern is based on the prosodic properties of the source word is complicated. However, in the case of loans or potential loans from Jula, the tones in Bobo do not correspond well to the original tones in Jula. In the case of French loans, the addition of epenthetic vowels means that the correspondence between tones and stressed syllables in French is not easy to evaluate. A straightforward mapping between the prosody of the source language to tones in Bobo does not exist, so it cannot explain the tone patterns in these loans.

this should not be the case, as it would be the number of morphemes and not the number of syllables that determined the number of tone melodies present in a word. Tone patterns such as L-M-L (e.g. *a¹ra³ba¹* 'Wednesday'), L-H-M (e.g. *ɲo¹ŋo⁵ma³* 'camel') can be found in monomorphemic trisyllabic word—but not bisyllabic ones. This cannot be explained due to a restriction against contour tones in bisyllabic words, which do exist in words with an L-HL melody. Importantly, these more varied tone patterns can involve tone sequences that are not found in bisyllabic words, namely M-H and H-M. These are not valid tone melodies; they do not appear in bisyllabic words. In trisyllabic words, they only appear when they would cross a foot boundary, as in *ta³ba³le⁵* ''table': (ta³ba³)(le⁵). The tone patterns found in monomorphemic trisyllabic words can always be decomposed into sequences of valid tone melodies in this way. Together, the more varied tone sequences allowed on monomorphemic trisyllabic words, as well as the fact that they can always be decomposed in this way, provide evidence that tone melodies are associated with the foot. They also provide evidence that footing is exhaustive, as the final syllable is assigned its own melody instead of remaining toneless or copying a tone from the previous syllable. There are no toneless syllables in Bobo, and no other evidence of unfooted syllables.

3.4.4 Summary of prosodic foot structure

In the previous sections I proposed that the prosodic foot is the domain of lexical stress and tone melody in Bobo. This accounts for the distribution of weak vowels and for the distribution of lexical tones. The following are the properties of the prosodic foot that are needed to account for the data:

- Feet are preferentially binary. If there was no restriction against unary feet, then monomorphemic words could potentially have as many prosodic feet as they have syllables. This could not account for the distribution of weak vowels, even with the

additional stipulation that unary feet cannot contain a weak vowel. A footing such as (σ́)(σ.σ́) would be possible, allowing a weak vowel on the second syllable of a trisyllabic word—which never occurs. Unary feet would additionally allow monomorphemic words to have any logical possible sequence of tones, but instead, there are restrictions on sequences of tones that are predicted by a binary foot. A trinary foot also would not account the more complex tone sequences in trisyllabic words or the cases in which an M-H or H-M tone sequence is allowable.

- Feet are iambic. This predicts the position of weak vowels within the binary foot.
- Words are footed from the left. This is only relevant in trisyllabic words, where footing from the right predicts the wrong distribution of weak vowels and tone sequences.
- Foot boundaries do not cross morpheme boundaries. This predicts that weak vowels can occur in a word-medial syllable if that syllable is the first syllable in a morpheme and allows for compound words such as *(ko¹-)(pə³ra³)* 'seven'.
- Words are exhaustively footed. Words with an odd number of syllables do not have toneless syllables or copy tones from the previous foot; instead, a foot can be unary and be assigned a tone melody if a single syllable is all that is available.

3.5 Summary of Word Prosodic Structure

In this chapter I have argued that the lexical tone inventory of Bobo is L, M, H, HL, ML, LM, HL. This provides additional support for including the M-L and HL tones into the tone inventory of Bobo, about which previous sources (Morse, 1976; Le Bris & Prost, 1981; Sanou, 1993) disagree. The existence of ML and HL tones, in turn, means that there is a near equivalence between the inventory of contour tones and the inventory of tone melodies found on monomorphemic bisyllabic words (with the exception being L-HL, which is not found on

monosyllabic words). This relationship was noted by Morse but was not examined in terms of prosodic foot structure. My own analysis is that the domain of tone melodies is the foot, which is preferentially binary and iambic. This contradicts Morse (1976), for whom the domain of tone melody is the morpheme or word and for whom stress is assigned according to tone melody rather than position within a foot. I found no clear phonetic evidence of stress; neither duration nor amplitude of vowels in bisyllabic words patterns according to tone melody or position. The evidence of iambic foot structure presented here is the distribution of weak vowels; /ə/ and the reduced forms of /i, u, ɛ/ only occur in the first syllable of the foot. Iambic foot structures have been described for other Mande languages, including Bamana (Green 2010; Green et al. 2012) and Bobo's close neighbor, Seenku (McPherson, 2017; 2020). The proposal that tone melodies are assigned to the foot is similar to Leben's proposal for tonal feet in Bamana (Leben, 2001; 2003), although Bamana has a much smaller inventory of melodies (only H and LH).

Chapter 4

Phrase-Level Prosody

4.1 Overview

This chapter is an investigation into the prosodic phrase structure of Bobo. Prosodic phrasing is the grouping of words into hierarchically organized phonological domains that are defined by various phonetic properties, such as changes in F0, duration, and pauses (Ladd, 2008; Féry 2017). Different types of prosodic phrases correspond to different levels of the prosodic hierarchy.[19] For example, following Beckman and Pierrehumbert (1986), a standard analysis of English is that it has an intonational phrase (IP), which contains one or more intermediate phrases (ip). The IP is defined by the presence of a final boundary tone and by final lengthening. The ip is defined by the presence of a final phrase accent, final lengthening, and by being the domain of downstep. (These properties will be discussed in more detail in §4.2.) The levels of the prosodic hierarchy and the properties that define them are language specific. Additionally, it is not possible to predict the phrase-level prosody of language from its word-level prosody (Jun, 2005b). This means that when investigating the prosodic structure of a new language, only a limited number of

[19] If the analysis includes recursive prosodic structure, such that a prosodic phrase can contain another prosodic phrase of its same type, then the correspondence between prosodic phrase type and level of the prosodic hierarchy is not necessarily one-to-one for any given utterance. Whether recursive prosodic structure is possible is a subject of ongoing discussion; the hypothesis that it is not recursive is the 'strict-layer hypothesis' (Selkirk 1986; Nespor & Vogel 1986). For some arguments against the strict-layer hypothesis, see e.g. Hyman et al. 1987; Wagner 2005; Ladd 2008 p. 290-299. Independently of which analysis ultimately turns out to be correct, the statement that smaller phrase types are nested within larger phrase types is true; an IP will never be nested inside an ip, for example.

assumptions can be made—for example, that all languages have some form of prosodic phrasing (Ladd, 2008). Instead of looking for specific prosodic phenomena that define a specific type of phrase, the identification of phrase types is based on general criteria: Each phrase type will have a unique phonetic or phonological property that distinguishes it from other phrase types (Beckman et al., 2005).

There is not much previous description of phrase-level prosody in Bobo. Le Bris and Prost (1981) and Sanou (1993) do not address phrase-level prosody. Morse (1976) describes it briefly. According to Morse, Bobo has two types of prosodic phrases: the pause group and the utterance (or sentence) group. The pause group is the smallest phrase. It is delineated by a pause, which Morse defines as "a slight pause or break in the flow of speech" (p. 96.). She does not describe the pause group as having any other characteristics. The utterance is the largest phrase. It is the domain of "automatic phonetic downdrift" and has one of four terminal contours: (a) the period contour[20], characterized by a terminal glottal stop; (b) the question contour, characterized by the absence of the terminal glottal stop and slight lengthening of the final vowel; (c) the negative contour, characterized by a higher pitch level throughout the utterance, the presence of the phrase-final negative marker *=ga*^{5}*/ŋa*^{5}, and possibly the addition of an H tone to the subject; and (d) the exclamation contour, characterized by an extreme raising of pitch and an increase in speaking rate. There is little additional detail that is not included here, since this section is very short. Her analysis does not correspond straightforwardly to later theoretical frameworks on prosody and the lack of detail makes it difficult to connect them. For example, Morse does not distinguish between a

[20] I have kept Morse's original terminology when summarizing her work. According to Morse, the period contour "indicates the absence of any strong emotional state and is used to mark simple declarative statements and narration." It is the default intonational contour. It is not clear whether it could also be used for utterance types she does not discuss, for example imperatives.

speaker's emotional affect and the prosodic structure of an utterance, which can be seen in her description of the exclamation contour; both speech rate and emotional pitch variation are typically not considered to be part of linguistic prosody. This raises the question of whether the raised pitch of a negative utterance is due to the speaker's emotional affect, rather than underlying linguistic structure.

The approach in this chapter is informed by a preliminary investigation of the phrase-level prosody of Bobo that I completed in 2014 (Sherwood, 2016). The investigation was narrower in scope than Morse (1976), but where it overlapped the results were frequently inconsistent. I found that questions (both polar and *wh*-) are distinguished from statements by lengthening, breathy phonation, and a raised F0 at the end of a phrase, none of which are mentioned by Morse. The design of the elicitation sentences that were collected limited the possible analyses, however. In particular, since most of the elicitation sentences ended in a grammatical L tone, it was not possible to determine whether there was a phrase-final L% boundary tone.[21] Additionally, only two speakers participated, and further work with the language suggested that Bobo intonation is much more variable than this small data set could capture. This data is described in §2.2.1 and is revisited here.

The specific questions that guide the investigation that follows are: (a) What are the phonetic and phonological properties of intonational phrases? (b) How is prosody used to mark

[21] The reason for this design was that when I was designing the sentences I did not yet know enough about grammatical tone to reliably create elicitation sentences that ended in other tones. Since Bobo is S-O-V-X, and transitive verbs are almost always L-toned, simple transitive sentences will end in an L tone unless there is an X constituent. At the time I wanted to avoid using an X constituent because I did not know whether the possible X constituents involved grammatical tone alternations or whether they would be set apart in their own prosodic phrase, which would complicate the analysis. I did not use intransitive sentences because creating intransitive sentences more than a few syllables long involves grammatical constructions that have grammatical tone alternations, which conflicted with criterion of using mostly lexical M tones.

utterance type? (c) Can other phrase types be identified, and if so, what are their properties?

4.2 Research Questions

4.2.1 What are the properties of intonational phrases?

All languages have prosodic phrasing and therefore all languages have at least one type of prosodic phrase. Under the Autosegmental-Metrical (AM) model of intonational phonology (e.g. Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Pierrehumbert & Hirschberg, 1990; Ladd, 2008), the intonational phrase (IP) is the highest level of the hierarchy that can be consistently identified as a distinct phonological category.[22] All languages are assumed to have an IP, although the phonological properties that encode the IP can vary. Therefore, I assume that Bobo has an IP, with properties that have yet to be determined. The data collection was designed with cross-linguistically common properties in mind, but any phenomena that appear to be associated with the intonational phrase were considered relevant during the analysis.

Since Bobo is a lexical tone language, intonational properties such as boundary tone, phrase accents, and downtrends are of particular interest. Phrase accents are associated with an ip boundary, while boundary tones are associated with an IP boundary. A final boundary tone

[22] Higher levels of the hierarchy have been proposed in various frameworks, the utterance being the most notable (see Shattuck-Hufnagel & Turk 1996 for an overview of different proposals). However, the evidence provided for the utterance as a separate phonological category has been gradient rather than categorical. Phonetic properties associated with the utterance level are also associated with lower levels of the hierarchy and differ only in the strength of their expression. For example, utterance boundaries might be associated with an L% that is phonetically realized with a lower F0 than the L% associated with the boundary of an IP (e.g. Lovick & Tuttle 2012) or by nested declination or downstep trends (e.g. Beck & Bennett 2007). Since both L% and declination are properties of the IP in these languages, this data can also be explained as the result of nested IPs if the Strict Layer Hypothesis is not assumed. The utterance level has also been proposed as the domain of certain segmental processes (e.g. Nespor & Vogel 1989; Hayes 1989), but later experimental work has shown that these types of segmental phenomena often apply at multiple levels of the hierarchy and not just at the utterance level (e.g. Pelekanou & Arvaniti 2002; see also Wagner 2005 and Krivokapić & Byrd 2012 for a discussion of this question). This means that the utterance cannot be systematically distinguished from lower levels. The difficulty of finding reliable phonological criteria for establishing the utterance level means that it has not gained widespread acceptance among researchers using an AM framework (see discussion in Jun 2005b and Beckman et al. 2005).

(notated as T%) is an obligatory element of an IP in a variety of genetically and typologically diverse languages (e.g. Pierrehumbert, 1980 for English; Myers, 1996 for Chichewa; Bishop, 2002 for Bininj Gun-Wok; Holton, 2005 for Tanacross). However, languages vary in whether they have boundary tones and whether, if they have them, boundary tones are obligatory. A lack of obligatory boundary tone has been found for an increasing number of lexical tone languages, including several of West Africa. For example, in Bàsàá, a Bantu language of Cameroon, there might be no boundary tones at all (Makasso et al., 2016). A lack of boundary tone has also been observed for Kɔnni, a Gur language of Ghana (Cahill, 2016). In some other languages, boundary tones only occur in specific pragmatic or phonological contexts. In Akan, a Kwa language of Ghana, L% is used to mark polar questions, but declarative statements do not have boundary tones (Kügler, 2016).[23] In Kammu, an Austroasiatic language of Laos, an H% boundary tone is characteristic of IPs, but is neutralized when it would obscure a lexical HL tone contour (Karlsson et al., 2012). There is very little work on intonation in other Mande languages, but boundary tone has been described in Kakabe (Vydrina, 2017) and Guinean Maninka (Vydrin & Diané, 2016). Both of these languages have two tone levels, as opposed to Bobo's three.

In the absence of intonational tones it is possible for the F0 contour of an utterance to be determined by the string of lexical tones alone. However, it is more common that lexical tones will undergo tonal processes that affect their realization. These processes are frequently sensitive to the phrase-level prosodic structure. Downtrends are particularly relevant here since they are common in African tone languages. *Downtrend* is a cover term for processes that result in the lowering of F0 over the course of the phrase, namely declination and downstep (sometimes called downdrift).

[23] Declarative statements have final tone neutralization that results in the lowering of final H tones. Kügler (2017) does not analyze this as boundary tone because it only affects H tones.

Downtrends have been previously observed in Bobo; as mentioned above, Morse (1976) claims that the utterance group is the domain of "automatic phonetic downdrift. Definitions of these terms and theoretical perspectives on the phenomena that they describe vary considerably. The definitions that I use here follow Connell and Ladd (1990), which are more common (although not universal) within the Africanist literature.

Declination is a gradual decrease in pitch throughout a prosodic phrase. It is sometimes considered to be a universal or near-universal process (e.g. Ladd, 1984; Maddieson, 1996; Clark et al., 2007), although it is not necessarily present in every phrase. Declination can be suspended to mark sentence level contrasts such as utterance type (e.g. Lindau, 1986 on Hausa; Armstrong 2010, on Puerto Rican Spanish), or in specific constructions (e.g. Caron et al., 2015 on Zaar). There are two prominent competing views of declination. One view is that it is a phonetic process by which F0 targets are scaled to a gradually decreasing reference line, resulting in decreasing F0 throughout the domain (e.g. Connell & Ladd, 1990; Connell, 2001). This is the definition that I adopt here. A different view is that declination is largely—or entirely—the result of successive applications of downstep (see Ladd, 2008 for a discussion). This view is more common in the literature on intonation in non-lexical tone languages. In a lexical tone language, declination is more easily distinguished from downstep than in a non-lexical tone language due to the dense specification of tone since in a lexical tone language there is less interpolation, making the stepwise nature of downstep more apparent.

Downstep is a phonological process that lowers the F0 of an affected H tone. This is the general definition that I follow. Thus, repeated application of downstep within a domain results in the gradual lowering of F0. It is distinguished from declination in that downstep occurs after a phonological trigger, usually an L tone, and this is the definition of downstep that I follow. The

concept of downstep originated in the literature on African tone languages, but perspectives (and research focus) diverged when it was adopted by those working on intonation in non-tone languages. In the Africanist literature, downstep is typically divided into *automatic* downstep, which is triggered by an overt L tone preceding the affected H tone, and *non-automatic* downstep, which occurs without an overt L tone and is instead conditioned by specific words (i.e. it is lexically contrastive) or grammatical constructions. In many cases, non-automatic downstep has been argued to be the result of a floating L tone, but this is no longer thought to be universally the case (Connell, 2001; 2011). Whether downstep affects M tones, or can be triggered by M tones, is language specific (Connell, 2011). The typical case is that of an H tone being lowered after an L tone. Downstep is commonly analyzed as a change in pitch register triggered by an L tone (e.g. Hyman 1993, 2013; Downing & Rialland 2016b). *Downdrift*, the term that Morse uses, has variously referred to both declination and downstep. It is likely that she is referring to automatic downstep, but since she provides no further detail or examples this is not certain.[24]

In addition to intonational properties, final lengthening and non-modal phonation are also likely to be relevant. My earlier work (Sherwood, 2016) suggests that these properties vary by utterance type in Bobo. Final lengthening (the lengthening of segments phrase-finally) is found in diverse languages (as examples, see Wightman et al., 1992 and Turk & Shattuck-Hufnagel, 2007 on English; Hockey & Fagyal, 1998 and Gósy & Krepsz, 2018 on Hungarian; Xu & Wang, 2009 and Yang & Wang, 2002 on Mandarin; Rialland & Robert, 2001 on Wolof; Coulter, 1993 on ASL;

[24] Connell (2001) uses *downdrift* to refer to downstep in which there is no establishment of a new F0 ceiling. Under Connell's definition, downdrift allows for later H tones within the domain to reach a higher F0 than the downstepped H, whereas downstep does not. This is not the definition I use here and I do not have data that would distinguish between downstep and downdrift in this sense. Therefore I will focus on whether there exists a gradual, global decrease in F0 (declination) and/or a phonologically conditioned decrease in F0 regardless of the establishment of a new F0 ceiling (downstep).

and Fletcher, 2010, p. 541 for an extensive list). It is likely a universal property of phrase-level prosody (Maddieson et al., 1996; Fletcher, 2010). Phrase-final lengthening can be affected by language-specific factors, however. For example, in languages with phonemic vowel length, the degree of phrase-final lengthening can be influenced by the need to preserve lexical contrast (e.g. Nakai et al. 2009 and Nakai et al. 2012 on Finnish). As mentioned above, questions in Bobo appear to have a longer final segment than statements (Morse, 1976; Sherwood, 2016). However, phrase-final lengthening in statements has not been investigated yet.

Non-modal vowel phonation at final phrase boundaries is also found in diverse languages (Gordon & Ladefoged 2001; Epstein 2002; Kohler 2000; Féry 2017). The functions of non-modal vowel phonation and the degree to which it is grammaticalized vary. In some languages, non-modal vowel phonation is a social or stylistic variable that is used to convey information such as the speaker's identity, stance or mood, as it does in English (Henton & Bladon, 1988; Yuasa, 2010; Mendoza-Denton, 2011, Podesva, 2013). It can also be used to help organize discourse; for example, in Finnish, creaky voice is associated with the end of a speaker's turn (Ogden, 2004). On the other hand, in some languages it carries sentence-level linguistic contrast, placing it within the strict definition of linguistic prosody. I will focus on the question of whether this is true in Bobo, as assessing the more pragmatic and paralinguistic functions of non-modal voicing is outside the scope (or possibilities) of this project. My previous work (Sherwood, 2016) indicates that breathy phonation may a component of question prosody in Bobo. However, since non-modal vowel phonation has a range of paralinguistic and linguistic functions, additional data is needed to assess how regularly it occurs—i.e. whether it is an obligatory component of question marking or an optional one.

The properties of intonational phrases described above will be discussed in detail in the

results sections. Other properties are not specifically targeted in the design of the elicited data, but any prosodic characteristics that could not be explained by the lexical or grammatical content of the phrase were considered relevant (e.g. any pitch excursions that are not due to lexical or grammatical tone). No other properties were apparent in the data, whether due to the nature of the data or because they are not present in the language.

4.2.2 How is prosody used to mark utterance type?

It is sometimes stated that in a language with lexical tone, sentence-level contrasts such as utterance type are more likely to be marked by means other than intonation (e.g. Yip, 2002). However, it is clear that many tone languages *do* use intonation to mark utterance type, even while others do not.[25] Investigating the prosodic structure of different types in a variety of tone languages allows us to assess the extent to which they do use intonation and, if they do, the extent which they use the same intonational means. Additionally, utterance type is a salient sentence-level contrast that is relatively easy to elicit. If a language does use prosody to encode utterance type, including different utterance types in the investigation of its prosodic structure provides a wider view than an investigation that only includes declarative statements.

The most obvious and well-known example of using prosody to mark utterance type is the use of an H% boundary tone to mark polar questions in English and in many other languages. Using some form of rising intonation to mark questions is cross-linguistically common, although not universal (Ladd, 2008; Gussenhoven, 2004). Many languages of Africa are exceptions.

[25] Tone languages in which there is no intonational difference at all between utterance types appear to be less common than those where there are intonational differences. The clearest example of a tone language without intonational differences between utterance types is Mambila (Connell 2016), which is unusual in having four lexical tone levels. There are also outstanding questions about the extent to which intonational differences between utterance types are grammaticalized in some tone languages, for example Vietnamese (Brunelle et al. 2012).

Rialland (2009) notes that African languages can be divided into two categories based on how they use prosody to mark polar questions: those that use some form of high pitch, and those that do not. Those that do not are clustered in a band below the Sahelian belt that extends from Sierra Leone to Eastern Sudan. They share a constellation of features: (a) a falling pitch contour due to a final L tone or L% boundary tone, (b) a sentence-final low vowel, (c) a prolongation of the final vowel, and (d) a breathy utterance termination. The term "lax question prosody" has been useful to refer to question prosody of this type but might not be accurate for all of the languages that it is intended to describe. The term reflects a proposed reduction in laryngeal tension or speech effort at the end of the utterance, which results in properties like the lowering of phrase-final F0 and breathy termination. Subsequent acoustic work has shown that this is not always accurate. For example, Genzel (2018) presents an analysis of Akan question prosody that does not show greater breathiness or reduced speech effort, although it does have falling intonation and prolonged final vowels in polar questions and has been counted as a "lax question prosody" language for that reason. The languages that have these characteristics belong to different language families, including Gur, Kwa, and Mande, making this an areal rather than genetic phenomenon. Bobo is not included in the sample of languages discussed by Rialland, but it is spoken just within the boundaries of this area. As described in the previous section, it appears to have the prolonged final vowel and breathy utterance termination in questions (Sherwood, 2016). However, this research was based on a limited amount of data and it is clear that there is more variation in question prosody in Bobo than the data could capture.

4.2.3 Can other phrase types be identified?

There is considerable cross-linguistic variation in the levels of the prosodic hierarchy between the phonological word and the IP. Not all languages have an ip, and some languages have

additional prosodic phrases. For example, in this position, Mandarin has been claimed to have the tone sandhi group, which is the domain of tone sandhi (Peng et al., 2005); Korean has been claimed to have the accentual phrase, which is the domain of a specific accentual (intonational) melody (Jun, 1993; Jun, 1998; Jun, 2000); and Cantonese has been claimed to show no evidence of a level between the phonological word and the IP at all (Wong et al., 2005). There is also cross-linguistic variation at a smaller scale, e.g. between phrases that are defined by similar processes and have been given the same label. For example, both Korean and French have an accentual phrase that is defined by the presence of an accentual melody, but details differ. The accentual melody of French is /LHLH/, while in Korean it is either /LHLH/ or /HHLH/, depending on the initial consonant of the phrase (Jun & Fougeron, 2000). Additionally, the accentual melody is the domain of primary and secondary stress in French, while Korean has no stress (Jun & Fougeron, 2000; Jun, 1993).

An important point is that within the AM framework, it is the categorical phonological properties of a prosodic phrase that differentiates it from phrases of another type, and not the degree of phonetic disjuncture, e.g. the duration of final lengthening or a pause. This is because the relationship between degree of phonetic disjuncture and the type of prosodic phrase is not one-to-one. Speakers are sensitive to more distinctions in degree of phonetic disjuncture than there are prosodic phrase types (Krivokapić & Byrd, 2012), and the degree of phonetic disjuncture can be affected by factors other than the level of the hierarchy, such as the length and syntactic complexity of the phrase (Krivokapić 2007; Fletcher 2010). Using the degree of phonetic disjuncture is not possible unless speakers either produce or perceive these continuous variables in a categorical manner (for example as is the case for VOT). Establishing whether or not speakers of Bobo do perceive these variables categorically goes beyond the scope of this study. This study focuses on the presence or absence of phonological elements associated with the prosodic phrase. This

approach has the consequence that there must be at least one unique and obligatory phonological property of the IP if smaller phrase types are to be identified. Otherwise, any boundary can be an IP boundary.

4.3 Declarative Sentences

4.3.1 Analysis

This section examines the prosodic phrase structure of elicited declarative statements. The elicited context, while not "neutral," reduces the variation in pragmatic and discourse factors that could affect the prosody of an utterance. Any obligatory properties of IPs will be present. This establishes a baseline against which additional data can be compared. Unless otherwise stated, the analyses in this section are based on the declarative statements collected in the elicitation sets described in §2.2.2.2 and §2.2.2.3 (phrase prosody elicitation tasks 1 and 2). These elicitation tasks consist of matched declarative statements and questions (both polar and *wh*-questions in task 1, and polar questions only in task 2). Together, 381 tokens recorded as part of these tasks were included in the project and 161 of these are declarative statements. Eight tokens were excluded from the analysis of F0 due to breathy or creaky phonation obscuring the F0 contour of the final syllable of the target word, and six tokens were excluded from all analyses of duration because the phrase-final target word could not be segmented. The presence of an IP boundary was identified by the presence of a perceptible pause that is not due to a disfluency. This is a complex issue, as there are no comprehensive criteria for identifying pauses that are due to prosodic boundaries as opposed to pauses that are due to speech planning or production errors (Ferreira, 2007). However, pauses are strongly correlated with major phrase boundaries (Ladd, 2001; Cruttenden, 1997); the presence of a pause strongly indicates that there is an IP boundary. Although the converse is not true and the lack of a pause does not indicate the lack of a boundary, there are no better criteria for

identifying IP boundaries in a language where the properties of the IP are as yet unknown. Pauses were counted as a disfluency if they were adjacent to an identifiable error, such as the repetition of a word or speaker self-correction. This left zero well-formed pauses in these elicitation tasks. Therefore, all of the final IP boundaries examined were also sentence-final.

4.3.2 Boundary tone

The shape of the F0 contour of the final syllable was coded as falling, level, rising, or rising-falling based on an audiovisual inspection of the recording in Praat. The coding was based on the following principles for consistency: If the difference in F0 between the first glottal pulse after the onset consonant and the last glottal pulse where pitch could be reliably tracked was greater than 5Hz, the contour was coded as rising, falling, or rising-falling depending on the direction of the trajectory. If the difference was less than 5Hz it was coded as level, *unless* there was a clearly audible rise or fall in pitch that could not be accurately tracked in Praat due to non-modal phonation or unless the rise or fall began during the onset consonant (i.e. if there was good reason to believe the measurement was inaccurate). If it was at all difficult to determine whether there was an audible rise or fall in F0, the token was excluded.

At the end of an elicited declarative statement, the shape of the F0 contour is predictable from the lexical tone. An L tone is expected to surface as either a fall in pitch or a low level F0, an M tone is expected to surface as a (mostly) level F0, and an H tone is expected to surface as either a rising or high level F0, depending on the preceding lexical tone. As shown in Table 7, this is what occurs in the majority of the statements with words ending in L, M, or H tones. There is no change in F0 trajectory that would suggest the existence of a boundary tone that is realized after the string of lexical tones. Instead, the shape of the final F0 contour is a straightforward realization of the lexical tone.

Table 7

F0 Contour in the Final Syllable of Words Ending in L, M, or H Tone

	Total	Falling	Level	Rising	Rising-Falling
L	59	53	6	0	0
M	32	12	20	0	0
H	46	0	8	36	2

Three representative examples are shown in Figure 23-Figure 24, showing elicitation sentences that end with a final lexical L, M, and H tone respectively. In Figure 23, the final lexical L tone of *ba¹ra¹* 'tear' realizes as a fall in F0. The fall in F0 is gradual and starts at the beginning of the string of L tones, in the first *ja¹ra¹* 'women', which is typical for a string of consecutive L tones. In Figure 24, the final lexical M tone of *fã³nã³* 'also' realizes as a level F0, at the same level as the first M-toned syllable of the word. In Figure 25, the final lexical H tone of *mẽ⁵* 'also' is realized as a rising F0, since the H tone is rising from a previous L tone.

In twelve cases, the final lexical M tone realizes with a slight fall in F0. This is not an indication of a final L% boundary tone. The labeling scheme used above to categorize F0 contour of the final syllable of the phrase is simple and does not distinguish between different degrees of a fall or rise. A typical example of a final M tone with a fall in F0 is shown below in Figure 26. Compare this to Figure 27, in which there is a final L tone. The fall in F0 in Figure 26 is less than would be expected if there was an L (or L%) in the tonal string. It also occurs irregularly, suggesting that it is phonetic variability in the realization of tone rather than phonologically structured. Note that in Figure 26, a slight fall in F0 can also be seen in *kɛ³kɛ³rɛ³* 'goat', which is not at the final boundary of the IP. Although M tones are typically level, this type of variation is not unusual, regardless of the position within the IP.

Figure 23

Declarative statement ending with L tone (S3, M)

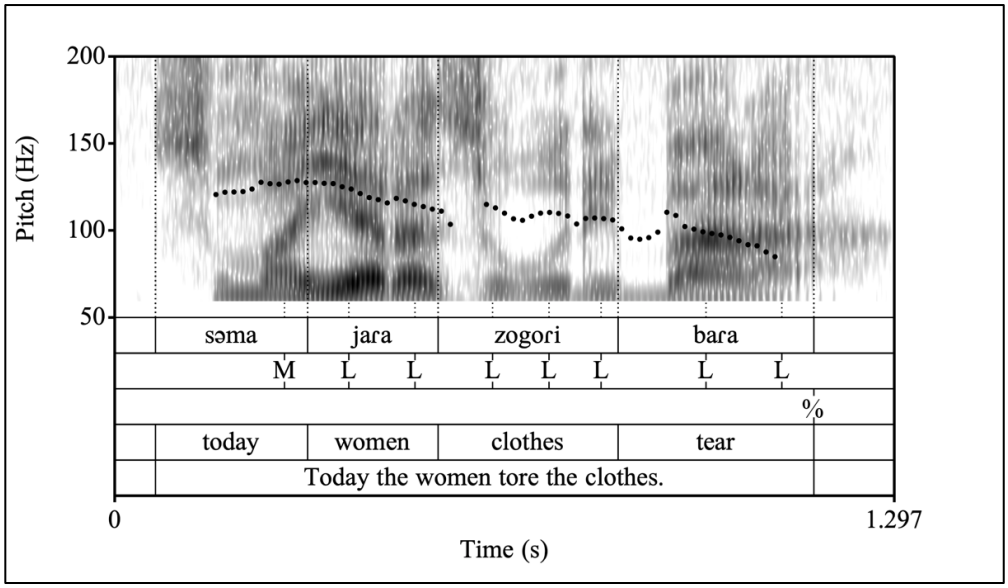


Figure 24

Declarative statement ending with M tone (S3, M)

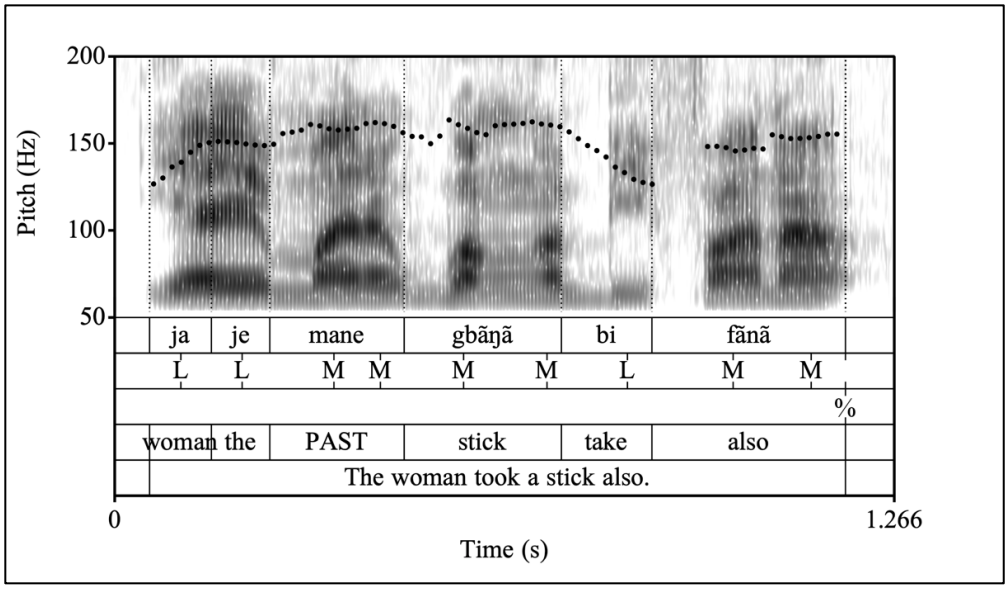


Figure 25

Declarative statement ending with H tone (S3, M)

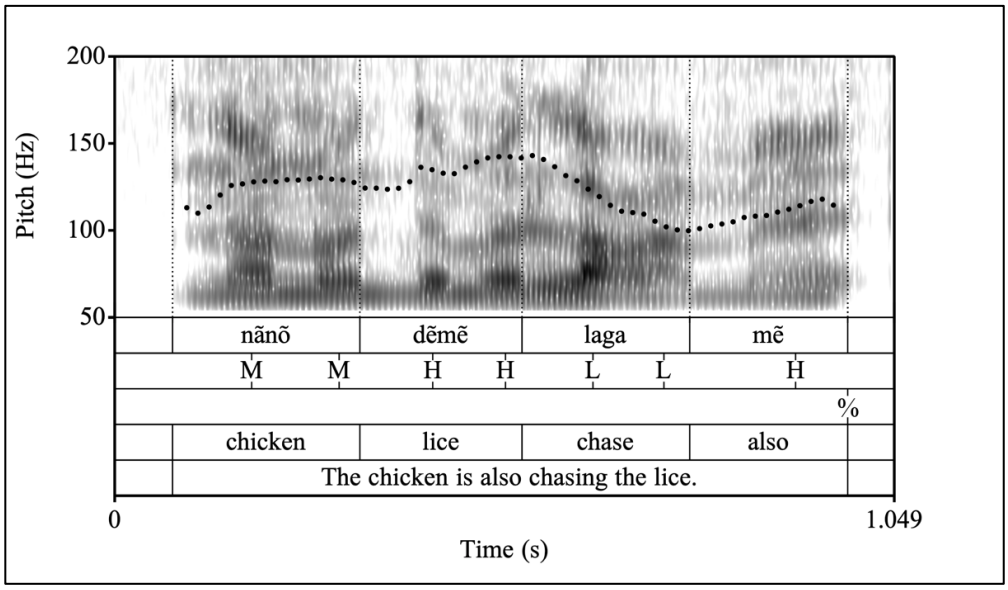


Figure 26

Declarative statement ending with M tone with final fall in F0 (S3, M)

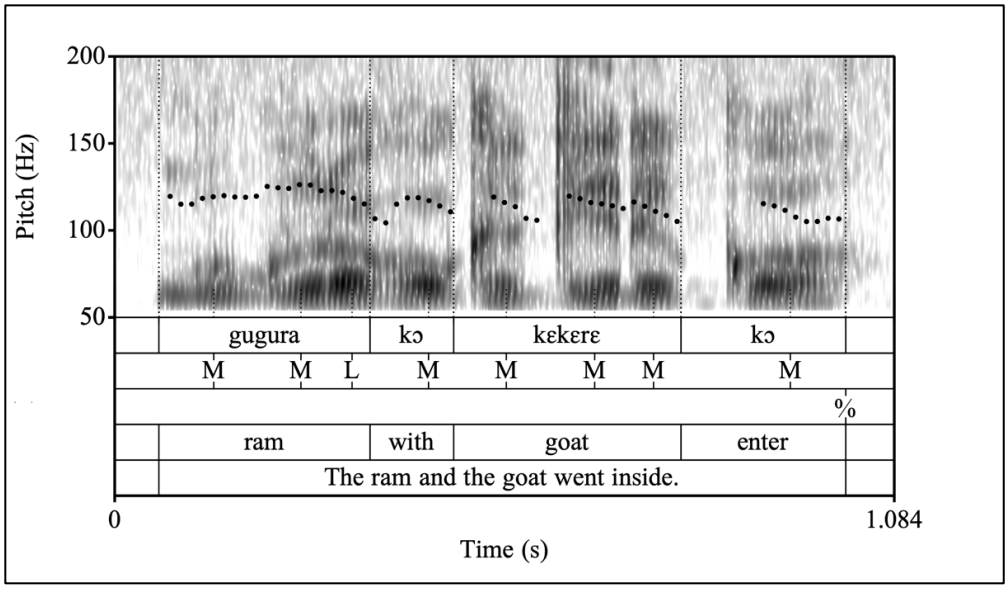
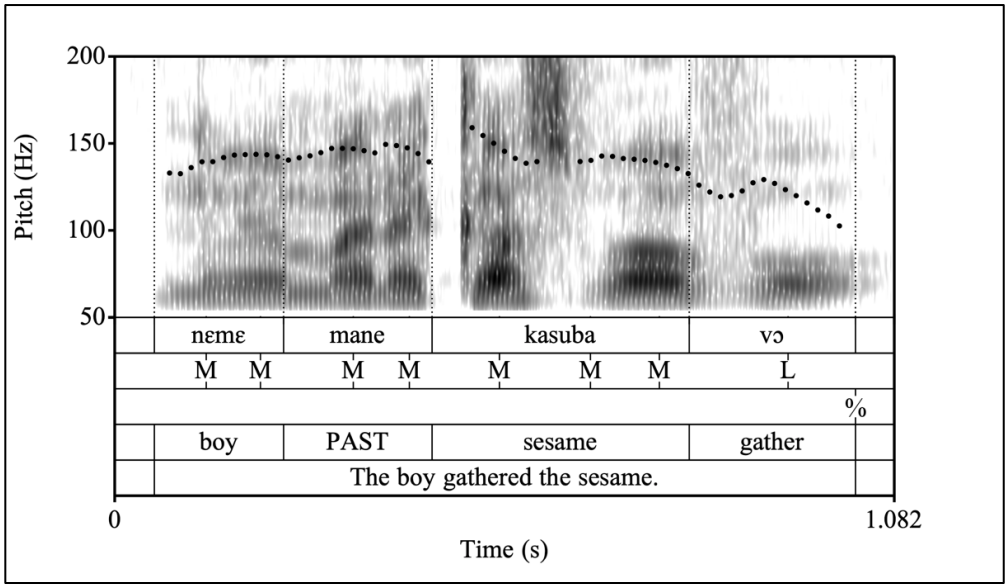


Figure 27

Declarative statement ending with L tone (S3, M)



Boundary tones do not always realize as a change in the trajectory of F0. Downing and Rialland (2016) describe *superimposition* as another possible realization. When boundary tones are superimposed, an L% or H% boundary tone lowers or raises the F0 range of the final lexical tone(s), rather than being realized after. They cite Embosi (Rialland & Aborobongui, 2016), Bemba (Kula & Hamann, 2016), and Shindgazidja (Patin, 2016) as examples but this is likely more widespread and described using different terminology. My earlier work on Bobo (Sherwood, 2016) found that the final lexical item of some questions appears to have a raised F0, which would be a likely case of superimposition as well. Therefore, it is necessary to consider the F0 range as well as the F0 contour when investigating boundary tone. Figure 23-Figure 25 are representative of the F0 range of final lexical tones as well as their F0 trajectories. The most informative cases are sentences ending in lexical M tones, as either lowering or raising of F0 should be easily observable. Note, however, that in Figure 24 the F0 of the final M-toned word is nearly identical to the F0 of M-toned words earlier in the sentence; there is no change in pitch register. The

realization of final lexical L tones also does not indicate the presence of a superimposed boundary tone. In Figure 23, the fall in F0 that occurs in *ba¹ɾa*¹ 'tear' is a continuation of the steady decline in F0 that began at the start of the string of L tones, without a change in slope or register. On the other hand, when the utterance ends in lexical H tone, the pitch peak for the final lexical H tone is lower than the pitch peak for the phrase medial H tone. Taken in isolation, the behavior of phrase-final lexical H tones could suggest the presence of a superimposed L% that results in lowered phrase-final tones. However, this would require an analysis in which this L% only affects H tones, and not L or M tones. This L% would also combine with the string of lexical tones differently than the L% that is used to mark questions (discussed in section §4.4). The lowering of H tones occurs regularly, suggesting that it is phonologically structured rather than the result of phonetic variation. However, it is more likely that the lowered H tone is the result of some type of downtrend rather than an L% that only affects H tones (downtrends are discussed in §4.3.3 below).

The realization of contour tones in phrase-final position is more variable than level tones. The data is also more limited since they were only included in one elicitation task (§2.2.2.3) and sentences with contour toned target words disproportionately excluded due to the errors mentioned in §4.3.1. Table 8 below shows the final F0 contour of sentences ending in a contour toned target word. (ML was not included due to a lack of ML-toned lexical items that could be placed in phrase-final position without undergoing grammatical tone changes.) LM and LH are expected to be realized with a falling-rising (if following an M or H tone) or rising F0 (if following an L tone) while HL is expected to be realized with a rising-falling (if following an L or M tone) or falling F0 (if following an H tone). This does occur in the narrow majority of cases. Deviations from the expected realization are difficult to explain in terms of grammatical, lexical, phonological, or

pragmatic context, since they are not patterned.[26] The realization of contour tones in phrase-final position also does not indicate the presence of boundary tone.

Table 8

F0 contour in the final syllable of words with final LM, LH, or HL Tone

	Total	Falling	Level	Rising	Rising-Falling	Falling-Rising
LM	4	1	2	0	0	1
LH	9	3	1	4	1	0
HL	5	5	0	0	0	0

The realization of all tones is more variable in the spontaneous speech data. For this data, I inspected a sample of 659 phrases from the narratives, monologues, and map tasks. All speakers except for S4 are represented in this data. Phrases were defined by the presence of a pause, since there are no other known criteria that can be used to determine whether a phrase boundary is present. The F0 contour in the final syllable of the phrase was coded for its shape in order to assess overall patterns of realization, using the same system as in the table above. Table 9 shows the results for words ending in an L, M, or H tone. Only phrases that are part of declarative statements are shown. Although this data is highly variable, there is an overall pattern: The final lexical tone of a phrase surfaces with its expected F0 contour the majority of the time: falling or level for L tones, level for M tones, and level or rising for H tones. Final M and H tones surface with a falling F0 contour in some cases. There are only a handful of cases where there are unexpected rises in F0—too few to base any conclusions on. This data is consistent with an analysis in which intonational phrases do not have an obligatory boundary tone.

Whether there is an optional L% boundary tone is unclear. The cases in which there is an unexpected fall in F0 do not have any obvious grammatical, phonological, or pragmatic properties

[26] This kind of variability is typical of contour tones in phrase-medial positions as well. It is extremely difficult to make generalizations about how contour tones are realized in context.

Table 9

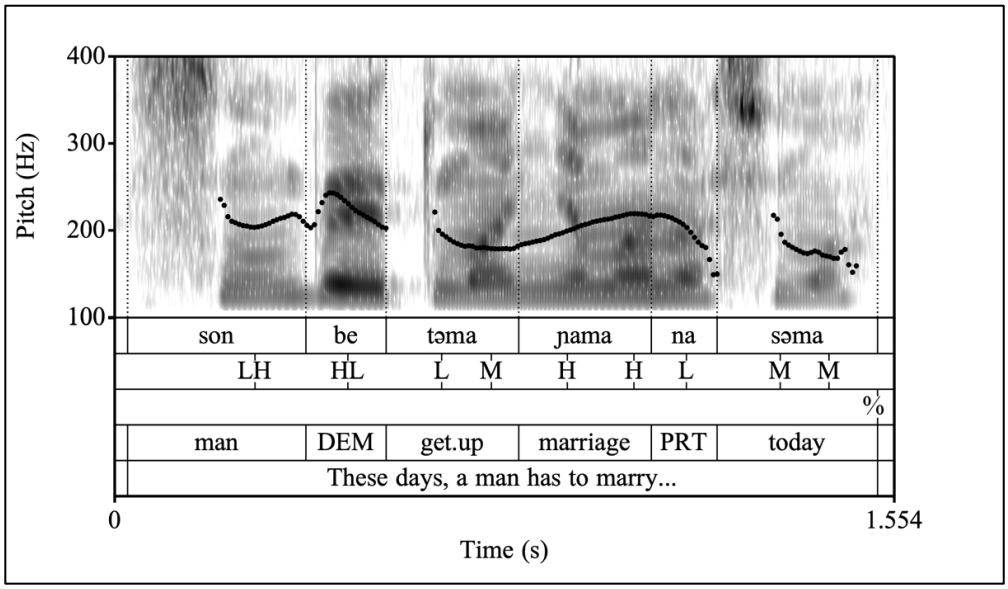
F0 Contour of Phrase-final Level Tones in Spontaneous Speech

	Total	Falling	Level	Rising	Rising-Falling	Falling-Rising
L	239	188	45	0	0	6
M	87	26	55	3	2	1
H	52	10	30	12	0	0

in common. Additionally, as discussed above regarding the elicited data, the categorization of the phrase-final F0 contours does not capture differences in degree of the fall in F0. In most cases the fall is slight. Figure 28 is a typical example. This does not rule out the possibility of an underlying L% boundary tone, as the realization of phrase-final L tones in the spontaneous speech data does not always involve an extremely pronounced fall in F0. However, it is also not clearly a case of an underlying L tone target; it could also be phonetic variation in the expression of the tone. If the fall in F0 occurred in predictable contexts, it would be easier to draw the conclusion that these phrases possess an L% boundary tone. Since it is not predictable, that question cannot be answered with this data.

Figure 28

Intonational phrase ending in M tone from an interview task (S2, F)



There is no evidence of boundary tone in the elicited declarative statements, while the spontaneous speech data is more equivocal; it is possible that there is an L% boundary tone, but that is not demonstrable at this time. However, it is clear that a boundary tone cannot be an *obligatory* element of the IP. Furthermore, there is no indication of other intonational tones, such as pitch accent. That is, there are no patterned deviations from the pitch contour expected from the lexical tones alone. In the elicited statements, the F0 contours of intonational phrases are determined by the lexical tones, and, in the case of H tones, some kind of downtrend. There are no other regular deviations in F0 that suggest additional intonational phenomena. (Whether there are pitch accents or other intonational phenomena in focused items will be discussed in Chapter 5; the spontaneous speech data is too complex for this kind of analysis.) Bobo appears to be a language that has limited obligatory intonation in declarative statements.

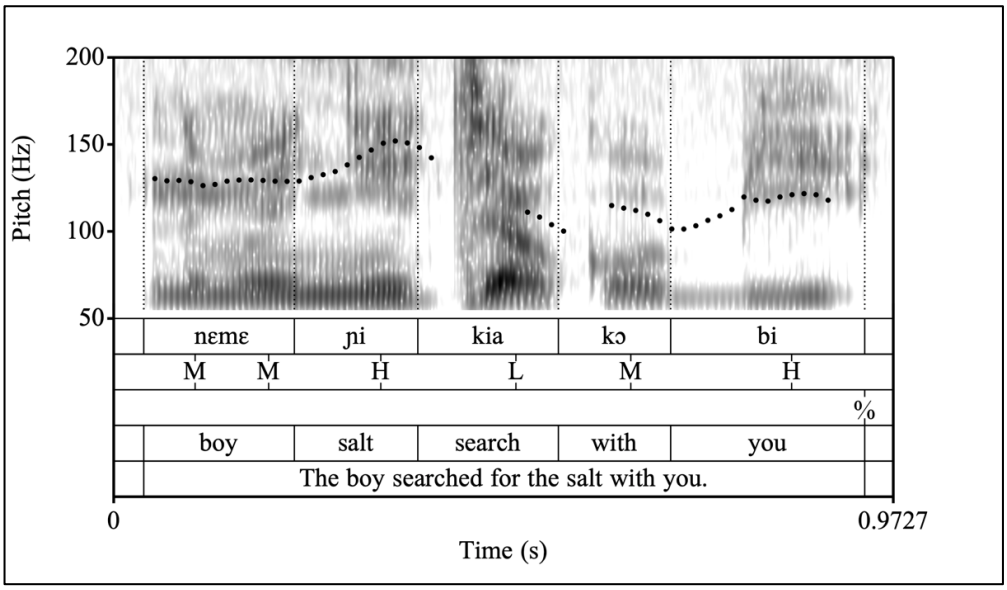
4.3.3 Downtrends

A full investigation into downtrends is not possible with the existing data, but some initial characterizations and issues can be stated. In this section, I incorporate some examples from the elicitation task described in §2.2.2.2 as well as an earlier elicitation task that was designed to provide data on the realization of tone melodies different tone contexts. This elicitation task was recorded in 2015 with the same speakers and in the same manner as the elicitation tasks described in Chapter 2. From the elicitation data, it is clear that some process results in the lowering of phrase-final H tones, as is seen in Figure 25. This example is illustrative of a general pattern in which the F0 of later H tones do not reach the same height as earlier H tones. Figure 29-Figure 31

are additional examples of this pattern.[27] Note that in Figure 31 the lowered H tones are not in phrase-final position.

Figure 29

Declarative statement with lowered F0 during final H tone (S3, M)

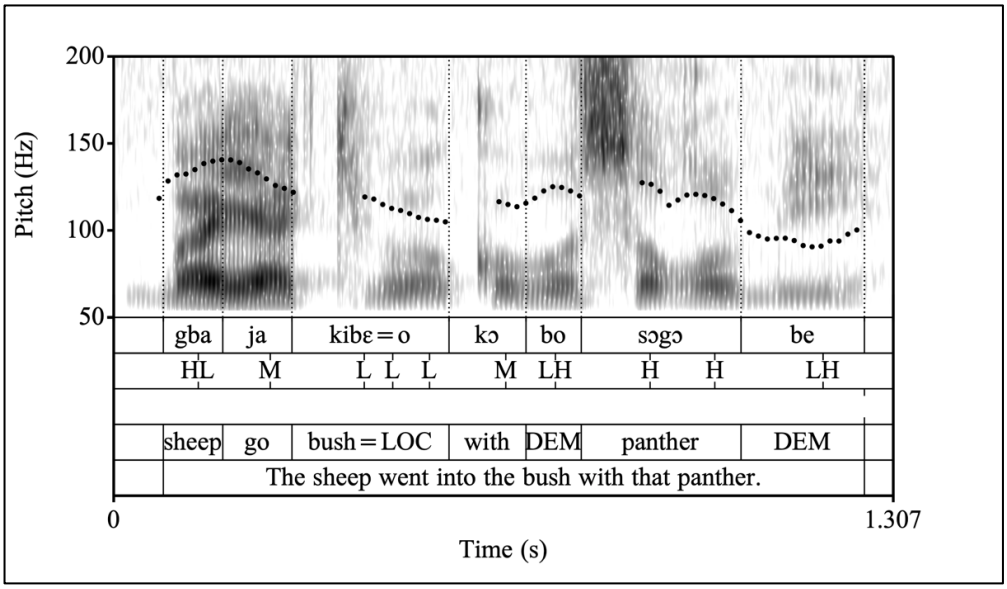


The primary question is whether this is due to declination or downstep. These analyses make distinct predictions. In the case of declination, we expect to see a gradual decrease in F0 over the course of the domain regardless of its phonological or grammatical content. This decrease would be most easily observable in strings of like M or H tones.[28] Such strings are typically realized with a level F0, however, which is what is expected without declination. Figure 32-Figure 33 show two examples for M tones and Figure 34-Figure 35 show two examples for H tones. There is no evidence of a gradual decrease in F0 across the phrase. Instead, the F0 only decreases when

[27] A quantitative analysis of the difference in F0 between phrase-medial and phrase-final words was not possible due to errors in pitch tracking caused by breathy or non-modal phonation, which can be seen in these examples. The overall F0 contour is visible (and audible) but tracking becomes unreliable near the end of the phrase-final vowel.
[28] L tones typically realize as a fall in F0 unless the speaker is already near the bottom of their range.

Figure 30

Declarative statement with lowered F0 in phrase-medial and phrase-final LH and H tones (S3, M)



the lexical tone changes from H to M or L or from M to L. It does not appear that declination is a consistent property of the IP. Instead, the lowering appears to be conditioned by the phonological content of the phrase, indicating that it is caused by downstep. H tones are lowered but M tones are not. In the examples shown previously (Figure 16, Figure 20-22), an L tone occurs before—although not always immediately adjacent to—the lowered H tones. This is consistent with downstep.

As stated previously, downstep is typically analyzed as a change in register that is triggered by the presence of an L tone (Connell & Ladd, 1990; Hyman, 1993). There are disagreements about what a change of register entails. Some propose lowering of the tonal space range with respect to the speaker's baseline (Ladd, 1993), while others propose that the F0 ceiling is lowered, resulting in a narrowing of the tonal space (Pierrehumbert & Beckman, 1986). The latter analysis is the most consistent with the Bobo data, as a lowered pitch ceiling would not affect the realization

of M tones unless it was lowered beyond the M tone range.

Some instances of lowering are not directly accounted for by downstep, however. Lowering of H tones also can also sometimes occur after M tones, as is shown in Figure 36-Figure 37. The lowering of H tones after M tones is not substantial: In the above examples, it the F0 peaks of the lowered H tones are only between 10-11Hz lower than those of the initial H tones in the phrase. Compare this to the lowering of H tones in Figure 29-Figure 31, where the difference between the initial H tone and the first lowered H tone is between 15-30Hz. If M tones also trigger downstep, this does not explain why it occurs inconsistently (see Figure 29 for a counterexample) or the difference in the degree of lowering. There is no easy, parsimonious explanation. One possibility is that in addition to downstep, there is also a local coarticulatory process resulting in lowered H tones after L or M tones (i.e. "downdrift" in the sense of Connell, 2001; see footnote 25). However, this does not explain why M tones are not lowered after L tones. Until more data can be examined, the process that lowers H tones after M tones remains an open question.

Figure 31

Declarative statement with lowered F0 in a phrase-medial H tone (S3, M)

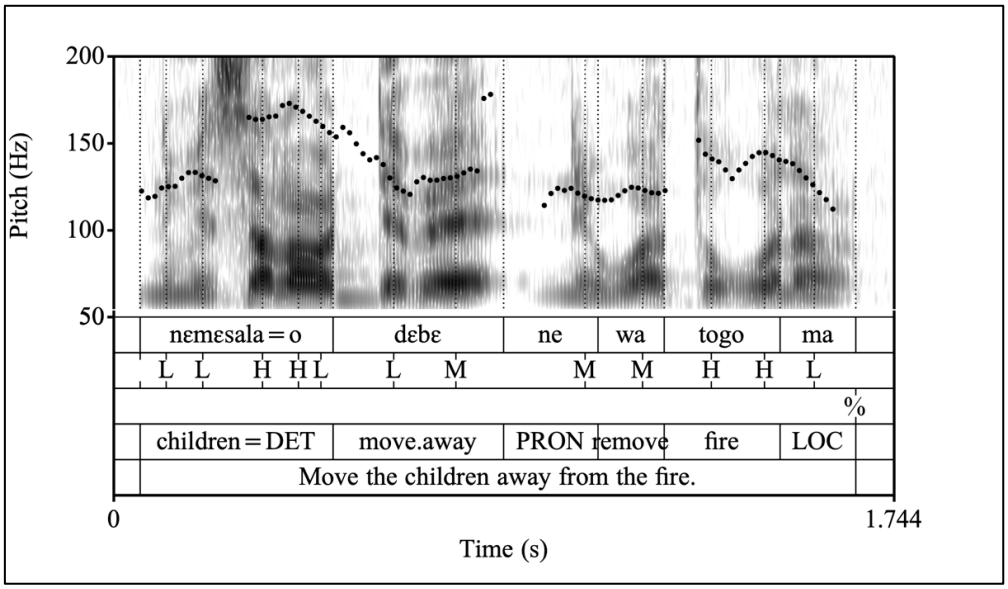


Figure 32

Declarative statement with no declination in a string of M tones (S3, M)

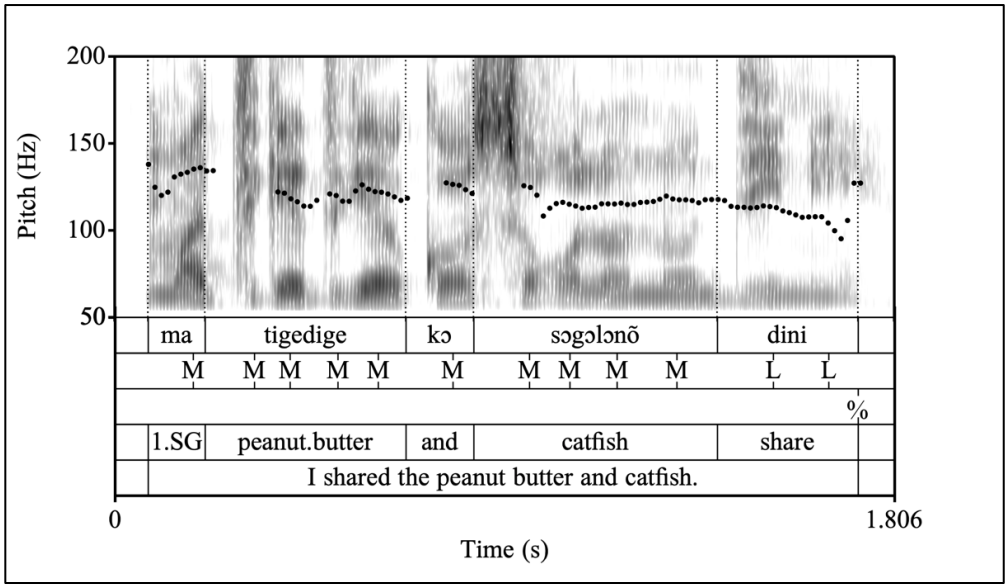


Figure 33

Declarative statement with no declination in a string of M tones (S3, M)

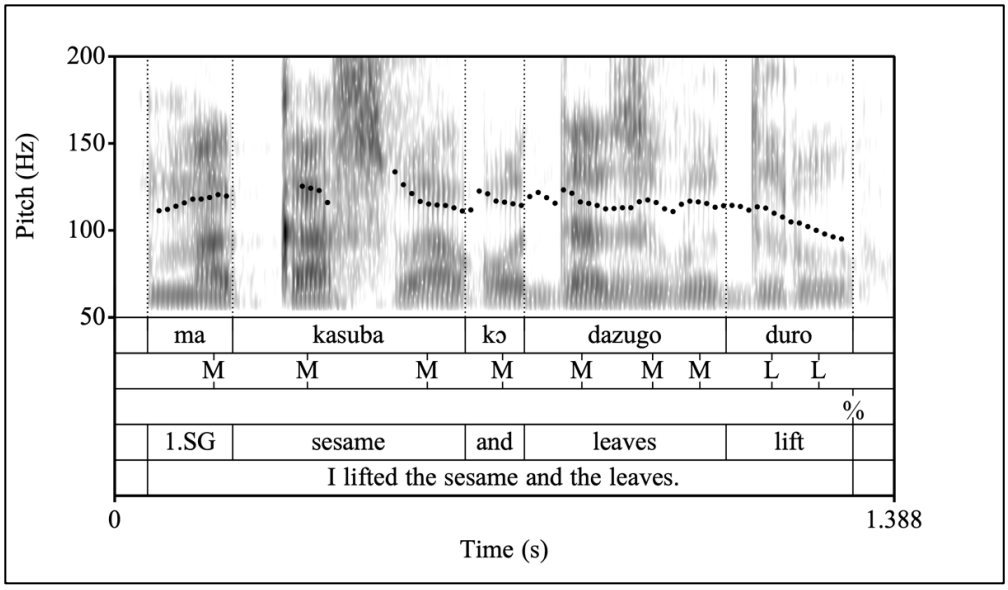


Figure 34

Declarative statement with no declination in a string of H tones (S2, F)

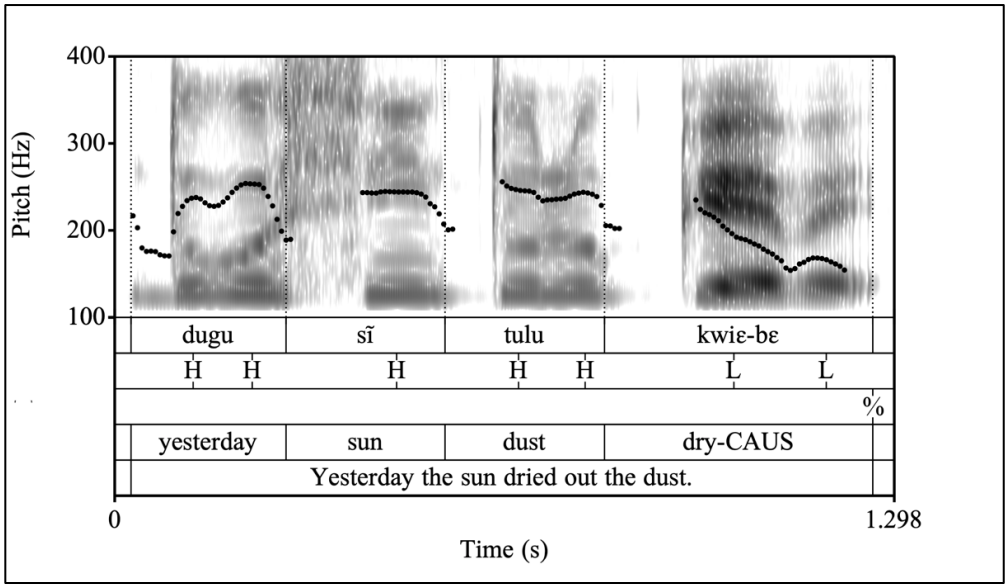


Figure 35

Declarative statement with no declination in a string of H tones (S3, M)

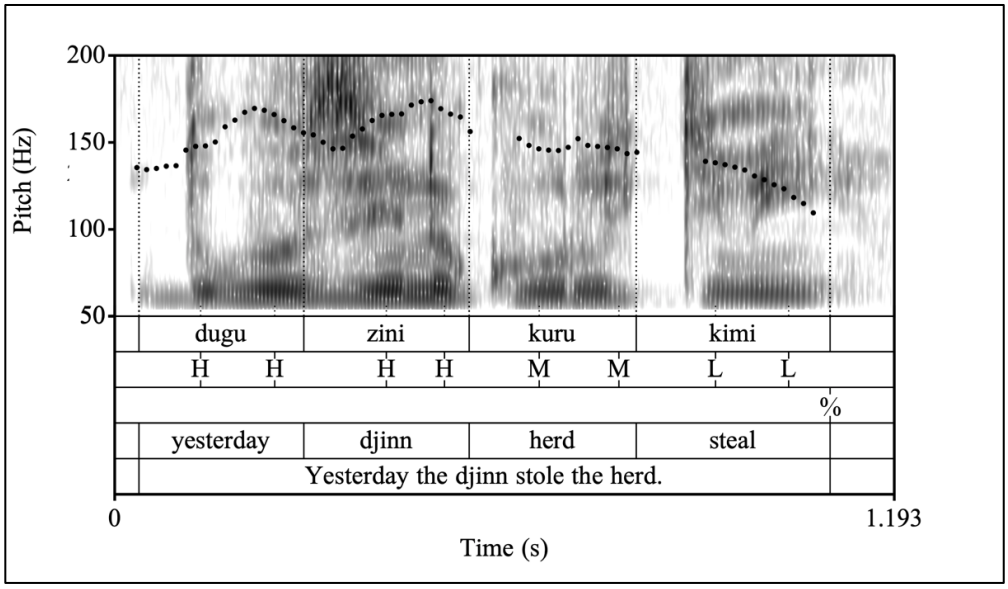


Figure 36

Declarative statement showing an H tone lowered after an M tone (S3, M)

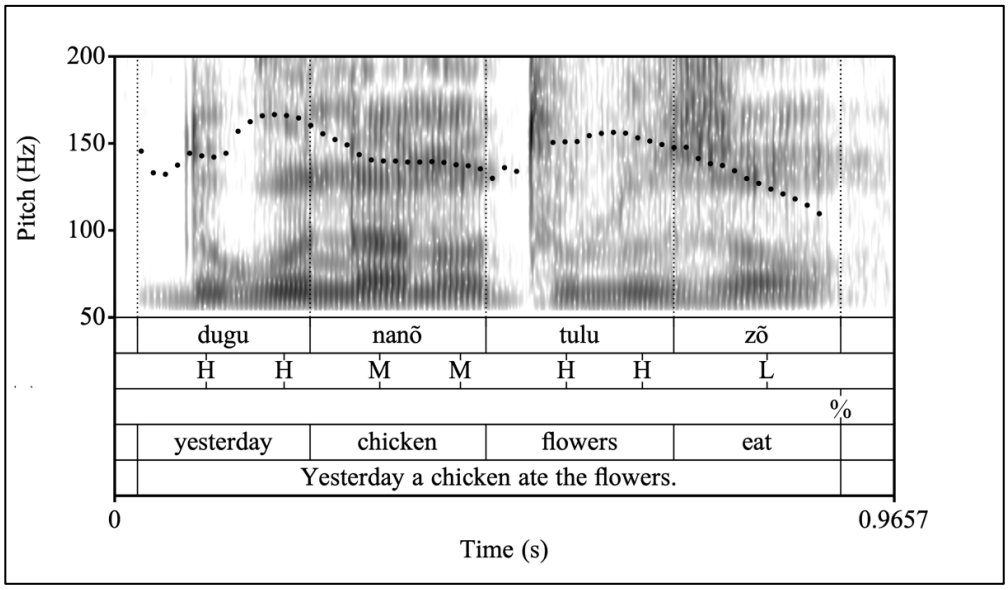
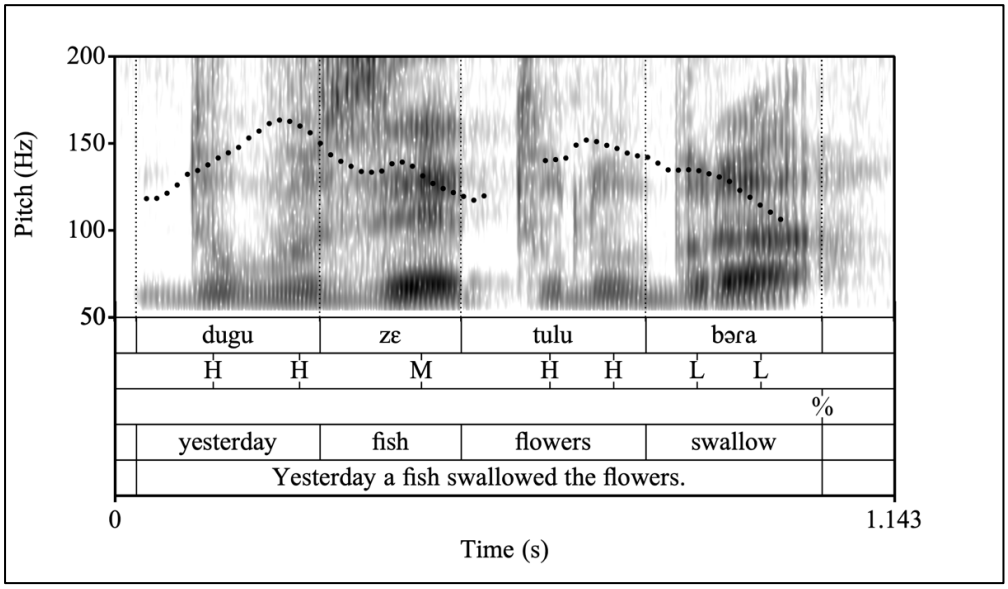


Figure 37

Declarative statement showing an H tone lowered after an M tone (S3, M)



In summary, Bobo does not appear to have declination as a consistent property of the IP. H tones are lowered after L tones, indicating a process of downstep. This is consistent with Morse

(1976), who notes that Bobo has "automatic downdrift"—although she does not provide any more details, she is most likely referring to lowering of H tones following an L tone, rather than declination. Lowering after M tones is not as pronounced or as regular as lowering after L tones, suggesting that it might be a coarticulatory rather than phonological process. More data is needed to determine the exact phonological conditions that trigger lowering of H tones: whether lowering after L and M tones is caused by the same process, and whether the lowering is a local effect or whether it operates over a larger phonological domain (such as a prosodic phrase).

4.3.4 Final lengthening

Since final lengthening is likely to be a universal property of prosodic phrase structure, it is expected that Bobo will have some degree of final lengthening in simple declarative statements. In order to verify this, the durations of identical phrase-medial and phrase-final vowels in the elicitation sentences described in §2.2.2.2 (phrase prosody elicitation task 1) were compared. This is the subset of elicitation sentences examined in this chapter that contain both phrase-medial and phrase-final target words. 60 of the sentences from this task that were included in the project are declarative statements. Eight of the statements were excluded from the analysis of duration in this section because either the final segment of the phrase-medial or phrase-final word could not be segmented. The phrase-medial and phrase-final vowels were segmented in Praat based on an analysis of the spectrogram and waveform and their durations measured with a Praat script. The durations of the vowels in statements were normalized by speaker using z-scores and a paired t-test performed.

The durations of phrase-medial and phrase final vowels indicate that, as expected, Bobo has phrase-final lengthening in declarative statements, $t(51)=4.97$, $p < .001$. Table 10 provides the means for the phrase-medial and phrase-final vowels and the mean difference between them. The

means for each speaker are based on raw values, while the pooled means are based on the normalized data. As can be seen by the standard deviations, the duration of the phrase-final and phrase-medial vowels are quite variable even within speaker. Some of this variation is due to averaging over syllables with different segmental content (as a reminder, while the vowels differed across sentences, the phrase-medial and the phrase-final vowel were always the same within each sentence). There are also interesting differences between speakers. For both S2 and S3, the phrase-final vowel is longer than the phrase-medial vowel, but for S4 it is not. Figure 38 shows the data by speaker.

Table 10

Mean Duration of Phrase-Medial and Phrase-Final Vowel in Declarative Statements

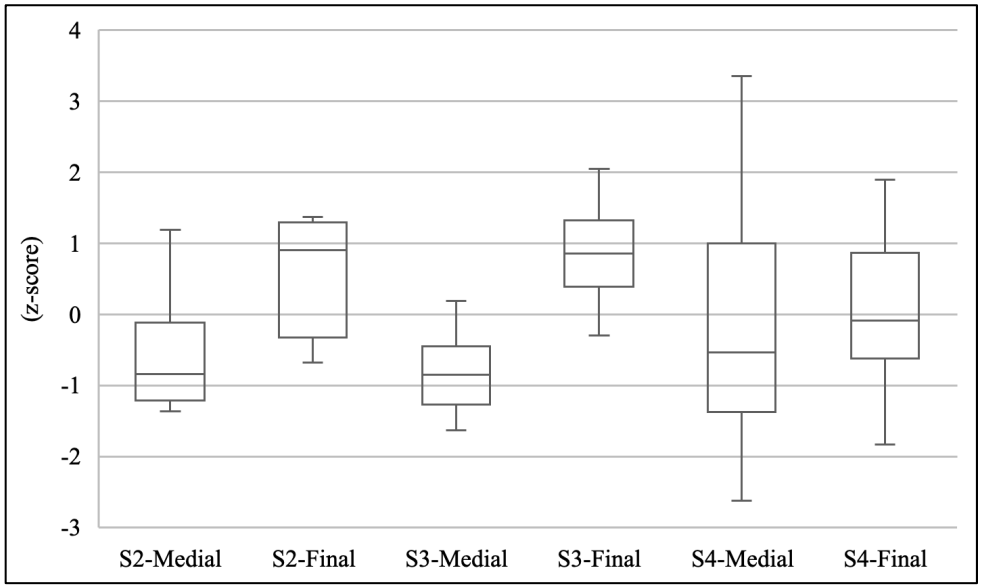
	Medial		**Final**		**Final-Medial**		**# of Sentences**
	M	*SD*	*M*	*SD*	*M*	*SD*	
S2	93	24	118	22	25	41	5
S3	67	19	138	25	72	33	24
S4	84	30	86	18	3	32	22
Pooled	-0.49	1.22	-0.49	0.88	0.98	1.5	51

Note. M=mean, SD=standard deviation. Final-medial is the mean difference between the final segment and medial segment within the same token. Means for individual speakers are calculated using raw values (ms) and pooled data using z-scores.

Despite the variation between speakers, when individual phrase-medial and phrase-final tokens are compared, the strong majority of sentences (85%) have a longer phrase-final vowel. In 73% of the sentences, the phrase-final vowel is longer than the phrase-medial vowel by more than 10ms. This tendency is also observable for S4 when his tokens are considered individually. There were eight sentences in which the phrase-medial vowel is longer than the phrase final vowel and seven of these were produced by S4 and one was produced by S2. The remainder of the sentences produced by S4 do show evidence of final lengthening: In eleven, the phrase-final vowel is longer than the phrase-medial vowel by more than 10 milliseconds, and in three it is longer but by less

Figure 38

Duration of phrase-medial and phrase-final vowels in statements by speaker (z-scores)



Note. Whiskers display minimum and maximum values.

than 10 milliseconds. This is obscured when the difference across his sentences is averaged because when the phrase-final vowels are shorter, they are shorter by a considerable amount: between ten to sixty-one milliseconds shorter. It is not clear at this point why these tokens are so different.

To conclude, although an individual utterance might not show evidence of phrase-final lengthening due to other factors affecting the duration of the vowels involved, the overall analysis of the data indicates that Bobo does have final lengthening.

4.3.5 Non-modal vowel phonation

Non-modal vowel phonation at the end of phrases is common across all of the speakers (S2, S3, S4) who participated in the elicitation tasks discussed in the previous sections. To assess how frequently they use non-modal vowel phonation, what kind of non-modal vowel phonation they use, and whether it is correlated with lexical tone, the elicitation tasks described in §2.2.2.2

(phrase prosody elicitation task 1) and §2.2.2.3 (phrase prosody elicitation task 2) were used. The phrase-final vowel of each token ending in a level-toned final syllable was coded for the presence of modal, creaky, or breathy voice based on an impressionistic transcription. Five tokens were excluded because it was difficult to determine their phonation type. The number of tokens showing each phonation type is given in Table 11.

Table 11

Vowel Phonation in Phrase-final Vowels in Statements

	Total	**Modal**	**Creaky**	**Breathy**
L	61	12	44	5
M	30	9	5	16
H	45	11	34	0
ALL	136	32	83	21

Approximately 60% of the phrase-final vowels have creaky phonation; of the remainder, approximately half have breathy phonation while half are modal. The extent of the non-modal voicing is variable. There appears to be some correlation with the final lexical tone, with creaky phonation occurring more frequently with L and H tones, and breathy phonation more frequently with M tones. The association of creak with both L and H tones could be due to different types of creak, e.g. vocal fry and tense voice (Keating et al., 2015). Overall, the distribution of non-modal voicing is similar to that of final lengthening: it occurs in the strong majority of statements, but not all. Since the type of non-modal voicing is associated with the final lexical tone, creaky or breathy phonation themselves cannot specifically be a requirement of the IP. Non-modal voicing in general, however, is a strong (but not required) cue.

Whether non-modal phonation occurs, and in what form it occurs, is not predictable based on the phonological or grammatical content of the utterance. One brief note: Although Morse (1976) claims that a final glottal stop is a property of the “period contour”, which these sentences

should have, there is no evidence of it in this data.

4.4 Questions

This section examines the prosodic properties of polar and *wh*-questions. The elicitation data that was collected in 2014 shows a different type of question intonation than the elicitation data that was collected in 2016-2017, so I will discuss the 2016-2017 data first (§4.4.2 and §4.4.3) and then discuss the 2014 data (§4.4.4) and spontaneous speech data (§4.4.5) separately.

4.4.1 Analysis

The 2017 elicitation data consists of the elicitation sentences collected as part of the tasks described in §2.2.2.2 and §2.2.2.3 (phrase prosody elicitation tasks 1 and 2). 165 of the tokens included from these tasks are polar questions and 63 are *wh*-questions. Five tokens (one polar, four *wh*-) were excluded from the analysis of F0 because non-modal phonation obscured the F0 contour of the final syllable of the target word. Eleven tokens (five polar, six *wh*-) were excluded from all analyses of duration because the final vowel of the phrase-final target word could not be segmented. The target vowels of all tokens were segmented in Praat. The duration was measured and F0 was coded in the same manner as for declarative statements (§4.3.1).

4.4.2 Polar Questions

Morse (1976) does not mention a specific intonational contour for questions. Her description only mentions the absence of an utterance-final glottal stop, in contrast to statements which do have an utterance-final glottal stop. I did not find that a final glottal stop is a characteristic of basic declarative statements, however, which means that questions cannot be distinguished from statements by the absence of one. Le Bris and Prost (1981) and Sanou (1993) do not discuss the prosody of questions at all, instead restricting their attention to the formation of questions through the use of particles or *wh*-question words.

The questions from the 2017 elicitation tasks were found to have the following characteristics: (a) a falling F0, (b) extreme lengthening of the final vowel, and (c) breathy phonation, as will be discussed below. These are three of the four prototypical characteristics of the "lax question prosody" described by Rialland (2009). The fourth characteristic, a sentence-final low vowel, is found in the "lax question prosody" languages that use question morphemes (such as particles). It is precluded here because no question particles were used in the design of the sentences. The final vowel of e sentence is therefore the final vowel of the final word, which is either a noun, verb, or adverb. Speakers report that Bobo does have a sentence-final question particle *wa¹*, but it does not occur in any of the recorded data. This particle is likely borrowed from Jula, which has an identical question particle.

Table 12 shows the number of sentences according to the shape of their final F0 contour. Final lexical L and M tones are realized with a falling F0 contour, while the majority of final lexical H tones are realized with a rising-falling F0 contour or a falling F0 contour. Three examples of the falling pitch contour are shown in Figure 39-Figure 41, for sentences ending in L, M, and H lexical tones respectively. For the M and H tones, this fall contrasts with the intonational contour of declarative statements, which do not have a final fall. For the L tones, whether there is a difference is not clear as an L tone is expected to surface with a falling F0 in declarative statements as well.

I interpret the decline in F0 during phrase-final M and H tones as the result of an L% boundary tone. The L% boundary tone is concatenated after the final lexical tone, rather than being superimposed. This is illustrated by a comparison between the question in Figure 41 with its declarative counterpart in Figure 25. The F0 peak of the final H tone in the question is not lowered; rather, the F0 falls after the peak is reached. The fall in F0 is also not due to the breathy phonation;

it begins before the onset of breathiness in most cases (including the examples shown here). The lowering also cannot be explained through purely articulatory factors, such as the loss of aerodynamic pressure in an extended final syllable. The pattern is extremely consistent, much more than would be expected than if it was not reflection of some underlying representation.

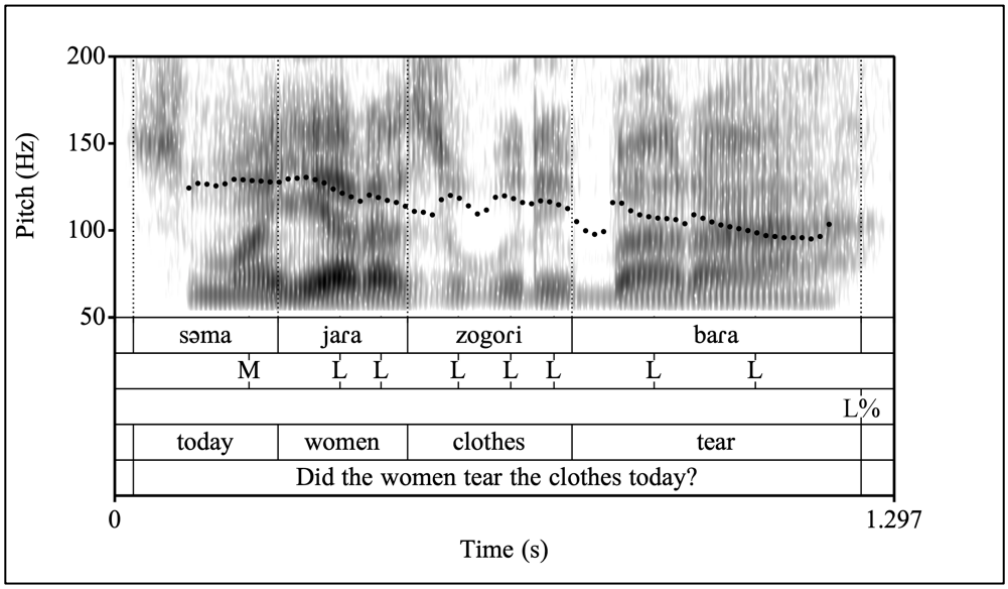
Table 12

F0 Contour in Final Syllable of Polar Questions

	Total	Falling	Level	Rising	Rising-Falling
L	65	65	0	0	0
M	33	26	7	0	0
H	47	15	0	1	31
LH	9	0	0	0	9[29]
LM	5	4	1	0	0
HL	5	5	0	0	0

Figure 39

Polar question ending in an L tone (S3, M)



[29] One of the LH token in this data set has a falling – rising – falling F0 contour, which I included under rising – falling because the relevant portion of the contour is the final fall.

Figure 40

Polar question ending in an M tone (S3, M)

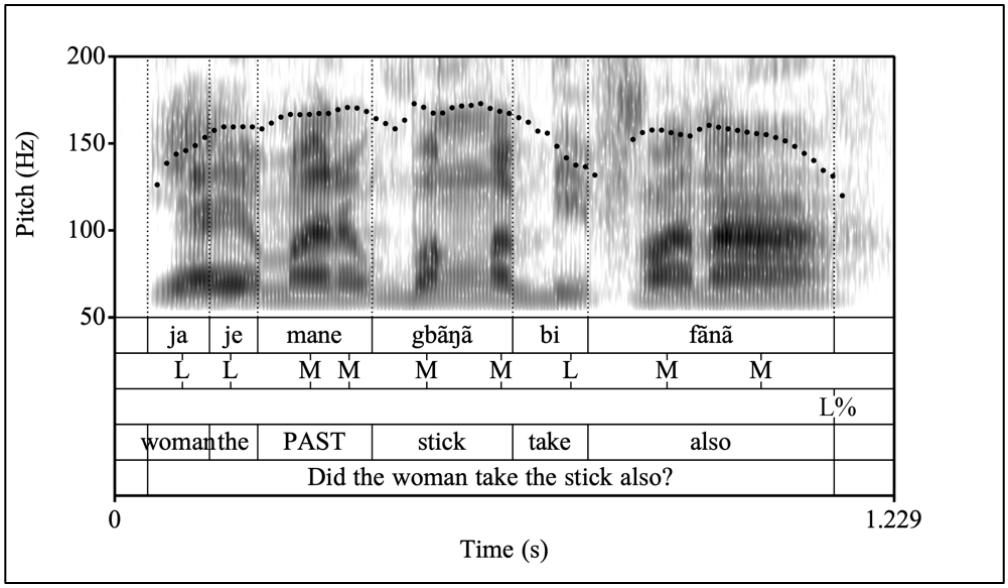
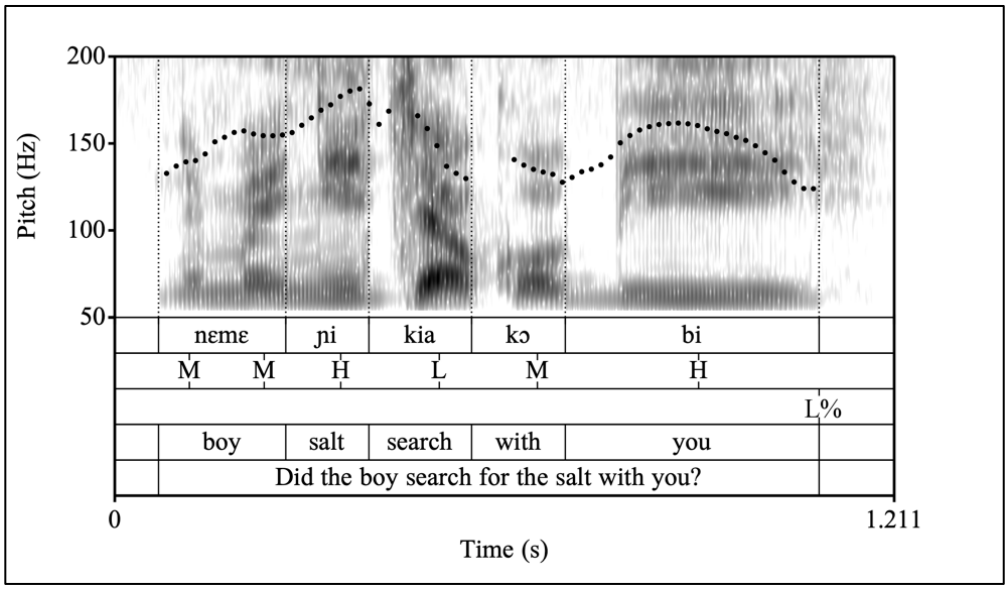


Figure 41

Polar question ending in an H tone (S3, M)



To test for final lengthening, the phrase-medial and phrase-final target vowels in polar questions from phrase prosody elicitation task 1 (§2.2.2.2) were compared. 63 of the sentences from this task that were included in this project are polar questions. Eight of these had to be

discarded because either the phrase-medial or phrase-final target word could not be segmented. The vowels were segmented and their durations measured following the same procedure as for statements (§4.3.4). The duration of the target vowels in polar questions were then normalized by speaker using z-scores and a paired t-test performed. The phrase-final vowel is significantly longer than the phrase medial vowel, $t(54)=23.7$, $p < .001$. Table 13 provides the means for each speaker and for the pooled, normalized data. Figure 42 displays each speaker's normalized data.

Table 13

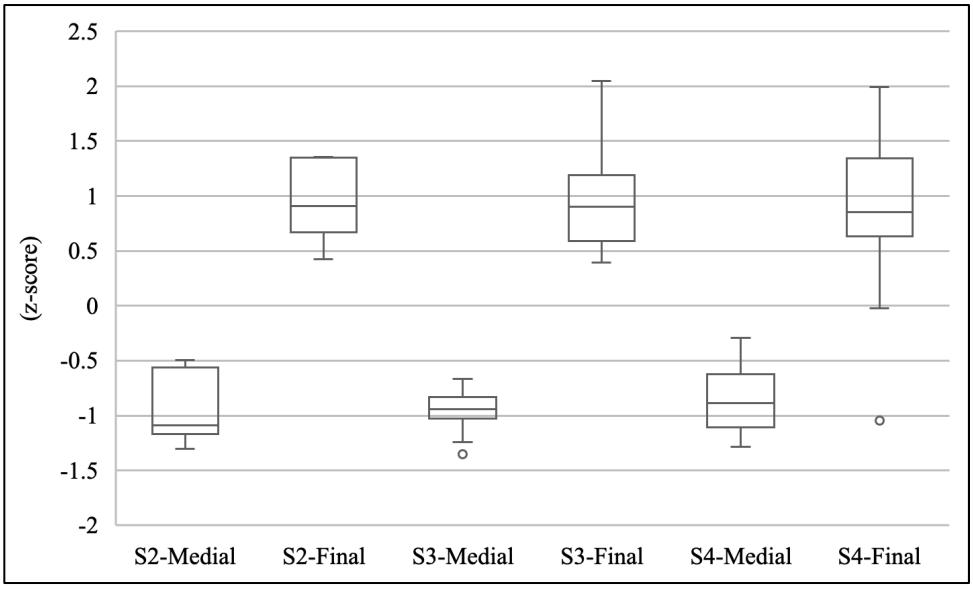
Mean Duration of Phrase-Medial and Phrase-Final Vowels in Polar Questions

	Medial		**Final**		**Final-Medial**		**# of Sentences**
	M	*SD*	*M*	*SD*	*M*	*SD*	
S2	84	30	278	34	193	50	6
S3	60	17	255	42	194	41	25
S4	88	23	214	56	126	51	24
Pooled	-0.89	0.25	0.89	57	1.79	0.56	55

Note. M=mean, SD=standard deviation. Final-medial is the mean difference between the final segment and medial segment within the same token. Means for individual speakers are calculated using raw values (ms) and pooled data using z-scores.

Figure 42

Duration of phrase-medial and phrase-final vowels in polar questions by speaker (z-scores)



Note. Whiskers display minimum and maximum values excluding outliers.

Polar questions in the elicitation data are marked through a more extreme degree of final lengthening than occurs in declarative statements. The final vowel of an elicited polar question is often approximately twice as long as in a corresponding statement—a difference that is quite audibly salient. To verify that this difference is significant, the statements and polar questions fromphrase prosody elicitation task 1 (2.2.2.2) and phrase prosody elicitation task 2 (2.2.2.3) were compared. The data from these tasks were normalized by speaker using z-scores and a two-sample t-test performed. The duration of the final segment in polar questions is significantly longer than the duration of the final segment in statements, $t(259)$, $p < .001$ (assuming unequal variances). Table 14 provides the means of the normalized final vowel durations in declarative statements versus polar questions.

Table 14

Duration of Final Vowel in Statements and Polar Questions (z-scores)

	Final Vowel		**# of Tokens**
	M	*SD*	
Statement	-0.88	0.35	155
Polar Question	0.86	0.60	160

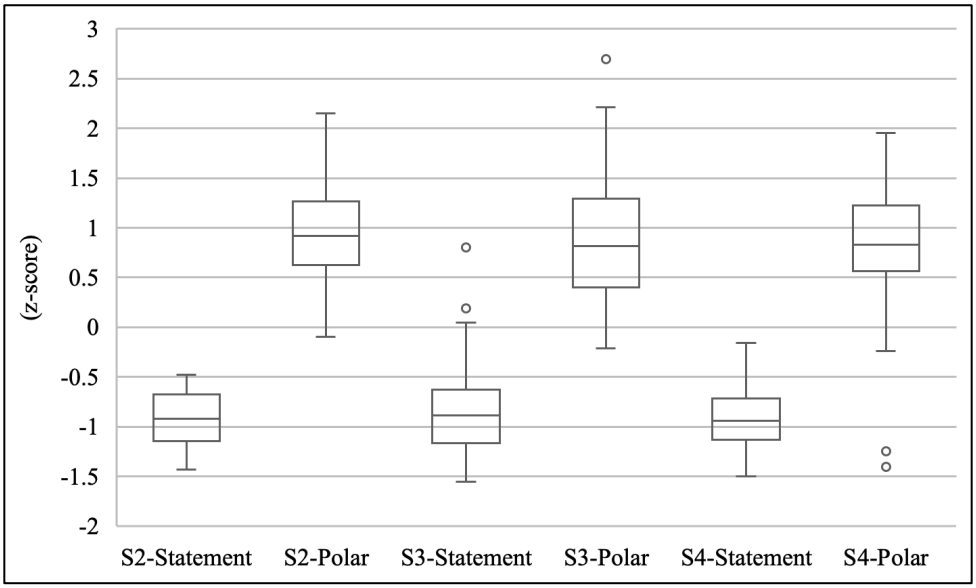
Note. M=mean, SD=standard deviation.

Figure 43 shows illustrates the normalized values for each speaker. The difference in duration between phrase-final vowels in statements and polar questions is strong for all three. It is worth noting that since these are elicitation sentences they are careful productions and the difference is not likely not as extreme in spontaneous speech.[30] To distinguish this extra degree of

[30] The speakers were aware of the purpose of the task—there was no practical way to obfuscate it. Although they were instructed to speak naturally, during elicitations they generally produce distinctions more clearly than in spontaneous speech, especially when they are being recorded. My impression from more informal elicitations is that polar questions do not always have lengthening that is quite this pronounced, although it is still present if a particle is not used. Polar questions without particles are rare in the recordings of spontaneous speech, so it is not possible to compare recordings.

Figure 43

Duration of phrase-final vowels in statements and polar questions by speaker (z-scores)



Note. Whiskers display minimum and maximum values excluding outliers.

final lengthening from the final lengthening found in statements, I will refer to it as *prolongation*—following Rialland's (2009) description of similar question prosodies in other West African languages. This is a terminological convenience and makes no claims about a difference in the underlying mechanism, since I can make no claims about the underlying mechanism at this point.

Non-modal vowel phonation is prevalent, but it patterns differently than the non-modal phonation found in the elicited statements. Table 15 shows the phonation type of level-toned phrase-final vowels in elicited polar questions. While the majority of declarative statements had a creaky termination, the majority of polar questions have a breathy termination. There is also no association of phonation type with lexical tone; approximately two-thirds of tokens ending with each lexical tone have breathy termination. The remainder have modal phonation throughout, except for two tokens with a creaky termination. Examples of breathy termination can be seen in Figure 39-Figure 41 above. The period of breathiness does not extend through the entire vowel.

Table 15

Final Vowel Phonation in Polar Questions Ending in Level Tones

	Total	**Modal**	**Creaky**	**Breathy**
L	60	19	0	41
M	33	10	1	22
H	50	16	1	33
ALL	143	45	2	96

The difference between elicited statements and polar questions is striking and requires an explanation, but there is not a satisfactory one. Breathy termination must be a cue for polar questions, but it does not occur on *all* polar questions; a large number of the tokens (47 out of 143) do not have it. Thus, it cannot be the direct result of a change to the underlying phonology of questions, since this would result in a more regular occurrence. Another potential explanation is that it is correlated with another phonological property. Rialland (2009) interprets the characteristics of "lax question prosody", including falling F0 contour and breathy utterance termination, as being the result of reduced vocal effort at the end of the phrase. Although she does not discuss grammaticalization of the properties of lax question prosody explicitly, in some cases it clearly must have occurred—for example, in languages that have a question particle ending in a low central vowel, since that vowel is phonological segment. Additionally, subsequent research on languages that fall within the geographic region on which lax question prosody is found, and that share some of the characteristics of this question prosody, does not always support the interpretation that reduced vocal effort causes these properties. For example, Akan questions have a falling F0 contour but not breathy termination (Genzel, 2018). At the end of a phrase, a falling F0 contour does generally take less vocal effort than sustaining a rising one, since it does not require maintaining the same degree of air pressure, but there is no independent evidence in Akan that less vocal effort is the *cause* of the falling F0 contour. In Bobo, breathy phonation also appears

to be independent of the falling F0 contour, as it also occurs in questions without a falling F0 contour (discussed in §4.4). It does not occur independently of prolongation, but prolongation is not a consequence of reduced vocal effort. Breathy termination is not an independent phonologized cue for polar questions, but the only property of polar questions on which it is possibly dependent does not straightforwardly explain its occurrence. It is likely also a part of the underlying representation of these polar questions.

In summary, prolongation of the final segment and a falling F0 contour are a robust, phonologized cue for polar questions. Prolongation of the final segment and a falling F0 contour occur in almost all of the elicited polar questions that were recorded in 2016-2017. Breathy utterance occurs nearly as regularly. Bobo therefore is another example of a "lax question prosody" language (Rialland, 2009) in terms of the properties of its question prosody, although the term itself, which presumes a reduced vocal effort at the end of the phrase, might not be accurate.

4.4.3 Wh-*questions*

In *wh*-questions, the presence of a question word signals that a question is being asked. Use of question prosody is therefore not necessary to mark the contrast between a statement and a question, which means that content questions might have a different prosody than polar questions (Ladd, 2008). In Bobo, this appears to be the case. The majority of *wh*-questions in the elicited data do not have an L% boundary tone. Like for declarative statements, the shape of the F0 contour at the final boundary is generally predictable based on the lexical tone alone. Table 16 shows the realization of phrase-final level tones in *wh*-question tokens. A phrase-final L tone generally surfaces with a falling F0 contour, M tone generally surfaces with a level F0 contour, and H tone generally surfaces with a rising F0 contour.

Intriguingly, ten tokens have an unexpected rising F0 contour: five of the L-toned tokens

Table 16

F0 Contour in Final Syllable of Wh-*Questions*

	Total	Falling	Level	Rising	Rising-Falling	Falling-Rising
L	23	13	5	2	0	3
M	18	3	10	1	0	4
H	18	2	0	12	4	0

Note. This table only summarizes the data in phrase prosody elicitation task 1 since task 2 did not include *wh*-questions. This task only includes phrase-final level tones.

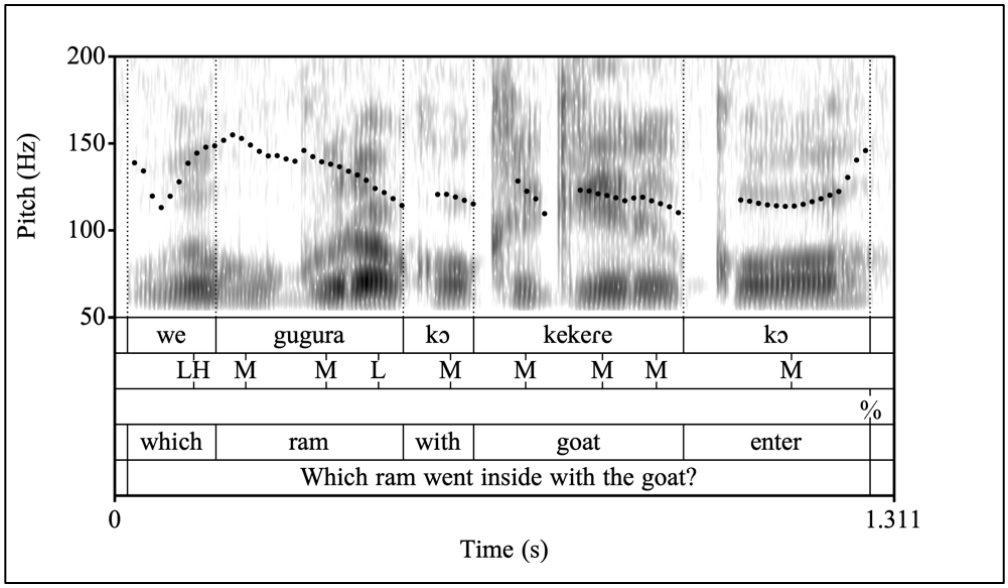
and five of the M-toned tokens have either a rising or a falling-rising contour. An example is shown in Figure 44.[31] It is difficult to explain the rising contour in terms of phonetic variability in the expression of lexical tone. Although it occurs in a minority of cases, the rise in F0 is quite salient and has no apparent explanation other than speaker intention. It is possible that this rise is caused by an optional H% boundary tone. However, since it only occurs in a small subset of the elicitation sentences for reasons that are unclear, it would be difficult to elicit more examples without directly instructing the speakers on what to produce, which could result in unnatural or inappropriate prosody.

Given that final lengthening is present in both declarative statements and polar questions, it is expected that it will be present in *wh*-questions too. To verify this, the phrase-medial and phrase-final target vowels of *wh*-questions from phrase prosody elicitation task 1 (§2.2.2.2) were compared following the same procedure as for statements and questions: The target vowels were segmented in Praat based on an analysis of the waveform and spectrogram and their durations were measured using a Praat script. There are 63 *wh*-questions in this task that were included in the project. Six were excluded because either the phrase-medial or phrase-final target vowel could not

[31] Despite the breathy termination, the F0 on these examples is tracked accurately during the modal portion of the final vowel and there is an audible rise in pitch.

Figure 44

Wh-*question with a final rise in F0 (S3, M)*



be segmented. The durations of the target vowels in wh-questions were normalized by speaker using z-scores and a paired t-test performed. The duration of phrase-final target vowels is significantly longer than the duration of phrase-medial target vowels, $t(56)=12.16$, $p < .001$. Table 17 provides the mean duration for each speaker based on raw values and the means of the pooled, normalized data. Figure 45 displays each speaker's normalized data.

Table 17

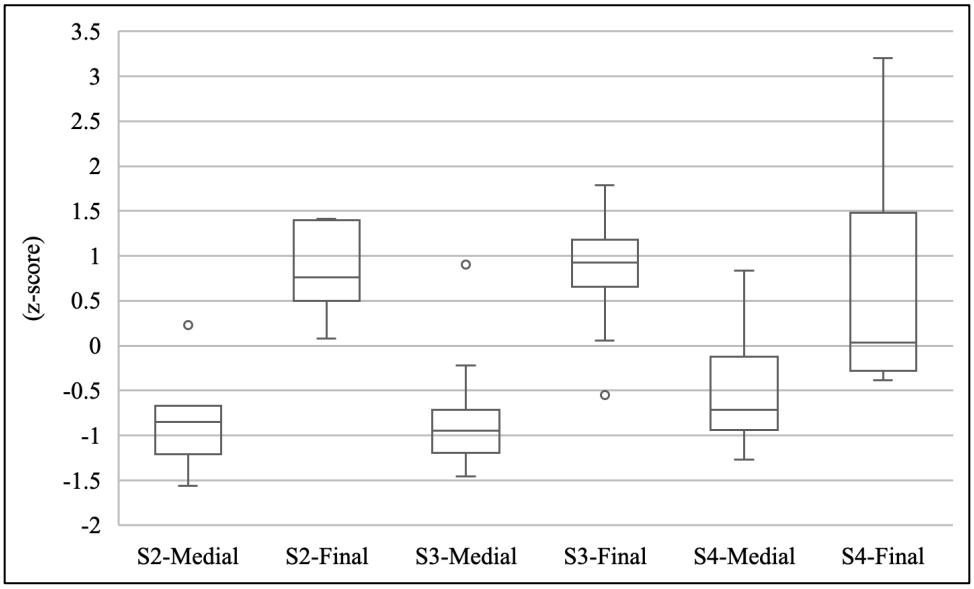
Mean Duration of Phrase-Medial and Phrase-Final Vowel in Wh-*Questions*

	Medial		**Final**		**Final-Medial**		**# of Sentences**
	M	*SD*	*M*	*SD*	*M*	*SD*	
S2	95	38	220	34	126	47	7
S3	74	29	182	32	108	42	26
S4	90	29	145	55	55	52	25
Pooled							

Note. M=mean, SD=standard deviation. Final-medial is the mean difference between the final segment and medial segment within the same token. Means for individual speakers are calculated using raw values (ms) and pooled data using z-scores.

Figure 45

Duration of phrase-medial and phrase-final vowels in wh-*questions by speaker (z-scores)*



Note. Whiskers display minimum and maximum values excluding outliers.

The degree of final lengthening in *wh*-questions is intermediate between declarative statements and polar questions. Table 18 provides the durations of the phrase-final and phrase-medial target vowels in the statements, polar questions, and *wh*-questions. Statistical testing was not performed because if all data are included, there are extremely unequal token numbers across the three sentence types. Figure 46 also displays the duration of phrase-final vowels in statements, *wh*-questions, and polar questions by speaker. All three produce *wh*-questions with a final vowel duration between statements and polar questions.

Breathy vowel phonation is used as frequently in *wh*-questions as in polar questions. Table 19 provides the number of tokens that terminate with modal, creaky, and breathy phonation. One *wh*-token was excluded because the phonation of the final vowel was difficult to determine.

Table 18

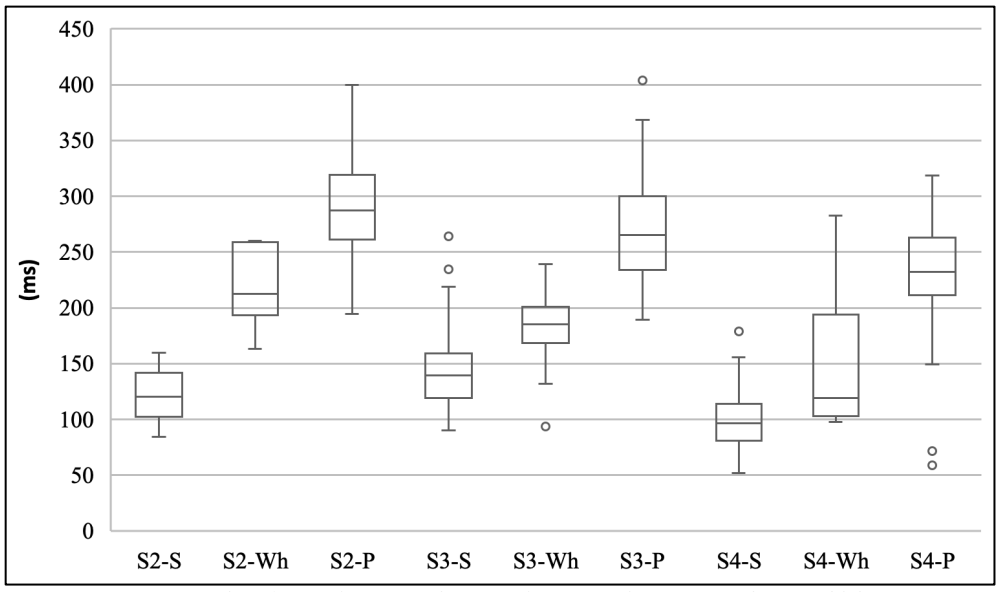
Mean Duration of Phrase-Final Vowels in Statements, Polar Questions, and Wh-*Questions*

	Statement		**Polar Question**		***Wh*-Question**	
	M	*SD*	*M*	*SD*	*M*	*SD*
S2	119	23	288	42	220	35
S3	141	30	268	43	182	32
S4	99	24	231	51	144	54

Note. M=mean, SD=standard deviation. Values are based on raw data (ms).

Figure 46

Duration of phrase-final vowels in statements, wh-*questions, and polar questions by speaker*



Note. S=statement, Wh=*wh*-question, P=polar question. Based on raw values. Whiskers represent minimum and maximum values excluding outliers.

Table 19

Final Vowel Phonation in Wh-*Questions*

	Total	**Modal**	**Creaky**	**Breathy**
L	26	3	2	21
M	18	4	0	14
H	18	0	0	18
Total	62	7	2	53

These *wh*-questions thus do have some properties of question prosody. They share F0 properties with declarative statements; there is no evidence of a boundary tone, except for a handful of tokens that potentially have an H% boundary tone. They do have increased final lengthening, as polar questions do, but it is to a lesser degree. Breathy phonation is a robust cue, occurring on 52 of the 62 total tokens. The properties of *wh*-questions are not simply intermediate or variable use of the question prosody found in polar questions. While the final lengthening is intermediate between declarative statements and polar questions, the F0 and phonation properties are not: the *wh*-questions *consistently* lack the L% found in polar questions, and *consistently* possess the breathy termination. This means that *wh*-questions have their own question prosody—a pattern which uses the same properties found in declarative statements and polar questions, but in a different combination.

4.4.4 Questions in the 2014 elicitation tasks

The data described in the previous sections was collected in 2016-2017 using a different method than the earlier data collection in 2014. During the 2016-2017 elicitations, the speakers read the elicitation prompts from a computer screen without another speaker present. During the 2014 data collection, S1 asked a question and S2 responded, and then they did the same task with reversed roles. Although they were also reading prompts during this elicitation, the question-and-response format meant that the task was more interactive than the elicitation tasks involving a single speaker. There is also a difference in the type of question prosody used. In the 2014 data, the shape of the final pitch contour is more variable in both polar questions and <u>*wh*</u>-questions than in the 2016-2017 data. Table 20 shows a summary of the shape of the final pitch contour in the 2014 data. The majority of the elicitation sentences end in a lexical L tone, and none end in a lexical H tone.

Based on the 2016-2017 data, we expect that polar questions will have an L% boundary tone and *wh*-questions will have no boundary tone. Furthermore, in the 2016-2017 data, L% is concatenated at the end of the string of lexical tones, resulting in a fall in F0 unless the speaker is already near the bottom of their range. Thus, we expect that the F0 contour of the final syllable of a polar question will usually end in a fall in F0, while the F0 contour of the final syllable of a wh-question will be determined by the lexical tone. In the 2014 data, the majority of questions end in a lexical or grammatical L tone, meaning we expect that the majority of both polar and *wh*-questions will end with a fall in F0. However, approximately half of the L-toned polar questions have no fall in F0, regardless of whether the speaker is already near the bottom of their range. This generalization includes the three polar questions ending in a lexical M tone, where the fall should be easily observable. Even these are realized with a completely level F0. There is no indication of L% in the polar questions from the 2014 data, and furthermore, in some of these L-toned polar questions the L tone itself is not being realized as expected. Instead, the fall in F0 that is expected due to the L tone is suspended and the final syllable is realized with a level F0 instead.

Table 20

F0 Contour of Final Syllables in Questions in 2014 Elicitation Data

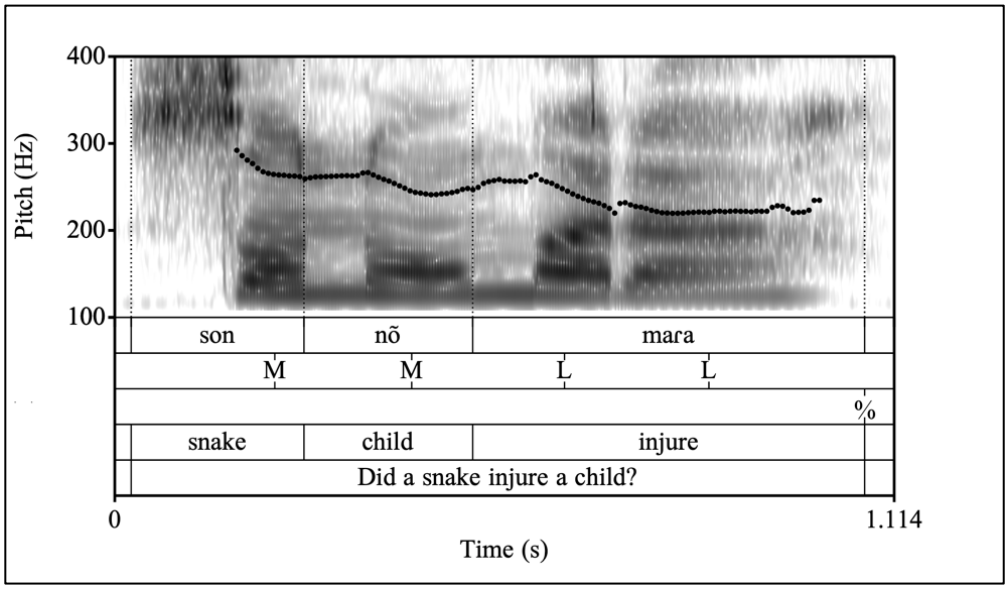
	Total	**Falling**	**Level**
Polar Questions			
L	26	14	12
M	3	3	3
Wh-*questions*			
L	19	15	4
M	0	0	0

The *wh*-questions, on the other hand, are realized consistently with the 2016-2017 data. All of these questions have a final L tone and all but four are realized with a falling F0. The four exceptions are cases in which F0 is near the bottom of the speaker's range before the end of the

utterance. The *wh*-questions do not have a specific *intonational* contour, as expected, but the polar questions have a different intonational contour than is found in the later data. While some of the level-toned polar question tokens are also cases in which the F0 of the final syllable is low and level, in the majority of tokens it is phonetically similar to an M tone. Figure 47 below shows one example. Here, S1 has not reached the bottom of her range, but there is no fall in F0 on the final L-toned syllable. Instead, the final syllable is realized with an extremely level F0.

Figure 47

Polar question with a final lexical L tone and no final fall in F0 (S1, F)



There is also individual variation that is not present in the 2016-2017 data, where the speakers who participated were mostly consistent with each other. In the 2014 data, S1 and S2 produce polar questions differently. (S1 did not participate in the 2016-2017 tasks.) The raising of final lexical L tones is driven primarily by S1. Ten of the twelve raised tokens were produced by S1, while only two were produced by S2. On the other hand, S1 and S2's productions of *wh*-questions are comparable. Table 21 shows a summary of the F0 contours by speaker.

My first analysis of the question prosody of Bobo, which was based on the 2014 data, proposed

that there is an H% tone whose tone target is superimposed on the tone target for the L lexical tone, resulting in an intermediate realization of F0 (Sherwood, 2016). In light of the 2016-2017 data, however, this analysis is not satisfying. The 2016-2017 data provided evidence for an L% that is concatenated to the end of the string of lexical tones, rather than superimposed on the lexical tone

Table 21

F0 Contour of Final Syllables in Questions in 2014 Elicitation Data by Speaker

	Polar Questions			***Wh*-Questions**		
	Total	*Falling*	*Level*	*Total*	*Falling*	*Level*
S1						
L	16	6	10	9	7	2
M	1	0	1	0	0	0
S2						
L	10	8	2	10	8	2
M	2	0	2	0	0	0

of the final syllable. This means that a superimposed H% tone would have different rules governing its realization than the concatenated L%. Additionally, there is some indication of H% boundary tones that are concatenated rather than superimposed—specifically, in some of the *wh*-questions (discussed in the previous section) and in some of the polar questions in the spontaneous speech tasks (discussed in the following section). If these concatenated H% are confirmed, then this difference in realization must be explained as well. The data is consistent with S1 raising the pitch register of the final syllable in these questions, but the underlying mechanism is as yet unknown.

The prolongation and breathy phonation is consistent between all of the elicitation tasks, suggesting that these are in fact the most important cues to polar questions when a particle is not present.

4.4.5 Questions in the spontaneous speech tasks

Since the spontaneous speech allowed the speakers the most freedom in how they produced questions, this data is the most variable. Each pattern found in the elicitation task is consistent with

at least some of the questions produced during the spontaneous speech tasks. However, morphosyntactic means of marking questions is much more common in the spontaneous speech tasks, meaning that there are few tokens in which the speakers rely on prosody alone to signal that a question is being asked. I discuss the different types of question formation in order of frequency of occurrence below. For the spontaneous speech data, sentences are identified as questions if this is how the speaker who assisted with the translation interpreted them.

The sample of the spontaneous speech that was examined closely contained 39 polar questions. The majority of these questions are formed using the final particle *kɛ¹* or the final negation particle *=ga⁵/ŋa⁵*. S3 heavily prefers using *kɛ¹*; of the 22 polar questions that he produced in this sample, 14 are formed using *kɛ¹*. None of the other speakers used *kɛ¹* in this way.[32] An example of *kɛ¹* used to form a question is shown in Figure 48. When S3 uses *kɛ¹* to form a question, there is no specific question prosody. Breathiness is sometimes present, as can be seen in Figure 48, but this is also true of statements. Prolongation is difficult to assess without a control for comparison, but there are no cases where it is clearly present. The L% boundary tone would not be observable if it were present, since *kɛ¹* is an L-toned particle. An H% or suspension of the falling F0 contour would be observable, but S3 does not use these cues in the elicited questions and there are also no examples of these cues in his spontaneous speech data either.

kɛ¹ is not solely used to mark questions. It is also used by both S3 and S5 in statements. Its meaning is difficult to ascertain; when assisting with the translation, speakers provided meanings that depended on the context. Figure 49 is an example from a narrative task where S5 appears to use it for "emphasis." Although there was another speaker present and listening, he did not respond,

[32] Only S1, S2, and S3, and S5 participated in the interactive spontaneous speech tasks. S4 did not.

Figure 48

Polar question from an interview using the final particle kɛ[1] *(S3, M)*

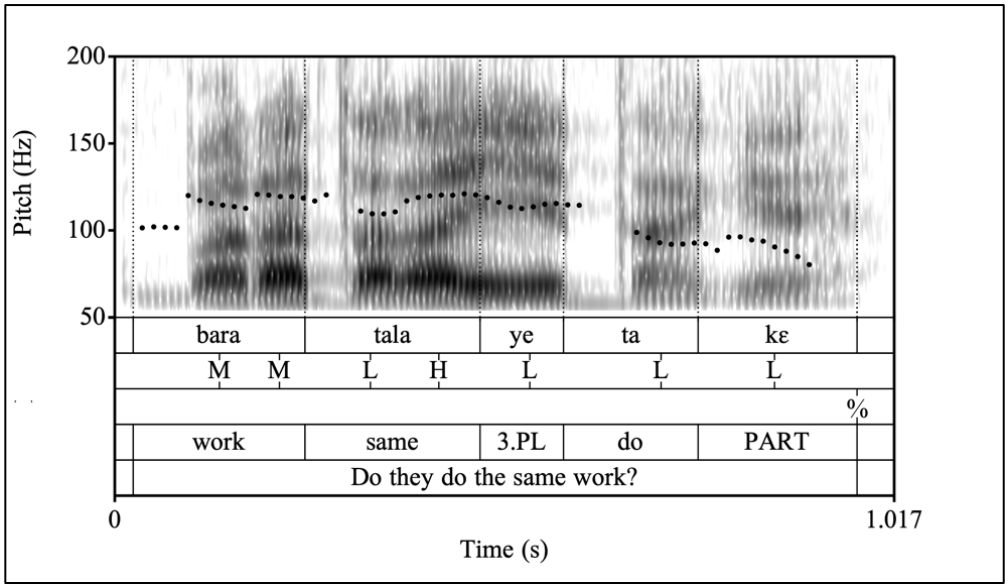
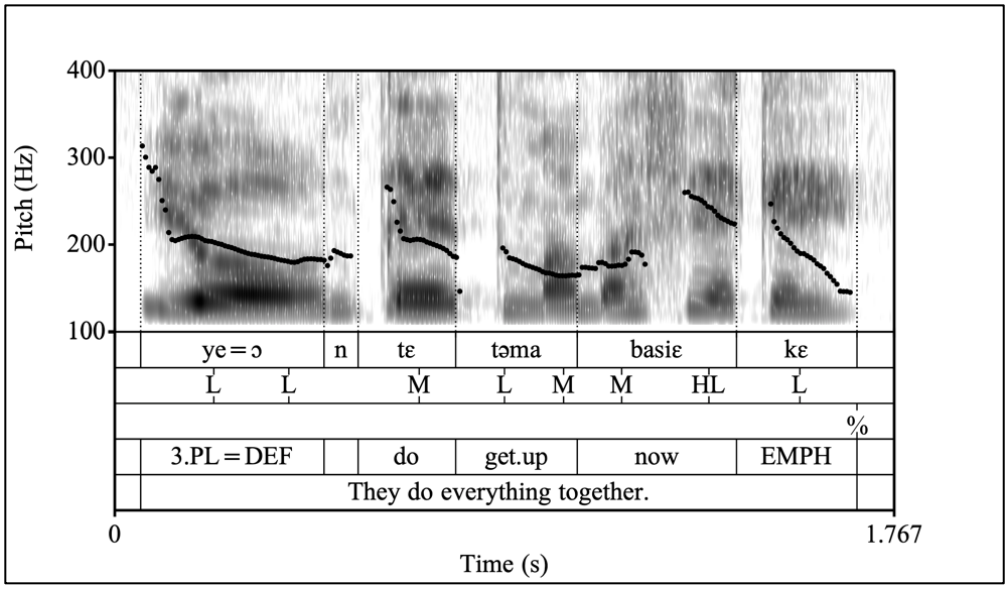


Figure 49

Final particle kɛ[1] *used in a statement (speaker S5, female)*



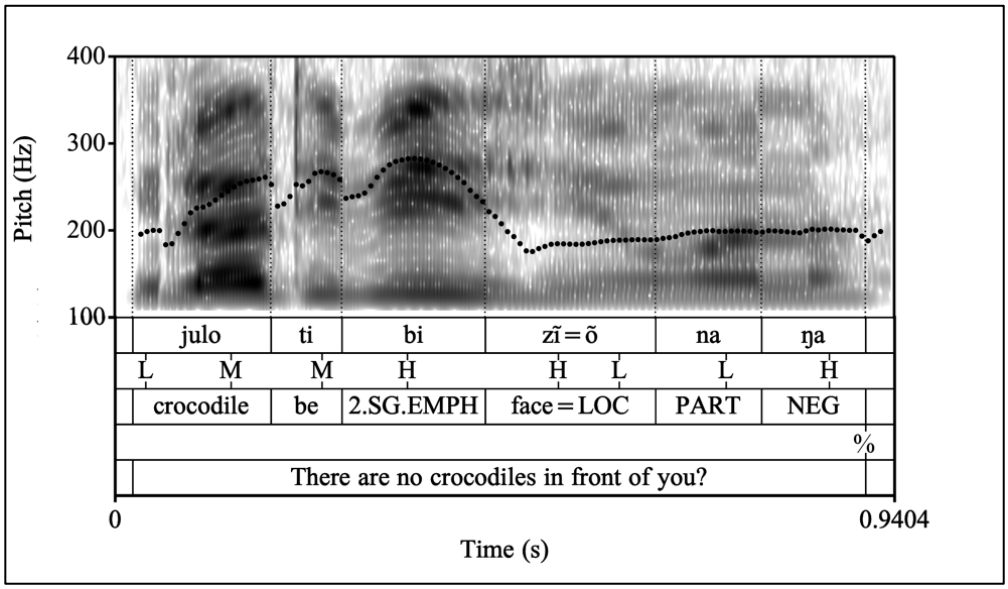
suggesting that the purpose of the particle is at least not always to prompt a response. There are no apparent differences in the prosodic properties of statements formed with *kɛ[1]* and questions formed with *kɛ[1]*, which raises the question of how they are distinguished. In the interactive speech tasks,

kɛ^{1} does prompt a response (usually an affirmation) from the other speaker. Whether it is identified as a question from the discourse context or some unknown property of the utterance itself is a question or future research.

S2 frequently uses the final negation particle = *ga*^{5}/*ŋa*^{5}. Of the 15 polar questions she produced, seven use this particle. The F0 characteristics are variable; two have falling F0, four have level F0, and three having rising F0. This is typical of negated statements ending in = *ga*^{5}/*ŋa*^{5} as well. At the end of a phrase, especially a longer phrase, the H tone is often lowered. As with *kɛ*^{1}, breathy phonation and prolongation of the final segment is also not consistently present, meaning that these are not reliable cues to the fact that these are questions. As with the questions formed with *kɛ*^{1}, it is not clear what indicates to the hearer that a question is being asked—although in this case, S2 asserting that S3 cannot see something would be strange given the context. Figure 50 shows an example. The speaker who assisted with the translation (S3 or S4) generally included the negation in the French (e.g. "*il n'ya pas des crocodiles devant toi?"* for Figure 50.)

Figure 50

Polar question ending in the negation particle =ga^{5}/ŋa^{5} *(S2, F)*



There are only 13 polar questions in the spontaneous speech tasks that do not have a phrase-final particle. The prosody of these questions is highly variable and difficult to generalize. Three of these questions do not unambiguously show evidence of question prosody. The remainder can be divided into two categories: Those that potentially have all three question prosody characteristics, and those that possess only a subset (sometimes only one). One complication here is that because of the relative frequency of lexical L tone in phrase-final position, there are only four polar questions in the spontaneous speech data that end in an H or M tone instead—too few to assess how frequent L% is in this data. What can be said is that it is sometimes but not always present. Two examples are shown below; Figure 51 has an L% boundary tone, while Figure 52 clearly does not.

Two of the polar questions in the spontaneous speech have an unexpected rise in F0. Although this is too few tokens to draw firm conclusions about, it is intriguing because both occur in a similar context: In the map task, one speaker has told the other speaker information that is contradicted by their own map, and they react in surprise. Figure 53 shows one of these questions. An unexpected rise in F0 at the end of a phrase is unusual in Bobo; phonetic variability in the expression of tone at the end of phrases tends to be in the form of falls in F0, e.g. a slightly falling rather than completely level mid tone. This is therefore likely to be intentional and connected in some way to the speaker's surprised affect during the production of these questions. The rise in F0 appears like a concatenated H% would be expected to. Figure 53 also shows how S3 has produced the question with a higher F0 range. The H tones reach 200Hz, which does not happen very often in his data.

Figure 51

Polar question with a final lexical H tone and an L% boundary tone (S2, F)

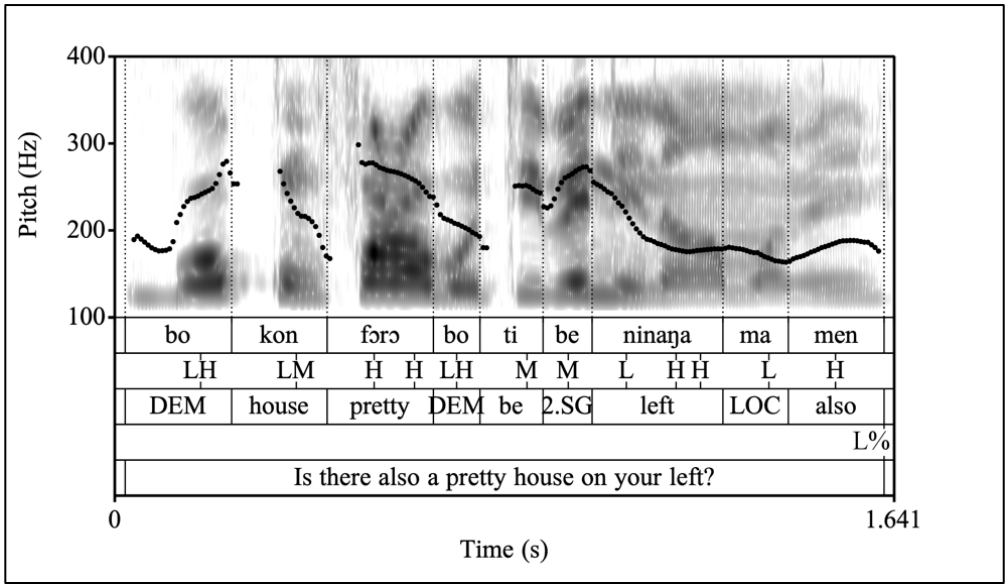


Figure 52

Polar question with a final lexical H tone and no L% boundary tone (S2, F)

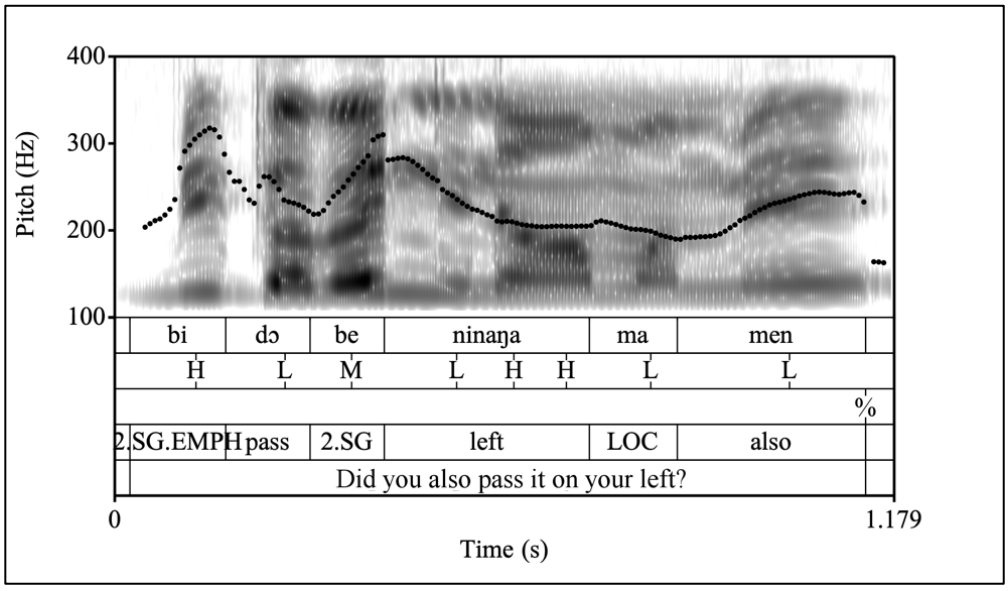
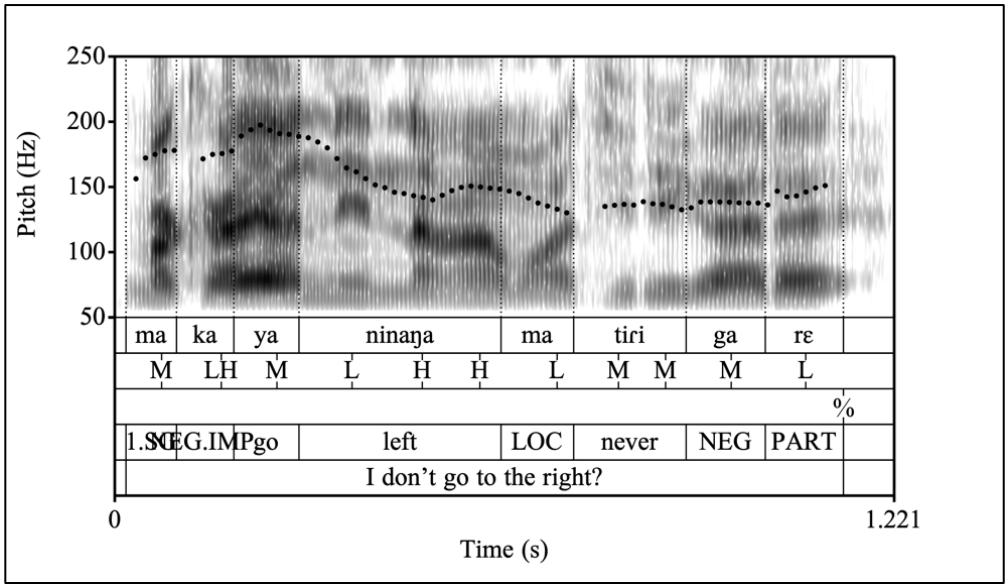


Figure 53

Polar question with a final lexical L tone and a final rise in F0 (S3, M)



Breathy termination and prolongation also occur in the polar questions in the spontaneous speech tasks, consistent with the elicitation tasks. Eight of the polar questions that do not end in a particle have breathy utterance termination, while seven have a prolonged final vowel. In five of the tokens, these cues co-occur. This reflects a pattern found in the elicitation tasks, which is that prolongation and breathy utterance termination appear to be associated with each other more strongly than they are associated with L%. Figure 54 shows an example of one of these tokens. The prolongation is not as extreme as in the elicitation tasks but is still quite audibly noticeable.

There are 14 *wh*-questions in the spontaneous speech tasks: eight with a final L tone, four with a final M tone, and two with an unknown final tone. Like the polar questions, the expression of the *wh*-question is difficult to describe in general terms because of the amount of variation. One of the M-toned tokens has an unexplained final fall in F0, but this is in reported speech and the grammatical tone of reported speech has not yet been described. Since the grammatical tone is unknown, this fall in F0 cannot be assigned to the presence of an L% boundary tone. The remainder

Figure 54

Polar question with prolongation and breathy termination (S3, M)

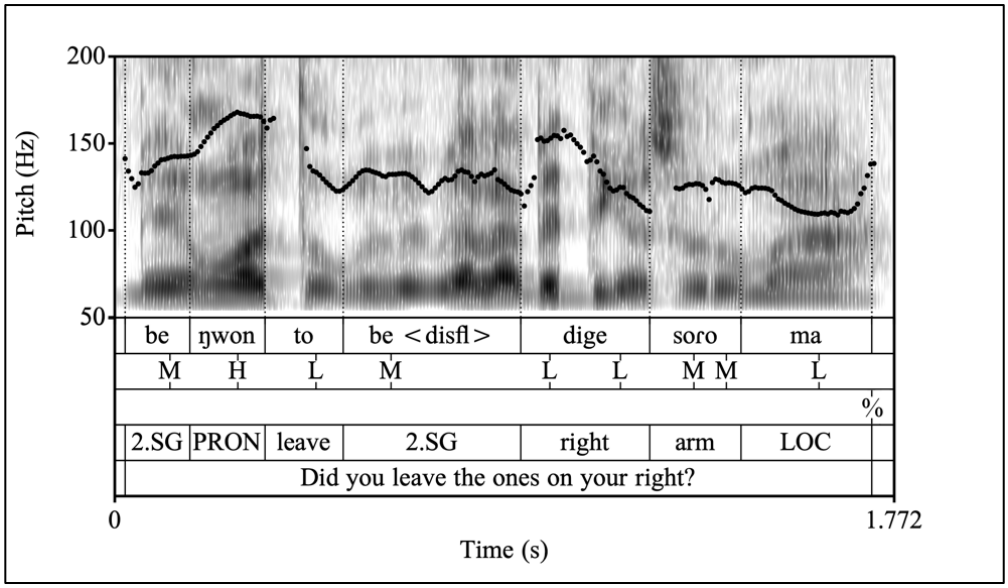
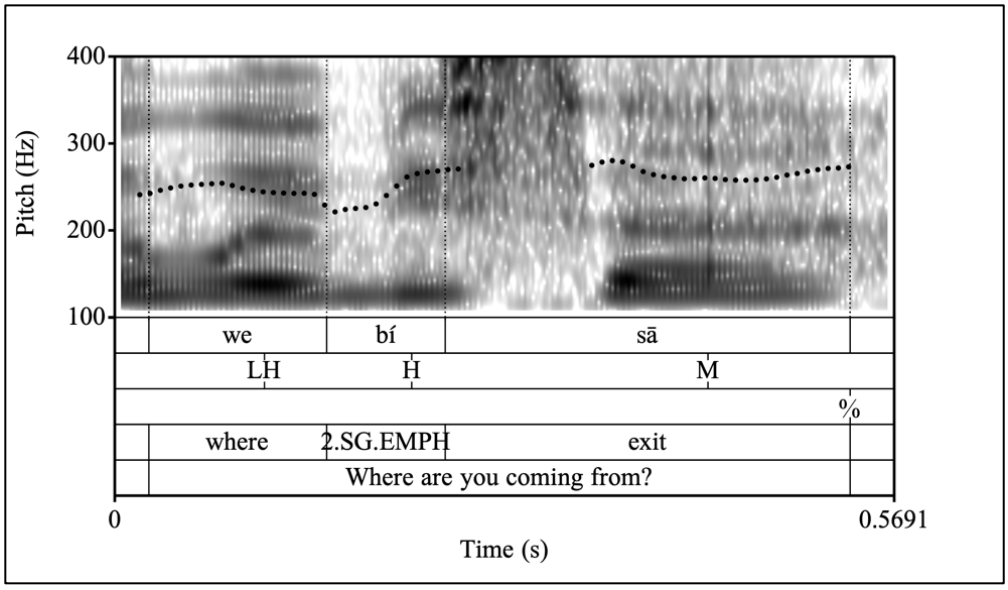


Figure 55

Wh-*question from an interview task ending in a final raised M tone (S1, F)*



of the M-toned tokens are realized with a level F0. Breathiness is frequent, with eleven of the tokens having some degree of breathy termination. Prolongation is gradient and inconsistent. The two questions produced by S1 have a raised F0 in the final syllable, consistent with her elicitation

data. The particularly interesting example is given in Figure 55, where it is an M tone that is raised. Here, the M-toned verb s*a*^{3} 'exit' is produced with a similar F0 as the previous H tone.

The questions in the spontaneous speech data show that there are multiple question-marking strategies. A question can be marked morphosyntactically, using a particle, or it can be marked using one or more of the prosodic cues that are available: breathy phonation, prolongation, and L%.

4.5 Negated Statements

Morse (1976, p.99) briefly describes negated sentences as having higher overall pitch than non-negated sentences. If this is true, then this would be one aspect of Bobo's intonation that could not be analyzed as the result of discrete intonational tones. Rather, it would be a global phenomenon affecting the entire phrase or utterance. There is reason to be skeptical of Morse's analysis, however. She did not use recordings, so her data is not available, and no F0 measurements were made. There is also an alternate explanation for a perception of higher pitch: In Bobo, negation is marked with a clause-final negative particle =*ga*^{5}/*ŋa*^{5}, which has a lexical H tone. Additionally, both she and Le Bris and Prost (1981) claim that negation is (or can be) marked with an H tone suffixed to either the subject or the TAM marker.

To investigate these claims about the prosody of negated statements, I examined the data described in §2.2.2.4. This data set consists of 22 pairs of matched negated and non-negated statements (44 total), which are identical except for the presence of the negation particle. The only TAM marker in Bobo is *ma*^{3}*ne*^{3}, which marks the past. Six of the sentence pairs were constructed with this TAM marker. 218 tokens were recorded, 39 of which were discarded due to a recording error and four of which were discarded due to a speaker error. The remaining 175 were included in the analysis in this section.

First, I evaluated the claim that an H tone is added to subjects or the TAM marker. In the majority of sentence pairs, neither the subjects nor the TAM show a difference in F0 between the non-negated or negated sentence; there is no addition of an H tone. Figure 56 and Figure 57 illustrate this with a sentence pair including *ma³ne³*. The F0 contour of both the subject *sio¹* 'horse' and *ma³ne³* are realized identically between the two sentence types; there is no addition of an H tone. Figure 58 and Figure 59 show a pair (declarative and negated sentence) without the TAM marker. Here there is also no difference between the realization of the subject between the non-negated and negated sentence.

However, in minority of cases there is indeed a rise in F0 on the final vowel of the subject. An example is shown in Figure 60 and Figure 61 with the mid-toned subject *nã³nõ³* 'chicken'. A total of 19 of the negated sentences show this unexpected rise in F0, out of 62 where the F0 of the subject could be tracked and did not have a final H tone.

Figure 56

Non-negated sentence with an L-toned subject and TAM marker (S3, M)

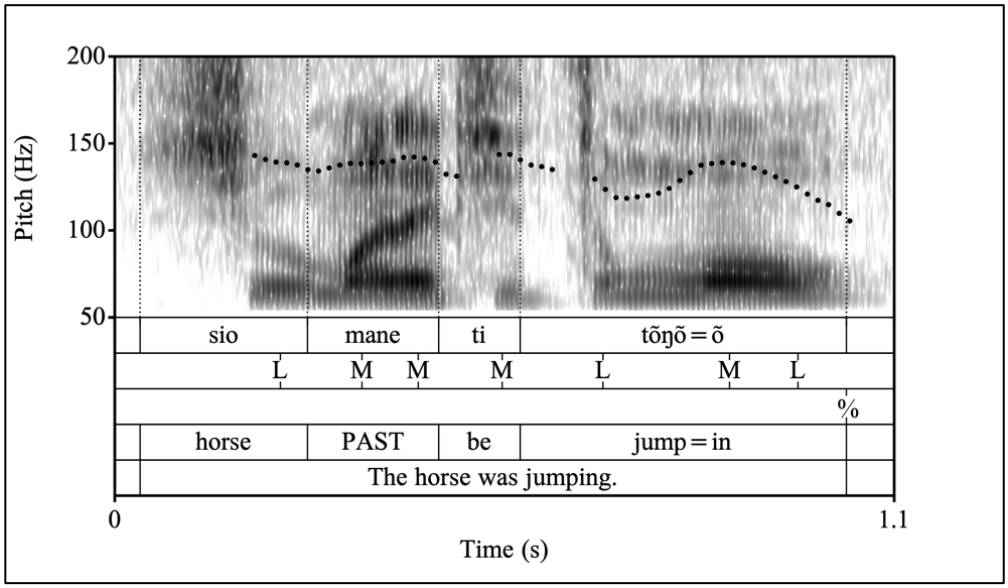


Figure 57

Negated sentence with L-toned subject and TAM marker (S3, M)

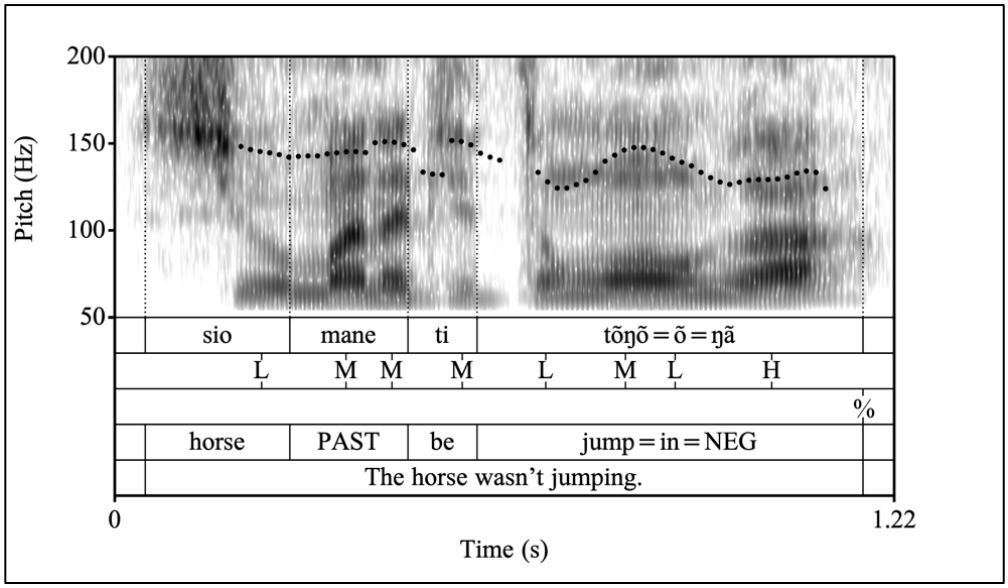


Figure 58

Non-negated sentence with an L-toned subject (S3, M)

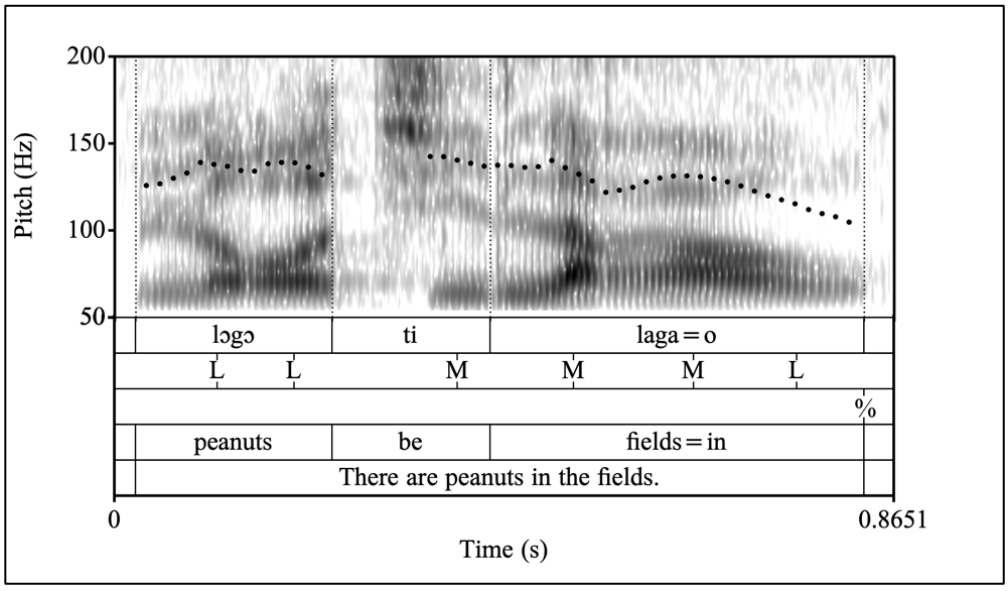


Figure 59

Negated sentence with an L-toned subject (S3, M)

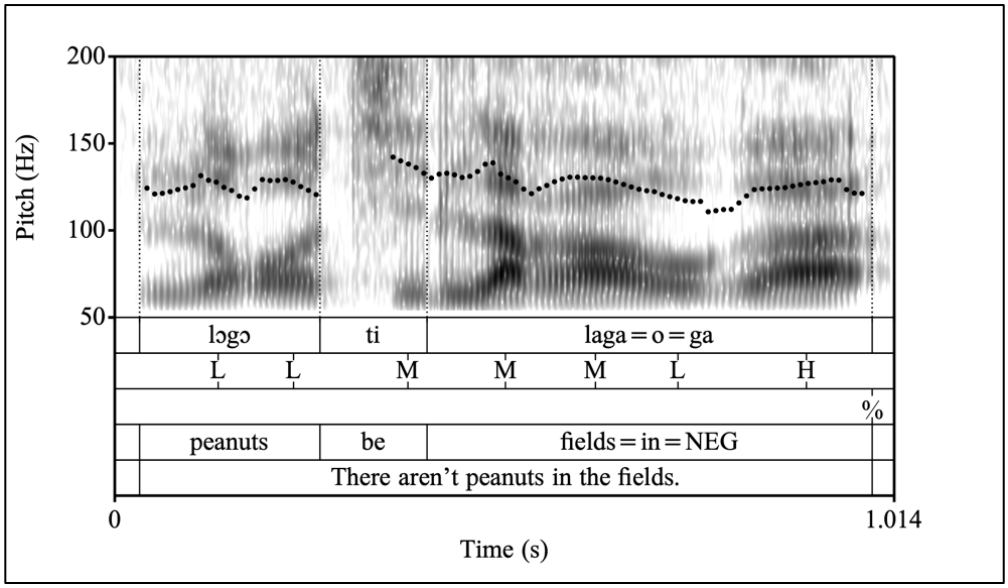


Figure 60

Non-negated sentence with an M-toned subject (S3, M)

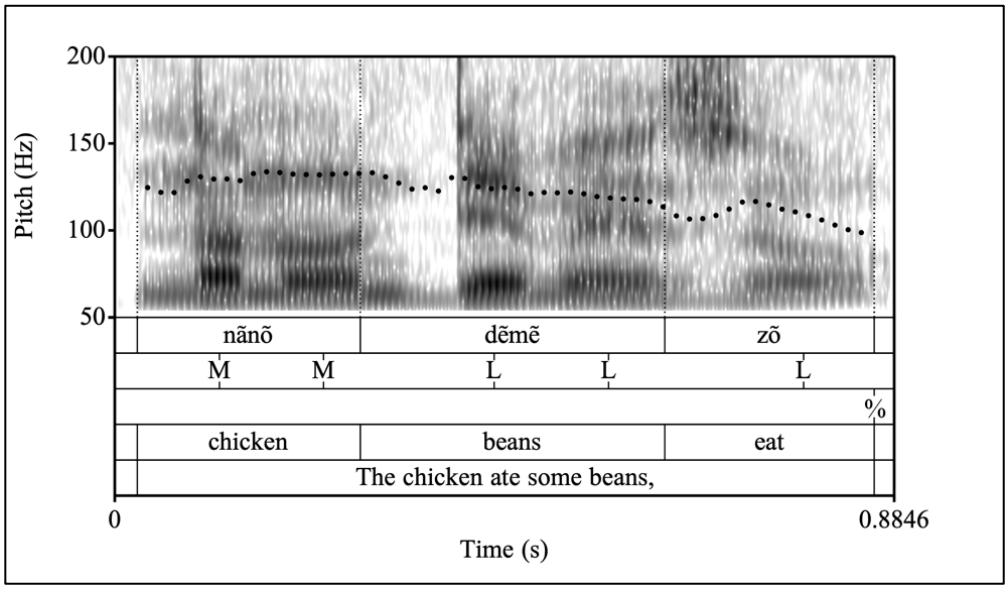
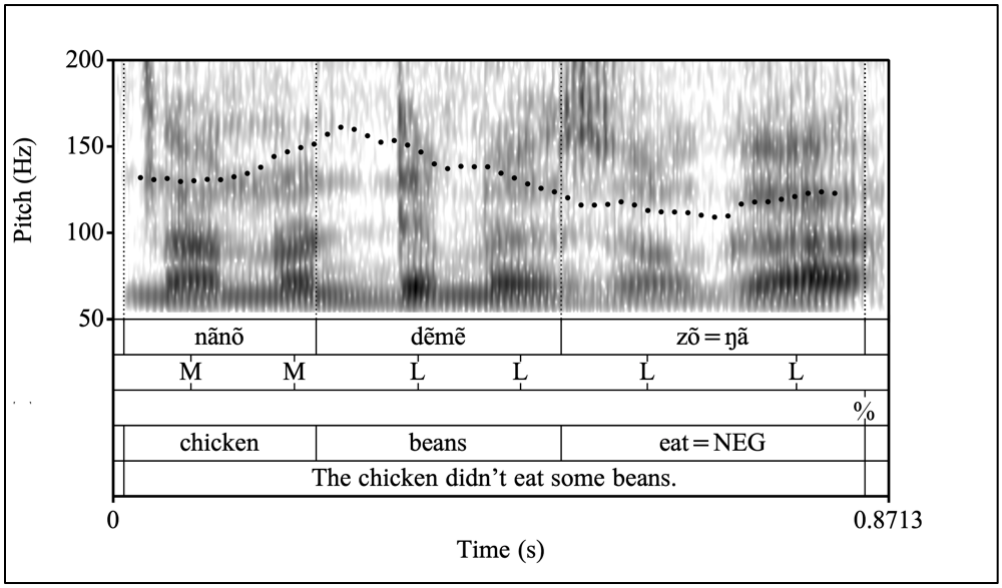


Figure 61

Negated sentence with an M-toned subject (S3, M)



There is no apparent pattern behind which subjects surface with a rise in F0 and which do not. The rise in F0 does not occur consistently on a particular subject type (noun or pronoun), on particular words, on a particular tonal melody, or with a particular speaker. There is clearly a modification of the tonal string in tokens like the above; the rise is sharp. The rise is not always so salient, however, and without being able to identify the context that triggers it, it is not possible to say whether it is an H tone suffix or an intonational tone. Neither Morse (1976) nor Le Bris and Prost (1981) discuss variation in the expression of the H tone suffix they claim is used to mark negation. Their description implies that the suffix is a grammatical marker of negation. This does not seem to be the case since it does not occur consistently, even within an elicitation context where variation is minimized.

Second, I evaluated the claim that negated sentences have raised pitch. I excluded the tokens in which there is possible H tone suffix or H-toned subject and then measured the average F0 of the utterance of the sentences that remained. The final two syllables of the non-negated

sentence were excluded, in order to control for potential effects of the phrase boundary, and the final three syllables (including the negative particle) were excluded from the negated sentences. Examples 23 and 24 below show which portions were included in the measurement (underlined):

(23) *ma³* *ma³ne³* *nu³ma³ne³* *za¹*
1.SG PAST children see
"I saw the children."

(24) *ma³* *ma³ne³* *nu³ma³ne³* *za¹* *ga⁵*
1.SG PAST children see NEG
"I didn't see the children."

The results by speaker are given in Table 22. For all three speakers, the mean F0 of non-negated and negated sentences are within 4Hz of each other, which is not a meaningful difference. Negated sentences do not have raised F0. The apparent H tone on the subject of some negated sentences is the only prosodic characteristic that differentiates the sentence types. However, since this does not occur on all negated sentences, and instead only occurs on a minority of them, it being a characteristic of negated sentences is questionable. At this point in time it does not appear that negated sentences have a specific prosody.

Table 22

Mean F0 (Hz) of Non-Negated and Negated Sentences

	Non-Negated		***Negative***	
	M	*SD*	*M*	*SD*
S2	217.3	11	217.9	14.2
S3	135.4	8.2	136.4	10.2
S4	106	7	110.2	5.6

Note. M=mean, SD=standard deviation.

4.6 Prosodic Phrase Types

I have not found evidence of any unique, obligatory characteristics of the IP in Bobo. The elicited declarative statements have minimal intonation: There is no evidence of pitch accent, phrase accent, or boundary tone. Downstep or downdrift may be a property of the IP, but due to

the relative infrequency of H tones, it is not observable in most prosodic phrases. It is also not clear to what extent it is obligatory; it does not occur in some phrases where it is expected, but the reasons why not need more investigation. Elicited declarative statements do have final lengthening, which is observable in mean durations of phrase-medial versus phrase-final segments. However, an individual token might not have observable final lengthening because the degree of lengthening is so variable. This means that there are as yet no criteria by which one can establish the absence of an IP boundary. The ability to do this is necessary in order to identify smaller phrase types—otherwise, any additional prosodic characteristics that are associated with that smaller phrase type could also be optional properties of the IP. Therefore, the IP is the only identifiable prosodic phrase type at this point.

4.7 Summary and Discussion

The phonetic and phonological properties of the IP in Bobo that have been identified are final lengthening, non-modal phonation, boundary tone, and downstep or downdrift. These properties are related to specific utterance types: Declarative sentences are the default type and do not have any markers of the boundary apart from final lengthening and non-modal vowel phonation. Negated statements do not have any specific prosody that distinguishes them from non-negated statements. Polar questions and *wh*-questions are both marked through specific phonetic properties, though even those are optional and depend on context and the speaker. Polar questions show three of the four characteristics of "lax question prosody" (Rialland, 2009), namely the prolongation of the final vowel, breathy utterance termination, and falling F0, while *wh*-questions share some prosodic properties with polar questions (vowel prolongation and breathy utterance termination) and some with declarative statements (lack of L% boundary tone). One open question is the mechanism behind the lengthening of the final vowel in questions. This lengthening has been

described in various ways for languages having this characteristic: as prolongation or lengthening of the final vowel (e.g. Rialland, 2009) or as prolongation or lengthening of the final segment (e.g. Cahill, 2016; Genzel, 2018). If the underlying mechanism is the same as final lengthening, just expressed to a greater degree, it should be the final segment that is lengthened regardless of what type of segment it is. Final lengthening does not "skip over" a phrase-final segment to affect the segment preceding it (Byrd & Saltzman, 2003). In a language like Bobo, which has only CV syllables, lengthening of the last vowel in a phrase and phrase-final lengthening make the same prediction. It will be difficult to determine whether the prolongation of the final vowel is a different degree of phrase-final lengthening or whether it is the result of a different underlying process. To my knowledge, there have been no quantitative studies looking at the degree of lengthening of final consonants in languages with this question prosody either within a single language or across languages.

The properties listed above vary in the consistency of their occurrence and the extent to which they occur when they do occur. This raises the question of which of these properties are a part of the underlying phonological representation of prosodic phrases. Under the AM framework, the prosodic phrase is a phonological unit, but it does not follow that every property associated with its boundaries must also be a realization of phonological categories (for example, various properties could be related to speaking style of pragmatic differences). The best candidates for phonological properties are those that occur consistently since this consistency needs an explanation. In Bobo, the most consistent prosodic cues associated with a phrase boundary are those that are used to mark *yes-no* questions—namely the extreme prolongation, breathy phonation, and L% boundary tone. They therefore seem to be a part of the grammatical (phonological) representation of questions. In spontaneous speech, they seem to trade off with other,

morphosyntactic means of marking questions (particles and negation) although this question needs to be further examined. The lowering of an H tone after an L tone also occurs fairly consistently; the best explanation is that this is a phonological process of downstep, although there remain many questions about its domain of application and whether the lowering of H tones after M tones is due to the same process.

Within the AM framework, the only phonological categories marking a boundary are tonal units (such as phrase accents and boundary tones). There has not been much consideration given to other phonetic correlates of boundaries within the AM framework, such as voice quality or temporal properties. The AM framework does not provide a straightforward way to include the breathy voice and prolongation used to mark *yes-no* questions into the prosodic hierarchy (beyond simply describing their existence). The presence of a juncture, expressed through final lengthening and pauses, can be the only indicator of the phrase boundary in Bobo. Since phrase boundaries were identified by the presence of a pause, it is not possible to assess how frequently pauses accompany a boundary. The *length* of the pause is extremely variable, as is the degree of final lengthening. This is completely expected; the strength of a juncture is influenced by multiple factors, many of which are non-phonological, such as speech planning and discourse context (see e.g., Krivokapić, 2007; Swerts and Geluykens, 1994). In order to accurately describe and explain the prosodic structure of languages like Bobo, more consideration needs to be given to the phonetic and phonological characteristics of phrase-level prosody in languages that do not make as much use of boundary tone as European languages.

The phrase-level prosodic hierarchy in Bobo likely consists only of the IP. This analysis could change through a better understanding of downtrends; if these are obligatory characteristics of the phrase, pitch reset could be used as diagnostic for distinguishing between an IP boundary

and a lower level phrase (if one exists). Otherwise, since there are no properties that can be used to consistently identify the presence of an IP boundary, there is no way to differentiate between an IP boundary and a hierarchically lower category.

Chapter 5

Focus Marking

5.1 Overview and Research Questions

In this chapter, I investigate focus marking in Southern Bobo Madaré, specifically whether focus is marked through prosody, through morphosyntax, or both. The motivation for including focus marking in this project is that, cross-linguistically, focus is frequently marked through prosodic prominence. Since there is no prior work on focus marking in Bobo, I also consider the possibility that focus is marked through morphosyntactic means instead of, or alongside, some form of prosodic prominence. Morphosyntactic focus marking is common in African languages (Fiedler et al., 2010; Aboh et al., 2007), including Mande languages (e.g. see Creissels, 2015 on Mandinka; Prokhorov, 2014 on Bamana; Diagana, 1987 on Soninke; Nikitina, 2018 on Wan; Bearth, 1992 on Tura). In Mande languages, morphosyntactic focus marking is frequently optional (Schreiber, 2008). The use of prosodic prominence to mark focus is much less well-described and it is not known how common it is within the Mande language family.

Before continuing, it is necessary to discuss what is meant by *focus, focus marking,* and *prosodic prominence*. *Focus* is a term with a proliferation of definitions (see Gundel, 1999 for some discussion). The definition that I will use here follows Jackendoff (1972) and many others (e.g. Rooth, 1992; Schwarzschild, 1999; Selkirk 2008; Büring 2010, 2016; Zimmerman & Onea 2011) in treating *focus* as a semantic or syntactic phenomenon related to the information structure of the discourse. The precise definition of focus varies by theoretical approach but generally relates

to the notion of new versus given information; a focused constituent is one that is imparting information that is expected to be new to the hearer. There is considerable debate about what types of focus exist, whether these types are universal, and what specific criteria determine when a constituent is focused. There are no criteria that can accurately and comprehensively account for focus phenomena cross-linguistically. However, in order to identify focus marking, I must be able to identify at least some focused constituents. For the purpose of this project, I have restricted my attention to two prototypical contexts in which constituents are expected to be focused: when they are the answer to a *wh*-question and when they are a correction to a previously mentioned constituent. These two contexts are included under most theoretical definitions of informational (*wh*-question) and contrastive (corrective) focus (see discussion in Güldemann et al., 2015).

Focus is crucially distinct from *focus marking*, which is the encoding of focus. While focus is often (but not always) taken to be a universal property of information structure, focus marking is not. Focus marking strategies vary by language (e.g. see overviews in Aboh et al., 2007; Büring, 2010; Zimmerman & Onea, 2011). A common strategy is to use some form of prosodic prominence, whether alone or in conjunction with morphosyntactic focus marking.

Prosodic prominence is distinct from focus and focus marking. A universal definition of prosodic prominence is difficult, since the term has been used to describe varied phenomena, including linguistic functions (e.g. prosodic focus, emphasis), phonological representations (e.g. pitch accent), and phonetic properties (e.g. lengthening, intensity, pitch). Wagner et al. (2015), in their discussion of definitional issues surrounding prominence, suggest that researchers adapt a definition of this form: "<We> say a <linguistic entity> is prosodically prominent when it <stands out> from <its environment> (by virtue of <its prosodic characteristics>)." (p. 3) Rather than proposing a single, universal definition of prominence, they propose that this variables within this

template be modified to precisely define the phenomena under investigation, e.g. by specifying which prosodic characteristics are of interest. This is not possible for Bobo since those characteristics, if any exist, are unknown. For the purpose of investigating focus marking in Bobo, I will specify that the linguistic entities in question are the syntactic constituents that are put under different focus conditions in the data, which includes both single lexical items and longer constituents such as NPs. The relevant prosodic characteristics are post-lexical, since the investigation is concerned with phrase-level rather than word-level prominence. Importantly, prosodic prominence is focus marking if and only if it encodes focus. A syntactic constituent can be prosodically prominent for reasons other than focus.

Cross-linguistically, prosodic prominence is frequently achieved through changes in pitch, duration, or amplitude of the prominent constituent, or through phrasing, where a prosodic boundary can be placed before or after a prominent constituent (Jun, 2005b; Büring, 2010). The majority of theoretical work on languages that mark focus through prosodic prominence has concentrated on languages that use pitch accent, such as English and German (for a summary, see Büring 2016). In African languages, pitch accent has not been widely reported. Research on focus marking in African languages has been concentrated on languages with morphological focus marking. However, some African languages have been reported to use phrasing (Güldemann et al., 2015).

Languages with prosodic focus marking can be divided into two types: those which mark the head of the prominent constituent and those which mark the edge of the prominent constituent (Jun, 2005b). A head-marking strategy is found in English and other pitch accent languages, which use pitch accent on the phrase head of a focused constituent to mark focus. Korean is an example of a language that uses edge prominence; there is no pitch accent and the edges of phrases alone

mark prominent constituents. Whether a language has head-prominent or edge-prominent focus marking cannot be predicted based on its word-level prosody. One might predict that edge-marking would be more common in lexical tone languages, so as to avoid prosodic focus marking obscuring the lexical tone of the focused item, but there are counterexamples. Mandarin and Cantonese, for example, use head-marking strategies. Focused constituents in these two languages have greater pitch range, amplitude, and duration (Peng et al., 2005; Wong et al., 2005). It has been proposed that the common principle behind all focus-marking strategies is that they make the focused item maximally prominent (Truckenbrodt, 1995; Büring, 2010), whether this is through pitch accent, through other means of placing of prosodic prominence on the focused constituent, or through morphosyntactic marking.

Prosodic focus marking is not universal and cannot be assumed for Bobo. For example, Rialland and Robert (2001) report that focused constituents are not marked prosodically in Wolof; prosodic focus marking is unnecessary since Wolof has a rich system of morphological focus markers. However, morphosyntactic focus marking does not preclude prosodic focus marking. There are relatively few studies of prosodic focus marking in languages that typically mark focus morphosyntactically, but it is clear that it is not always one or the other; both can be used. For example, in Chickasaw, which has morphological focus marking, contrastively focused elements are still produced with raised pitch and longer duration. Focus also has an effect on the placement of prosodic breaks (Gordon, 2008).

Focus marking itself also cannot be assumed. An increasing number of prosodic descriptions are reporting a lack of focus marking in expected contexts, especially in African languages. In some African languages, focus marking is optional or restricted to certain syntactic positions. For example, Hausa has both *ex-situ* and *in-situ* focus marking, with *ex-situ* focus

involving the fronting of the focused element. According to Hartmann and Zimmerman (2007), *ex-situ* focus is optional and its focus interpretation can be ambiguous. *In-situ* focus is also optional; it can be marked with a focus particle or not. When it is not marked with a focus particle, there is no focus marking at all—neither morphosyntactic marking nor prosodic marking. Buli, a Gur language, shows a similar pattern. Buli has morphological focus marking that can only occur in certain syntactic positions, resulting in significant ambiguities regarding the type and scope of focus, which are not distinguished prosodically (Schwarz, 2009). For some other African languages, researchers have found no focus marking at all. Zerbian (2007) reports that this is the case in Northern Sotho, a Bantu language of South Africa. She used question-and-answer dialogues in order to elicit a variety of focus structures, none of which showed evidence of focus marking. A follow-up perception experiment confirmed that speakers of Northern Sotho did not distinguish between statements that were expected to have focus marking and those that were not. In addition, Downing and Hyman (2016) argue that some constructions in Bantu languages that have been interpreted as focus marking are only correlated with focus, rather than directly encoding it. They go so far as to say, "As far as we know, focus does not directly condition the analogous use of obligatory prosodic prominence or prosodic restructuring in Bantu or African languages" (p. 16). "Analogous" here is referring to the well-studied phenomena of prosodic focus marking in European languages.

The non-universality of focus marking raises methodological and analytical questions: Focus marking and other features dependent on information structure can be difficult to elicit (Himmelmann & Ladd, 2008). If there is no evidence of focus marking in the data, it is possible that this is a valid result but it is also possible that it is due to a problem with the methodology. Many focus-eliciting tasks require speakers to read a prompt and produce natural-sounding

sentences in response, which does not always come easily—especially if the language does not have a written tradition or if speakers are not literate. They also often require the speaker to imagine themselves in a specific discourse context and act as if they are a part of it, which some speakers cannot do. There are ways to avoid or minimize these problems, but none are perfect. One can identify speakers that are comfortable with this type of task, or work with speakers who have become comfortable over time, but this limits the number of speakers that can participate. Using spontaneous speech avoids some of the problems of elicitation, but the lack of controlled comparisons in spontaneous speech can make it difficult to draw conclusions. If the language does not exhibit focus marking in expected contexts, then the researcher has the difficult task of proving a negative, namely, of demonstrating that this lack of focus marking in the data is not due to methodological problems. The best way to address this issue is to use tasks of different types so it less likely that a problem with a single task is leading to a false result. This is the approach taken here. I use three different tasks to elicit focus marking: An elicitation task, a questionnaire task, and a map task.

5.2 Analysis

The data consists of recordings of elicited question-and-answer pairs, the Questionnaire on Information Structure (QUIS) and a map task (Skopeteas et al., 2015). This data is described in detail in §2.2.1, §2.2.3 and §2.3. Each task is designed to elicit key words or phrases under different types of focus. Although the data are of different types, they were analyzed in the same way. First, focused constituents within the data were identified based on the criteria discussed in the previous section. I adopted the division of focus into the following three types (cf. Mücke & Grice, 2014):

- *Broad (or wide) focus* – When a constituent larger than a word is focused. In languages

like English, in which focus is marked by a discrete pitch accent, focus marking on one word within the focused constituent is used to mark focus of the entire constituent (for different views on how the placement of pitch accent is determined in these cases, see Selkirk, 1995; Büring, 2005, 2016).

- *Narrow focus* – When a single word is focused.
- *Contrastive focus* – When the focused constituent corrects (or contrasts with) a previously mentioned constituent. While broad and narrow focus refer to the scope of focus, contrastive refers to the information structure of the discourse, and thus can overlap with both broad and narrow focus.

The focused constituents were then examined for any potential correlate of focus, whether morphosyntactic or phonetic. Since there is no prior information about focus marking in Bobo, the examination was not restricted to any specific phonetic or morphosyntactic properties. Instead, I attempted to identify any differences between the same (or similar) constituents when placed under different types of focus. The focus conditions were compared with each other and also to utterances in which the same (or similar) constituent was *given*, i.e. when it was previously mentioned and therefore was assumed to be part of the common ground (Büring, 2016). Each identifiable instance of focus was coded according to focus condition in the following manner: Broad and narrow focus are identified as the answers to *wh*-questions. This context is a prototypical condition for focus that is shared across a variety of theoretical approaches (Büring, 2016).

The 2014 data and the QUIS data are designed as question-and-answer prompts, and as such, *wh*-questions and responses are well-represented. Contrastive focus is identified as the correction of a previously mentioned item. The division of contrastive focus into a separate category of focus is controversial; see Kiss (1998) for an argument for treating it as a separate

category, and Zimmerman (2007) for an argument against. Regardless of whether contrastive focus is a distinct semantic or syntactic category, contrastive focus is frequently more strongly marked than non-contrastive focus. For example, in English, contrastively focused items have higher intensity, longer duration, and greater pitch range than non-contrastively focused items (Breen et al., 2010). This makes contrastive focus particularly useful for the investigation of prosodic focus marking.

As stated above, there were no prior hypotheses about whether or how focus is marked in Bobo. Any morphosyntactic or phonetic characteristics of the surface realization that are associated with focus condition are potentially examples of focus marking. Special attention was also paid to any characteristics that deviated from the expected form of the sentence, e.g. non-canonical word order or unexpected F0 excursions. However, no such characteristics were found. The results are presented in two sections. In §5.3, I discuss the lack of any consistent phenomena associated with focus condition. In section §5.3.4, I discuss some problematic cases that complicate the conclusion that there is no focus marking, but that ultimately appear to be conditioned by factors other than focus.

5.3 Results

Obligatory morphosyntactic focus-marking can be easily ruled out. Most sentences are produced with their canonical word order regardless of focus condition. There are no morphemes that are candidates for focus marking. Two examples from the QUIS data are provided below, in examples 25 and 26. In example 25, the response sentence is all new and thus in broad focus, and in example 26 *dẽmẽn=õ* 'the beans' is in narrow focus. Both sentences are responses to a *wh*-question.

(25) a¹ ŋʷõ⁵nõ⁵ n tɛ³?
3.SG what be
“What happened?”

bo¹⁵ ja¹ dẽ¹mẽ¹=õ¹ zõ¹.
DEM woman beans=DEF eat
“A woman ate the beans.”

(26) ja¹=je¹ ŋʷõ⁵nõ⁵ n zõ¹?
woman=DEF what eat
“What did the woman eat?”

a¹ dẽ¹mẽ¹=õ¹ zõ¹.
3.SG Woman beans=DEF
“She ate the beans.”

In both cases the response is a simple transitive sentence in a canonical SOV order. There are no focus-marking morphemes. The only grammatical morphemes present, in this case the demonstrative *bo¹⁵* and the definite enclitic *=õ¹*, are used regardless of focus condition. There is also no evidence of optional morphosyntactic focus marking. The QUIS data shown above is representative; the morphosyntactic form of the responses does not differ between focus conditions. Example 27 below demonstrates the same lack of morphosyntactic differences between cases of broad focus and contrastive narrow focus. In this example, which is also from the QUIS data, S4 makes a statement which S3 then corrects. The morphosyntactic form of the statements is identical except for the focused item itself (*saa¹* ‘three’).

(27) da¹la⁵¹ pə³la³ ti³ ki¹ta³bu³ ka¹la³=o¹.
girls Two be book read=LOC
“Two girls are reading a book.”

a¹⁵o³, da¹la⁵¹ saa¹ ti³ ki¹ta³bu³ ka¹la³=o¹.
no, girls three be book read=LOC
“No, three girls are reading a book.”

The QUIS is designed to allow speakers the freedom to construct their responses with minimal direction from the researcher or interference from the contact language. Unlike

elicitations in which speakers read a pre-designed sentence aloud, the QUIS allows speakers to employ morphosyntactic constructions that the researcher might not have anticipated. It is unlikely that Bobo has morphosyntactic focus marking that the speakers chose not to use at any point in the questionnaire. This is true of all of the data types, including the map tasks, in which the speech was most spontaneous.

It is more difficult to rule out prosodic focus marking. For this it is necessary to have the same word (or a word with the same phonological and grammatical properties) in different types of focus, so that a comparison can be made. Ideally this comparison includes a quantitative analysis, especially if the use of gradient phenomena such as pitch range expansion is expected. This is not possible at this time. The QUIS and the map task are not designed with quantitative analysis in mind; the freedom that they offer speakers comes at the cost of controlled comparisons. The 2014 elicitation data contains too few tokens and does not include many examples of the same target word in different focus conditions. Therefore the analysis that follows is qualitative. The examples that I include are representative of the data.

5.3.1 Contrastive focus

There is no evidence of contrastive focus marking. Contrastively focused constituents are produced like those in other focus conditions. Figure 62 and Figure 63 below are provided as examples. Figure 62 shows the response to the statement "the woman called Fatou"; *wi¹ri¹* 'call' is in contrastive focus as a correction to the verb. Figure 63 shows the response to the question "who called Fanta?". Here *wi¹ri¹* 'call' is given and thus not focused. There is no evidence of focus marking in these examples; the focused constituent, in this case the verb, is produced similarly in both focus conditions, with no indication of prosodic focus marking on either the verb itself (i.e., there is no noticeable change in pitch or duration) and no indication of a phrase boundary before

Figure 62

Declarative statement with wi[1]ri[1] *'call' in contrastive focus (S3, M)*

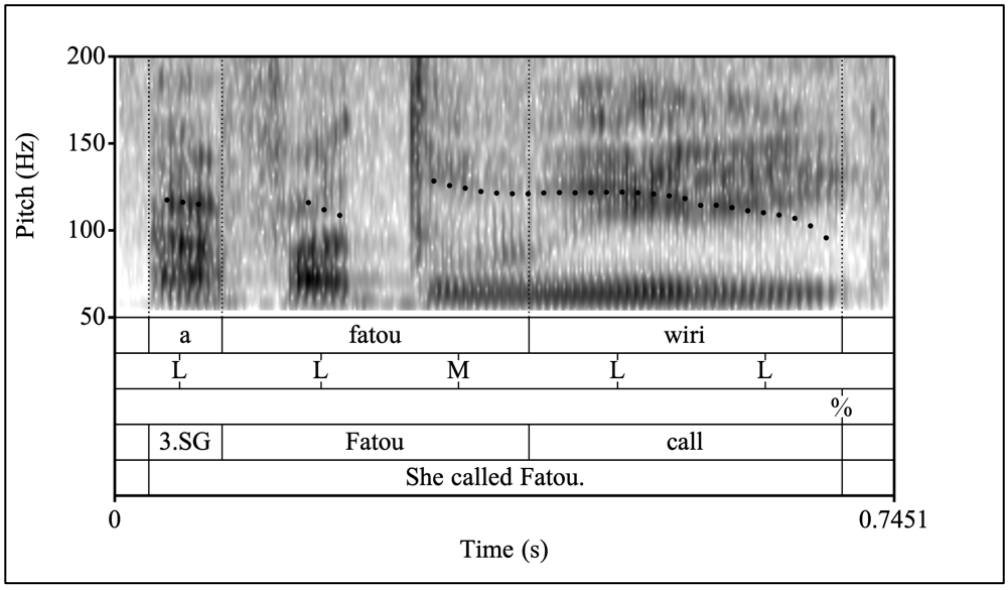
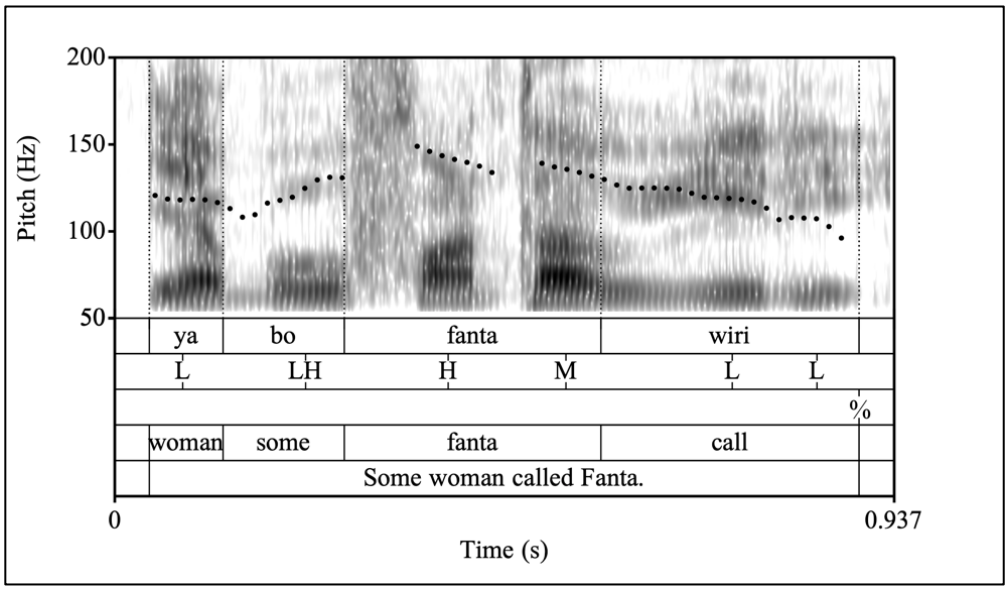


Figure 63

Declarative statement with wi[1]ri[1] *'call' not focused (S3, M)*

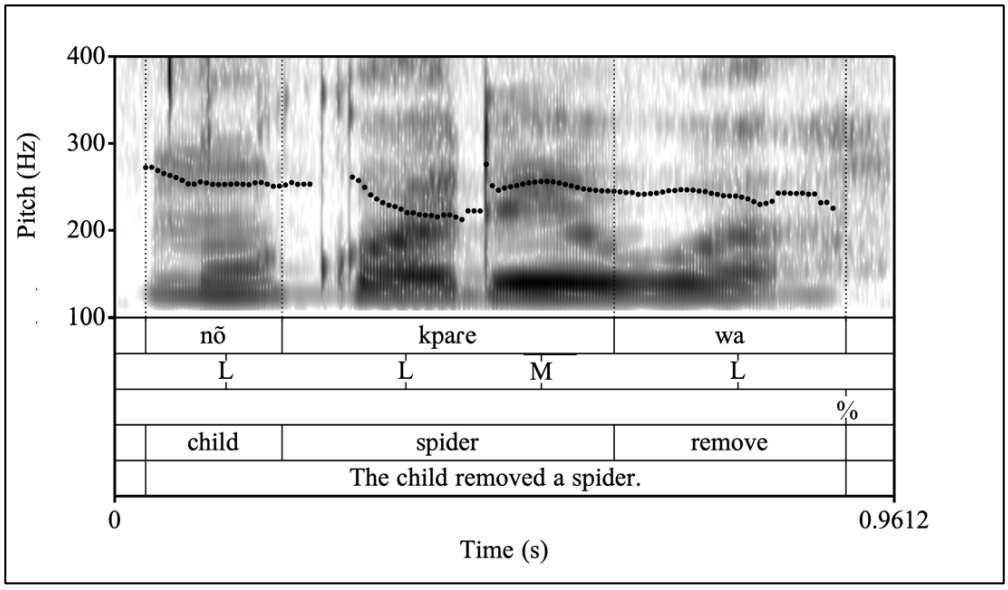


the verb. In this example, the contrastively focused item is utterance final, meaning that a boundary placed *after* the focused constituent would not be observable. Figure 64 below is an example of a

contrastively focused item that is not utterance final to illustrate that a boundary is not placed after the focused constituent. In this sentence, the object noun *kpa[1]ɾe[3]* 'spider' is in contrastive focus; it is a response to the question "did the child remove some grass?". There is no indication of a prosodic break between the focused object and the following verb.

Figure 64

Declarative statement with a contrastively focused object (S1, F)



Another example of the lack of contrastive focus marking is given in Figure 65 and Figure 66. The sentence shown in Figure 65 is a response to "did you remove the fonio?" where the object, *da[3]zu[3]go[3]* 'leaves', is in contrastive focus. The sentence shown in and Figure 66 was recorded as part of a separate elicitation task in which there was no preceding question; it is all new information and therefore *da[3]zu[3]go[3]* 'leaves' is in broad focus. The same item under different focus conditions is produced similarly in each sentence.

The prosodic characteristics of constituents in contrastive focus are predictable based on their word-level prosodic characteristics. There does not appear to be any post-lexical prosodic characteristic associated with contrastive focus. The examples above illustrate that there is no

Figure 65

Declarative statement with a contrastively focused object (S1, F)

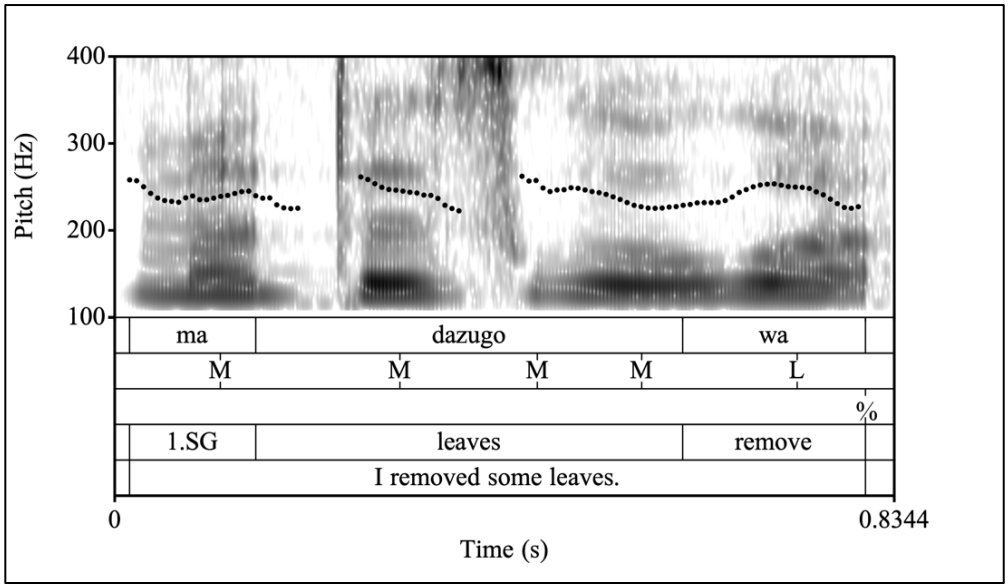
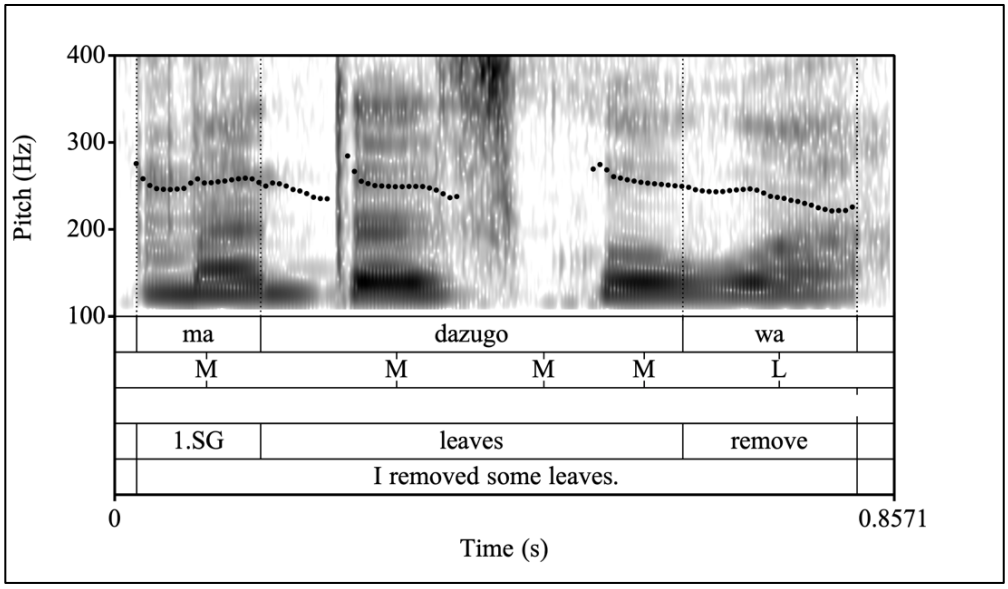


Figure 66

Declarative statement with a contrastively focused object (S1, F)



change in F0. Amplitude and duration are not visible, but they are also not audibly different. There appears to be no contrastive focus marking in this data, whether morphosyntactic or prosodic. In the future, the lack of prosodic focus marking could be confirmed with data collection designed to

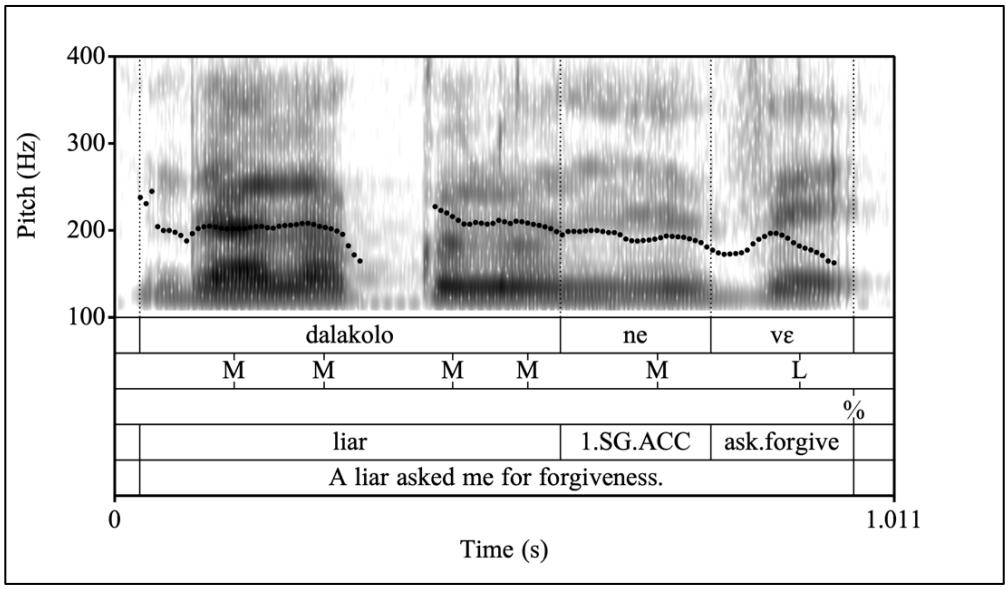
facilitate a quantitative analysis.

5.3.2 Narrow focus

There are also no correlates of narrow focus marking. This is expected, based on the data presented in the previous section; contrastive focus marking is generally the most salient. An example is provided below in Figure 67. In this sentence, the subject *da³la³ko³lo³* 'liar' is the response to the question "who asked you for forgiveness?".

Figure 67

Declarative statement with subject in narrow focus (S2, F)



There is no indication of focus marking on *da³la³ko³lo³.* The lexical M tones are realized with a level F0. The lack of focus marking can be seen even more clearly when compared to the sentence in Figure 68. This sentence was collected as part of an elicitation task in which there was no question prompt and is all-new. The M-toned subject of this sentence is also produced with a level F0. Another comparison is given in Figure 69 and Figure 70, with three sentences from the QUIS data. Figure 69 shows the response to the question *"What happened?"* with the name *Fa¹tou³* in broad focus. Figure 70 shows the response to the question "*Who ate the beans?"* with

the name *Fa¹tou³* in narrow focus. Figure 71 shows the response to the question *"Who hit Fatou where?*" with *Fa¹tou³* being given.

Figure 68

Declarative all-new statement with M toned subject (S2, F)

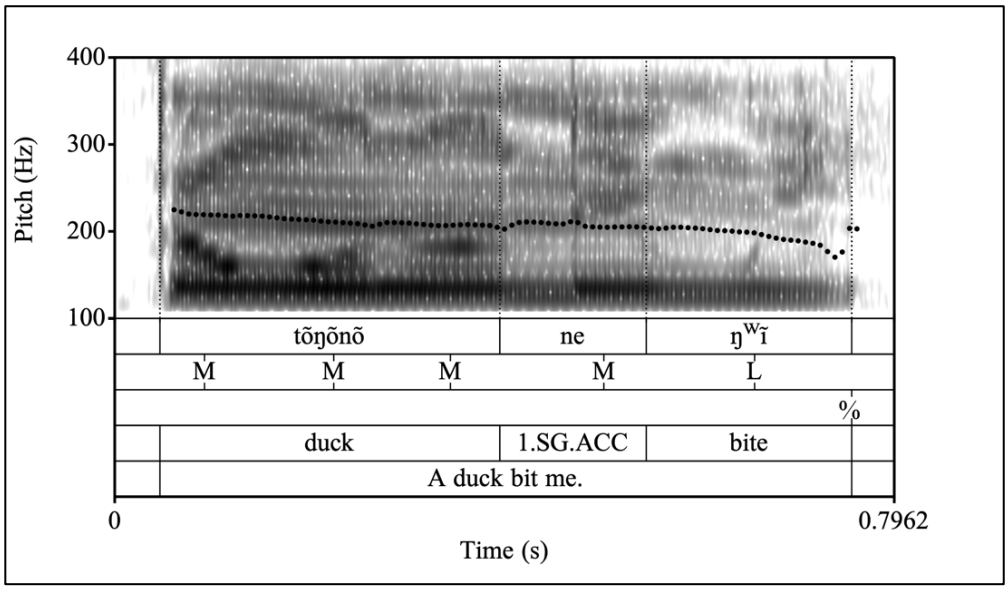


Figure 69

Declarative statement with Fa¹tou³ *in broad focus (S3, M)*

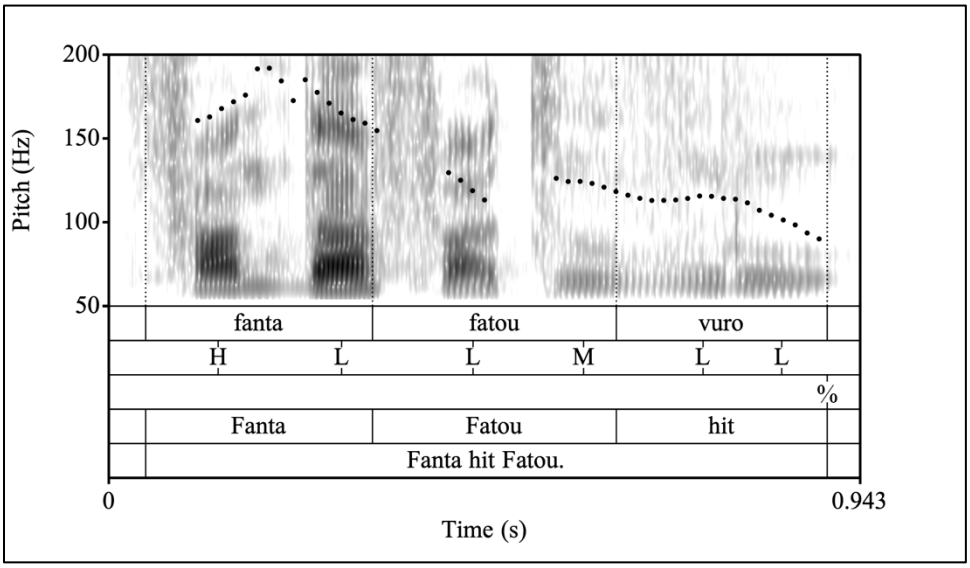


Figure 70

Declarative statement with Fa¹tou³ *in narrow focus (S3, M)*

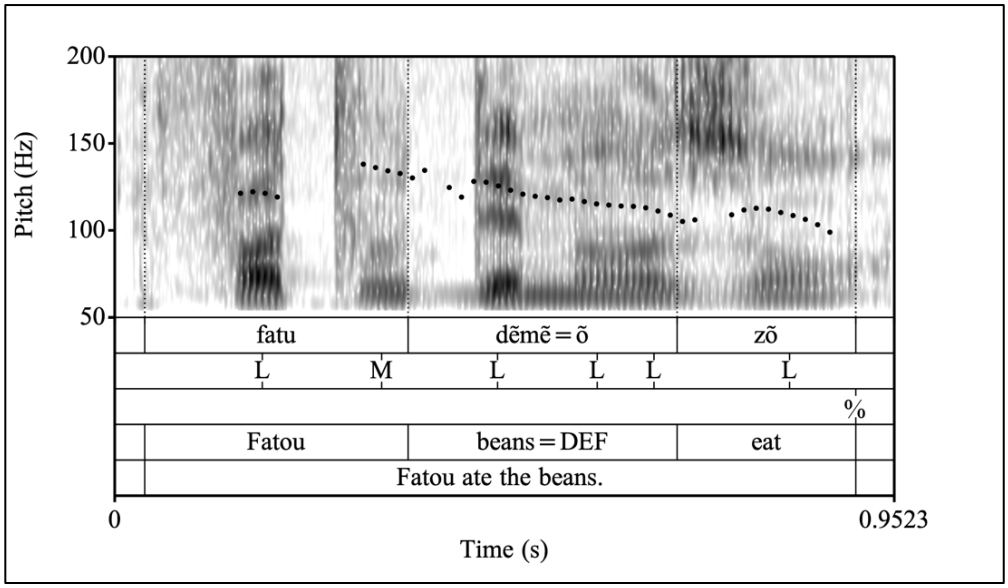
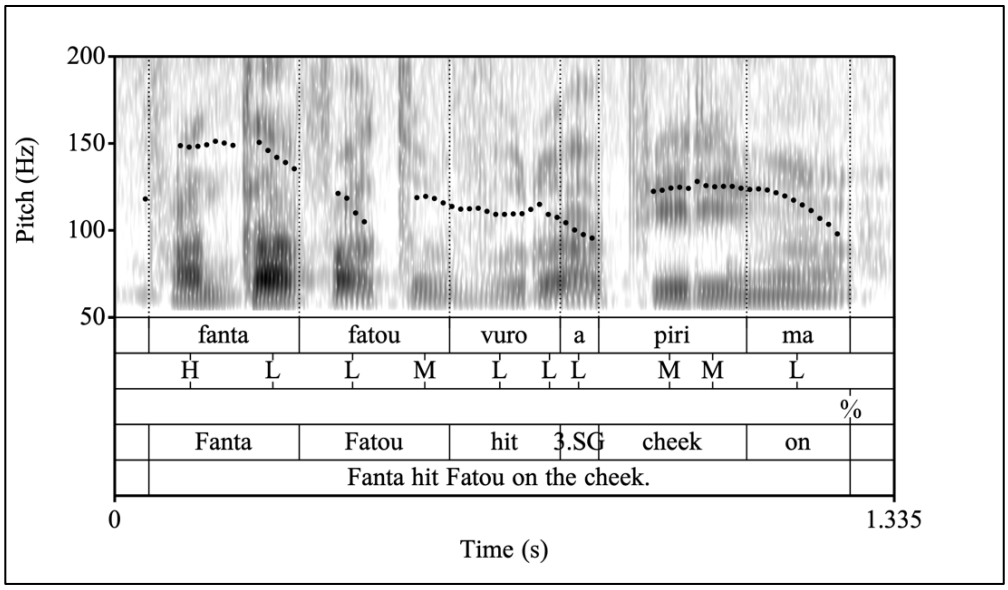


Figure 71

Declarative statement with Fa¹tou³ *not focused (S3, M)*



The realization of *Fa¹tou³* is similar in each case; there is no noticeable difference in F0, duration, or amplitude on the name itself and no prosodic break either before or after. Not only is there no difference between constituents under broad and narrow focus, there is also no difference

between constituents under broad focus and constituents that are given. Focus marking is not present.

5.3.3 Broad focus

In a language that has focus marking, it is important to determine how focus marking is realized when the focused constituent is larger than a single word. In a language without focus marking, this question is largely moot. It would be unusual for Bobo to have broad but not narrow focus marking. The examples of broad focus discussed above were all-new sentences in which the entire sentence is in focus. There were no indications of focus marking in these examples. Below are two examples of broad focus in which the focused constituent is smaller than the sentence, to illustrate that focus marking does not unexpectedly appear in these cases. Figure 72 shows the response to the question *"What did the woman eat?"*, with the response *fan^{5}ta^{1} ta^{1} dɛ̃^{1}mɛ̃^{1}=õ^{1}* 'Fanta's beans' in focus. Figure 73 shows an example of the VP *zio^{3} mɛ̃^{1}nɛ̃^{1}* 'drink water' in contrastive focus. In both examples, the focused constituent is realized as is expected without focus.

Figure 72

Declarative statement with fan^{5}ta^{1} ta^{1} dɛ̃^{1}mɛ̃^{1}=õ^{1} *in broad focus (S3, M)*

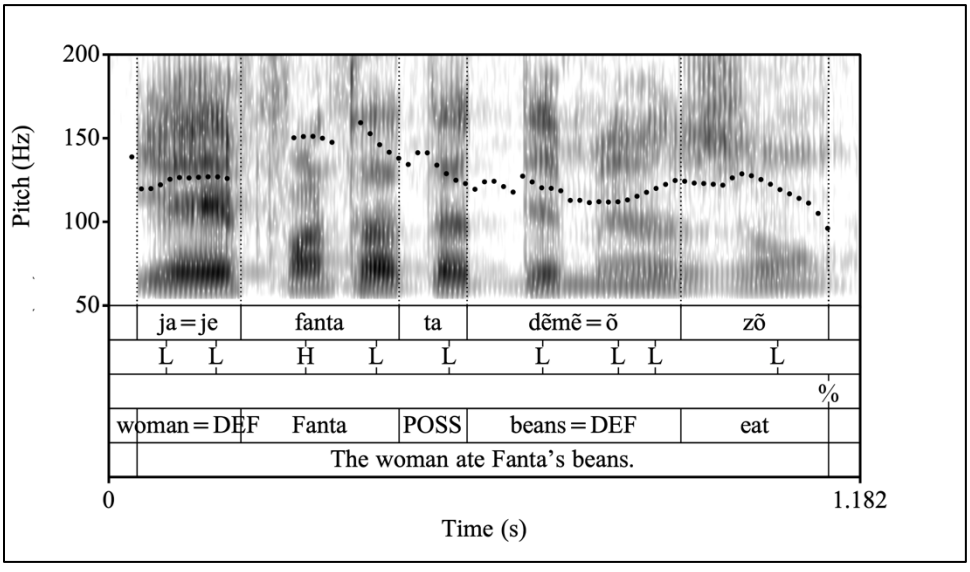
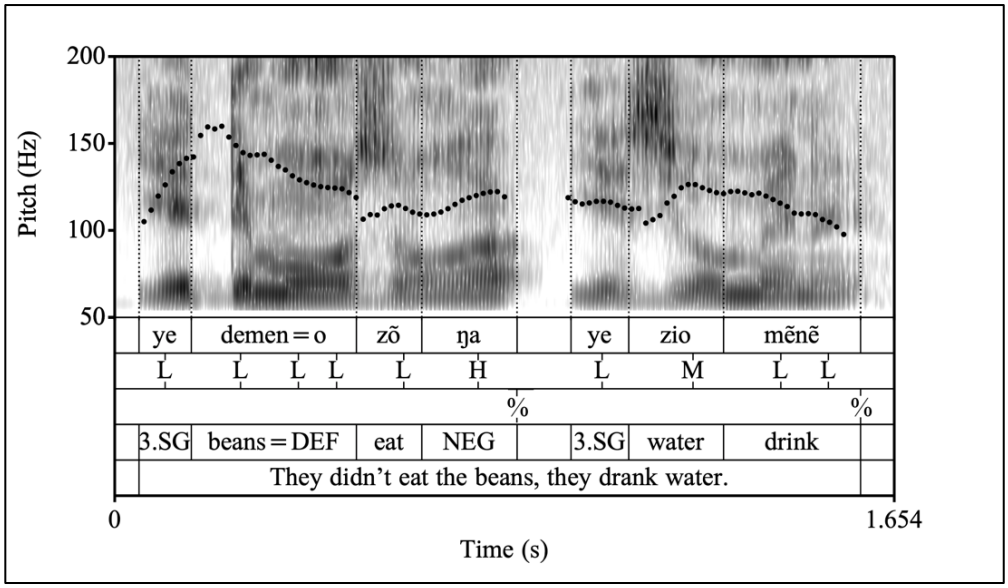


Figure 73

Declarative statement with zio^{3} mẽ^{1}nẽ^{1} *in broad contrastive focus (S3, M)*



5.3.4 Problematic cases

Some of the 2014 data presents a problem for the analysis presented above. In several of the responses, there is an unexplained fall in F0 before a focused verb, as is shown in Figure 74, which shows the response to the question "did you remove the fonio?" In this example, *zõ* 'eat' is contrastively focused. An unexplained fall in F0 can also occur after a focused object, as is shown in Figure 75. In this example, *fẽ*^{3} 'fonio' is contrastively focused in response to the question "did you remove the sorghum?"

A possible explanation for this pattern is that focus is being marked through prosodic phrasing in these examples. If the focused item is placed within its own prosodic phrase, we would expect there to be a prosodic break in this position when there is either a focused object or focused verb. However, there are reasons not to accept this explanation. First, the fall in F0 does not only occur when object or verb is focused, as is shown in Figure 76 and Figure 77. In Figure 76, the fall in pitch appears in a question, when neither the object nor verb is expected to be focused. In,

Figure 74

Declarative statement with zõ[1] *in contrastive focus (S2, F)*

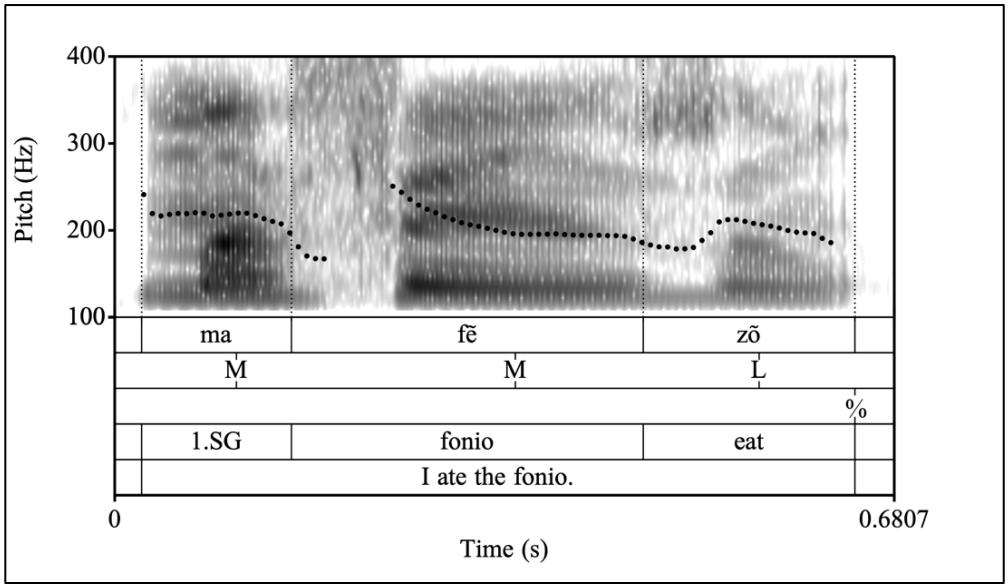


Figure 75

Declarative statement with fẽ[3] in contrastive focus (S2, F)

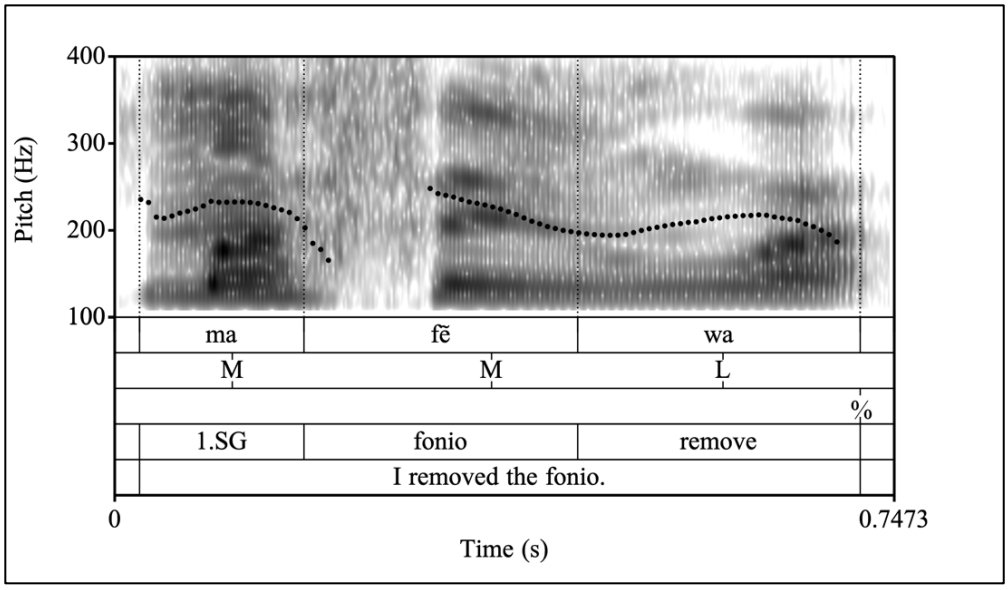


Figure 77, the fall occurs in the same position between the object and the verb, but it is the adverb *sa[1]le[3]* 'quickly' that is in focus. We have also already seen an example where the fall in pitch on the object occurs when there is *subject* focus, in Figure 67.

Figure 76

Polar question with unexpected fall in F0 on the object (S1, F)

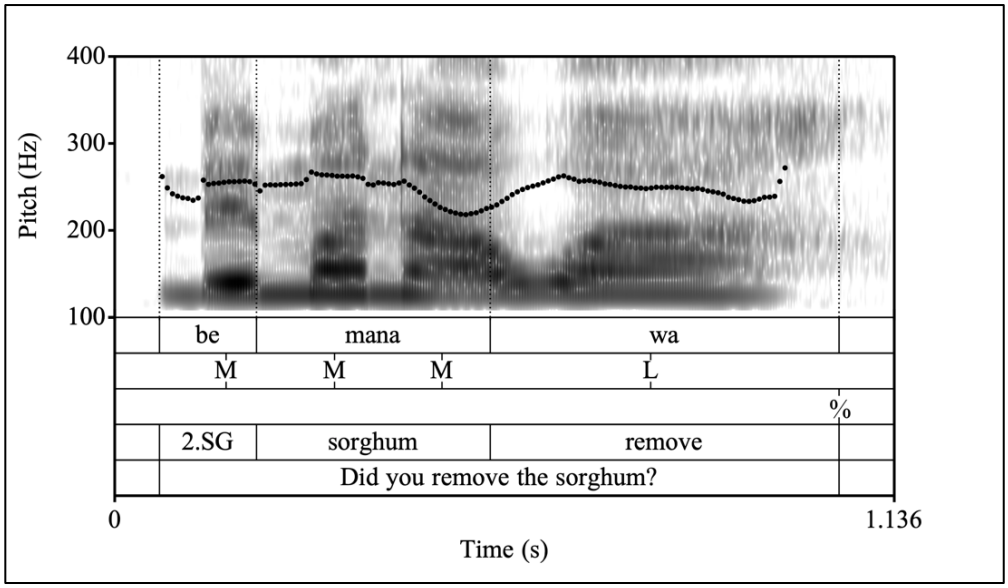
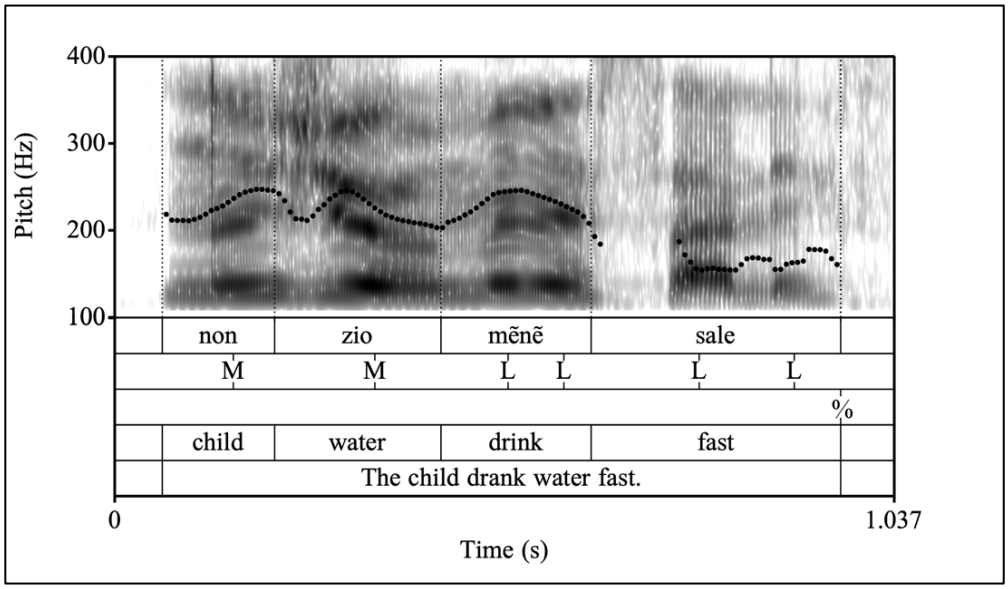


Figure 77

Declarative statement with Unexpected fall in F0 on the object when the adverb is focused (S2, F)



Importantly, this phenomenon occurs specifically in this task; it does not occur in the QUIS data or in the map tasks. This suggests that there is something about the task that is prompting it.

There is no firm answer here. One possible explanation is that the observed falls in F0 are due to the task involving read speech. In 2014, S1 and S2 had less practice reading Bobo prompts. Although the speech sounds natural, it is slow and deliberate; this could have prompted more prosodic breaks between words. The fact that this pattern does not appear in the QUIS data or in the map tasks, both of which are more spontaneous, supports this interpretation. Regardless of the explanation, the fall in pitch is not correlated to focus condition and is therefore unlikely to be focus marking.

5.4 Discussion

The purpose of this chapter was to investigate focus marking in Bobo, which was not found. The data examined here does not show evidence of focus marking in response to *wh*-questions or in responses containing a correction. Not only does it not exhibit *obligatory* focus marking in these contexts, there is also no evidence of *optional* focus marking as in Hausa (Zimmerman, 2007), or *restricted* focus marking, as in Buli (Schwarz, 2009). Rather, it shows no evidence of focus marking at all, similar to Northern Sotho (Zerbian, 2007) and Embosi (Rialland & Aborobongui, 2016). The question is how to interpret these results. It is unlikely that the lack of focus marking in this data is the result of a problem with the methodology, since it is consistent across all of the speakers and all of the tasks. Bobo can be added to the list of African languages that do not mark focus in these prototypical contexts.

Previous examples of languages that have been argued to not have focus marking in these contexts fall into two types. The first type are those languages in which *wh*-questions and contrastive focus conditions do elicit explicit markers of information structure, but the semantic or pragmatic contexts in which these markers are used do not correspond well to most definitions of "focus." There is considerable discussion of the underlying information structural categories that

these markers encode—for example, whether it is "focus" of a different type (e.g. Bearth, 1992 on Tura) or whether it is instead a category such as topic (e.g. Downing & Hyman, 2016). The second type are those languages in which there is no explicit marker of information structure in these contexts, such as Northern Sotho (Zerbian 2007) and Embosi (Downing & Aborobongui, 2016).[33] Languages like Tura and Embosi point to cross-linguistic differences in how information structure is grammaticalized, both in the degree of grammaticalization and in the grammatical categories that exist. This has led some to question whether "focus" itself is a coherent concept that can be applied cross-linguistically (e.g. Matić & Wedgewood, 2013).

It is not possible to say what Bobo can contribute to this discussion at this time, since only two potential focus contexts were examined here. However, negative result is intriguing and suggest that information structure in Bobo might also function differently than in Indo-European languages. Further work on information structure in Bobo is clearly needed.

[33] In Embosi, prosodic prominence exists in other contexts, but, crucially, the authors interpret it as "emphasis" and not focus. Emphasis is paralinguistic and gradient, determined by the amount of "emphasis" the speaker wishes to place on an item rather than by the grammatical structure of the language. Emphasis and focus are orthogonal concepts: An emphasized item is not always focused and a focused item is not always emphasized. The strength of the prosodic prominence on a focused item is determined by both the grammatical requirements of the language and the amount of emphasis. For more discussion of this division, see Hartmann (2008) and Downing & Pompino-Marschall (2013). Emphasis was not examined here because it is a difficult concept to operationalize for a language whose prosody is not yet well-studied.

Chapter 6

Conclusion

In this dissertation I investigate several aspects of the prosodic system of Southern Bobo Madaré. The research is based on data recorded in Bobo-Dioulasso between 2014 and 2017. Five speakers participated in various elicitation and spontaneous speech tasks. The investigation has two principle goals. The first goal is descriptive: to identify and describe the surface-level prosodic phenomena that exist in the language and to propose an analysis of their phonological structure. I adopt the framework of Autosegmental-Metrical Theory (Pierrehumbert, 1980) for the analysis of its phrase-level prosody. The adoption of AM theory serves the second goal, which is to contribute to the growing body of literature on prosodic typology—much of which shares this widely-understood framework. Since this project began, there have been significant additions to the literature on the intonational systems of African languages (notably Downing & Rialland, 2016), but Bobo is still the first Mande language to have a prosodic description with this level of detail and scope.

A fundamental assumption of AM theory is that prosodic structure consists of categorical phonological elements, such as intonational tones, that are used to encode linguistic structure and meaning. Patterns in phrase-level prosodic phenomena are therefore amenable to the same type of phonological analyses as word-level prosodic phenomena. A recurrent challenge in this project is the amount of variation in the data that cannot be attributed to differences in meaning. Even in an elicitation context, in which sentence-level meaning is as tightly controlled as possible, there is

not always a clear categorical relationship between the prosodic phenomena in the sentences speakers produce and the intended sentence-level meaning. An example is the breathy utterance termination that is found in approximately two-thirds of the elicited polar questions. It occurs frequently enough that an *increased likelihood* of breathy utterance termination can be said to be a characteristic of polar questions, but not frequently enough to say that breathy utterance termination marks polar questions. There is clearly a different (weaker) relationship between breathy utterance termination and its meaning than, for example, lexical tone and word meaning. One point worth noting is that these findings are difficult to account for within the AM model, which only incorporates tonal characteristics of prosody. Further research is needed to integrate the variation in voice quality into this theoretical model.

6.1 Main Findings

6.1.1 *Word-level prosody*

Previous work on word-level prosody of Southern Bobo Madaré has described its lexical tone system and little else (Morse, 1976; Le Bris & Prost, 1981; Sanou, 1993). Even then, there are significant disagreements about the inventory of contour tones. My analysis is that there are four: LM, ML, LH, and HL. However, the HL tone frequently surfaces as H in phrase-medial position. The reason for this is unknown; more data is needed regarding the behavior of the HL tone in different tone contexts. Similar alternations do not exist for the other contour tones.

In addition to lexical tone, I investigated the evidence for lexical stress—i.e. the existence of strong and weak syllables within a prosodic domain. The phonetic data that I collected is ambiguous, showing that lexical stress is not expressed straightforwardly as greater amplitude or duration. The distribution of reduced vowels and /ə/ on the other hand is best explained with reference to a prosodic foot. Specifically, their distribution is explained by a preferentially binary

iambic foot; reduced vowels only occur in the first syllable of this foot. The existence of a binary foot can also account for the sequences of tones that can (and cannot) be found within a word. Like many Mande languages, Bobo has a restricted inventory of lexical tone melodies. In Bobo, these melodies are associated with a bisyllabic domain. The tone sequences in words containing more than one prosodic foot can be decomposed into a sequence of lexical tone melodies, each associated with one foot. These facts suggest that the domain of tone melodies is in fact the foot, similar to the proposal for Bamana by Leben (2001; 2003).

6.1.2 Prosodic phrase structure

Previous work on the prosodic phrase structure of Bobo is limited to Morse (1976), which does not include much phonetic detail or explanation of the analysis. My own data collection and analysis therefore does not make prior assumptions about the prosodic phrase structure of Bobo. The purpose of this section of the dissertation is to investigate the prosodic properties of intonational phrases, how those properties are used to mark sentence type, and the possibility of other prosodic phrase types.

Overall, Bobo makes very little use of intonation, which is not surprising given its complex tone inventory. An intonational phrase does not have any obligatory intonational tones: There is no evidence of phrase or pitch accent and a boundary tone only occurs in marked contexts. Prosodic phrase structure is marked primarily through non-intonational properties. The final phrase boundary of a prosodic phrase is marked by final lengthening and in some cases non-modal vowel phonation. Downstep is also a property of the IP. It is clear that either downstep or downdrift affects H tones, but due to the relative infrequency of H tones, the conditions under which this can be observed do not occur in the majority of phrases. Further examination is needed in order to determine the phonological contexts that trigger downtrends and the domain over which they

extend. It is unclear whether downstep is an obligatory property of the IP. Not all IPs contain H tones, meaning that downstep does not occur in these phrases. More data is needed to determine whether it does indeed occur in every instance of an H tone following an L tone. Downstep does occur in the instances of H tone following an L tone examined here, but comprehensive data that could address whether downstep always occurs in different grammatical or pragmatic context was not collected.

Individual types of sentences differ in their prosodic characteristics. Polar questions and *wh*-questions have prosodic characteristics that distinguish them from declarative statements. However, none of these characteristics are present in all questions of a particular type. In elicitation contexts, polar questions are associated with an L% boundary tone, breathy utterance termination, and prolongation of the final vowel These are three of the four characteristics of the "lax question prosody" described by Rialland (2009). As mentioned above, breathy termination is not found in all of the questions. In spontaneous speech the prosody of polar questions is much more variable than in elicitation. In spontaneous speech polar questions are often marked morphosyntactically using the particle *kɛ*^{1} or *=ga*^{5}*/ŋa*^{5}—and in these cases there is no special question prosody. When polar questions are not marked with a particle, polar questions do not always possess all three of the characteristics of question prosody found in the elicitation sentences; any of the three characteristics can be absent. *Wh*-questions are similarly variable. In elicitation contexts, *wh*-questions are associated with breathy utterance termination and less extreme vowel prolongation than polar questions. They have no boundary tone. In spontaneous speech, the prolongation is often not present. It not possible to determine whether the prolongation is an extra degree of final lengthening than occurs in statements or whether it is caused by a different underlying process (see discussion in §4.7).

Some of the variability can be attributed to a difference between S1 and the four other speakers. Rather than using L% to mark polar questions, S1 raises the F0 of the final word. This presents a complication for a phonological interpretation of intonation in terms of discrete intonational tones. If the raising is due to an H% intonational tone, then different intonational tones must have different rules governing their realization: L% is concatenated to the end of the tonal string, while H% is superimposed, resulting in the final tone being raised. This problem would be compounded if H% tone is sometimes also sometimes concatenated. There is some evidence that H% is an optional boundary tone that is indeed realized in this way. It appears in some *wh*-questions and in some polar questions in the spontaneous speech data, where it seems to indicate surprise or disbelief. In these cases, it surfaces as a rise—that is, it is not superimposed on the lexical tone target.

Negated statements do not have special prosody, contrary to Morse (1976).

6.1.3 Focus

The most unusual finding in this dissertation is that Bobo does not mark focus—at least not in prototypical contexts. I investigated focus marking in three different types of task: in pre-designed question-and-answer elicitation sentence pairs, in a questionnaire in which the speakers were free to use any morphosyntactic focus marking they felt appropriate, and in a map task. In none of these tasks were there morphosyntactic or prosodic differences between sentences according to the presence or absence of semantic focus. This is an unexpected but not unprecedented result; focus marking in these contexts has also not been found in Northern Sotho (Zerbian, 2007) and Embosi (Rialland & Aborobongui, 2016). A lack of focus marking is likely to be more widespread than is currently known.

6.2 Future Directions

The methods used in this dissertation were broad-ranging and primarily qualitative, due to the lack of prior knowledge about the prosodic system of Bobo. The understanding of the prosodic system is now at the stage where specific phenomena can be investigated using more targeted methods. Focus marking is of particular interest. Obligatory morphosyntactic focus marking has been ruled out and optional morphosyntactic focus marking was not found. There is also no evidence of categorical phonological markers of focus such as pitch accent—and there appear to be no phonetic correlates of focus at all. A quantitative investigation using controlled data could confirm this. A perception experiment investigating whether speakers can identify the focus condition under which a sentence was produced would also be valuable, as this would confirm that there are no other cues that are not being captured by the methods being used.

Another question that can now be investigated in detail is the relationship between different prosodic phenomena and the meanings they encode. Although I have referred to the existence of "question prosody" in Bobo, this prosody has turned out to be quite variable, and also not exclusively (or obligatorily) used to marked questions. It is possible that some of these features have a broader meaning or a meaning that varies in context. It is also possible that there are additional prosodic phenomena that did not appear in the elicitation data. More directed collaborative speech tasks in addition to the map task that was used here could be illuminating. These tasks also have the benefit of (generally) not requiring the speakers to read a prompt, which limited the number of speakers that could participate in the elicitation tasks.

Finally, the question of the nature of downtrends in Bobo is still not settled. What is known is that an H tone following an L tone will generally have a lowered F0; thus, it is either a process of downdrift or downstep. What is not known is the domain of this process and the nature of the

change to F0 range. At the time the data for this project was collected, too little was known about the lexical tone system to design elicitation sentences that can distinguish between the different possibilities. It is now possible to design these sentences.

References

Abbott, C. (2006). *Oneida teaching grammar*. https://www.uwgb.edu/UWGBCMS/media/Oneida-Language/files/teaching-grammar-revised4.pdf

Aboh, E. O., Hartmann, K., & Zimmerman, M. (Eds.). (2007). *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*. De Gruyter Mouton.

Armstrong, M. E. (2010). Puerto Rican Spanish intonation. In P. Prieto & P. Roseano (Eds.), *Transcription of intonation in the Spanish language* (pp. 155–189). Lincom EUROPA.

Arvaniti, A., & Ladd, D. R. (2009). Greek wh-questions and the phonology of intonation. *Phonology*, *26*(1), 43–74.

Arvaniti, A. (2011). The representation of intonation. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 1–24). Wiley Blackwell.

Aikhenvald, A. Y. (2003). *A grammar of Tariana, from Northwest Amazonia*. Cambridge University Press.

Bearth, T. (1992). Constituent structure, natural focus hierarchy, and focus types in Toura. *Folia Linguistica*, *26*(1–2), 75–94.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 9–54). Oxford University Press.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.

Bennett, R. T. (2012). *Foot-conditioned phonotactics and prosodic constituency* [Doctoral dissertation, UC Santa Cruz].

Bishop, J. (2002). *Aspects of intonation and prosody in Bininj gun-wok: Autosegmental-metrical analysis* [Doctoral dissertation, University of Melbourne].

Boone, C. J. L. (2016). *Tone in the Bobo Madare North Noun System* [Master's thesis, Leiden University].

Brunelle, M., Hạ, K. P., & Grice, M. (2012). Intonation in Northern Vietnamese. *The Linguistic Review*, *29*, 3–36.

Büring, D. (2006). Focus projection and default prominence. In V. Molnár & S. Winkler (Eds.), *The Architecture of Focus* (pp. 321–346). De Gruyter Mouton.

Büring, D. (2010). Toward a typology of focus realization. In M. Zimmerman & C. Féry (Eds.), Information Structure: Theoretical, typological, and experimental perspectives (pp. 177–205). Oxford University Press.

Büring, D. (2016). *Intonation and Meaning*. Oxford University Press.

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*(7), 1044–1098.

Cahill, M. (2016). Kɔnni intonation. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 53–88). De Gruyter Mouton.

Caron, B., Lux, C., Manfredi, S., & Pereira, C. (2015). The intonation of topic and focus: Zaar (Nigeria),Tamasheq (Niger), Juba Arabic (South Sudan) and Tripoli Arabic (Libya). In A. Mettouchi, M. Vanhove, & D. Caubet (Eds.), *Corpus-based studies of lesser-described languages: The CorpAfroAs corpus of spoken AfroAsiatic languages* (Vol. 68, pp. 68–115). John Benjamins.

Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd ed.). Blackwell Publishing.

Connell, B. (2001). *Downdrift, downstep, and declination.* Paper presented at the Typology of African Prosodic Systems Workshop, Bielefeld University, Germany. https://www.researchgate.net/publication/228747173_Downdrift_downstep_and_declination

Connell, B. (2011). Downstep. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. D. Rice (Eds.), *The Blackwell Companion to Phonology: Vol. II. Suprasegmental and Prosodic Phonology*. Blackwell Publishing.

Connell, B. (2016). Tone and intonation in Mambila. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 131–166). De Gruyter Mouton.

Connell, B., & Ladd, D. R. (1990). Aspects of pitch realization in Yoruba. *Phonology*, *7*(1), 1–29.

Coulter, G. R. (1993). Phrase-level prosody in ASL: Final lengthening and phrasal contours. In G. R. Coulter (Ed.), *Current issues in ASL phonology* (Vol. 3, pp. 263–272). Elsevier.

Creissels, D. (2015). *A Sketch of Mandinka*. http://www.deniscreissels.fr/public/Creissels-sketch_of_Mandinka.pdf

Cruttenden, A. (1997). *Intonation* (2nd Edition). Cambridge University Press.

Diagana, Y. (1987). La focalisation en soninké. *Mandenkan*, *13*, 61–72.

Dimmendaal, G. J. (2011). *Historical linguistics and the comparative study of African languages*. John Benjamins.

Downing, L. J., & Hyman, L. (2016). Information structure in Bantu. In C. Féry & S. Ishihara (Eds.), *The Oxford Handbook of Information Structure* (pp. 790–813). Oxford University Press.

Downing, L. J., & Pompino-Marschall, B. (2013). The focus prosody of Chichewa and the Stress-Focus constraint: A response to Samek-Lodovici (2005). *Natural Language & Linguistic Theory*, *31*(3), 647–681.

Downing, L. J., & Rialland, A. (Eds.). (2016a). *Intonation in African tone languages*. De Gruyter Mouton.

Downing, L. J., & Rialland, A. (2016b). Introduction. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 1–16). De Gruyter Mouton.

Dumestre, G. (2003). Grammaire fondamentale du bambara. Éditions KARTHALA.

Dwyer, D. J. (1998). The place of Mande in Niger-Congo. In I. Maddieson & T. J. Hinnebusch (Eds.), *Language history and linguistic description in Africa: Selected papers of the 26th annual conference on African linguistics* (pp. 27–36). Africa World Press.

Dwyer, D. (2006). Mande languages. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (Second Edition, pp. 477–482). Elsevier. 10.1016/B0-08-044854-2/02302-6

Epstein, M. A. (2002). *Voice quality and prosody in English* [Doctoral dissertation, UCLA].

Féry, C. (2017). *Intonation and prosodic structure*. Cambridge University Press.

Ferreira, F. (2007). Prosody and performance in language production. *Language and Cognitive Processes*, *22*(8), 1151–1177.

Fiedler, I., Hartmann, K., Reineke, B., Schwarz, A., & Zimmerman, M. (2010). Subject focus in West African languages. In M. Zimmerman & C. Féry (Eds.), *Information structure: theoretical, typological, and experimental perspectives* (pp. 234–257). Oxford University Press.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Eds.), *The handbook of the phonetic sciences* (pp. 521–602). Wiley Blackwell.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at the edges of prosodic domains. *The Journal of the Acoustical Society of America*, *101*(6), 3728–3740.

Fujisaki, H., & Hirose, K. (1984). Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *The Journal of the Acoustical Society of Japan*, *5*(4), 233–242.

Genzel, S. (2018). The non-lax question prosody of Akan*, Sixth International Symposium on Tonal Aspects of Languages*, 2–6.

Glottolog. (n.d.) *Mande* [Language map]. Glottolog.org. Retrieved Jun 23, 2020, from https://glottolog.org/resource/languoid/id/mand1469

Goldsmith, J. A. (1976). *Autosegmental phonology* [Doctoral dissertation, Massachusetts Institute of Technology].

Gordon, M. (2008). The intonational realization of contrastive focus in Chickasaw. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: A cross-linguistic perspective* (pp. 65–78). Springer.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *2001*(29), 383–406.

Gósy, M., & Krepsz, V. (2018). Phrase-final lengthening of phonemically short and long vowels in Hungarian spontaneous speech across ages. In Maria Gósy & T. E. Gráczi (Eds.), *Challenges in analysis and processing of spontaneous speech* (pp. 99–126). Research Institute for Linguistics of Hungarian Academy of Sciences.

Green, C. R. (2010). *Prosodic phonology in Bamana (Bambara): Syllable complexity, metrical structure, and tone* [Doctoral dissertation, Indiana University].

Green, C. R., Davis, S., Diakite, B., & Baertsch, K. (2012). Domain-restricted reduction: A proposal for segmental feet in Bamana. In B. Connell & N. Rolle (Eds.), *Selected proceedings of the 41st annual conference on African linguistics* (pp. 1–9). Cascadilla Proceedings Project.

Güldemann, T., Zerbian, S., & Zimmerman, M. (2015). Variation in information structure with special reference to Africa. *The Annual Review of Linguistics*, *1*, 155–178.

Gundel, J. K. (1999). On different kinds of focus. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 293–305). Cambridge University Press.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.

Hartmann, K. (2008). Focus and emphasis in tone and intonation languages. In A. Steube (Ed.), *The discourse potential of underspecified structures* (pp. 389–412). De Gruyter.

Hartmann, K., & Zimmerman, M. (2007). In place—out of place: Focus in Hausa. In K. Schwabe & S. Winkler (Eds.), *On information structure, meaning, and form: Generalizations across languages* (pp. 365–403). John Benjamins.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Henton, C. G., & Bladon, A. (1988). Creak as a sociophonetic marker. In H. Hyman Larry M. & C. N. Li, (Eds.), *Language, speech and mind: Studies in honor of Victoria A. Fromkin* (pp. 3–29). Routledge.

Himmelmann, N. P., & Ladd, D. R. (2008). Prosodic description: An introduction for fieldworkers. *Language Documentation & Conservation*, *2*(2), 244–274.

Hockey, B. A., & Fagyal, Z. (1998). Pre-boundary lengthening: Universal or language-specific? The case of Hungarian. *University of Pennsylvania Working Papers in Linguistics*, *5*(1), 71–82.

Holton, G. (2005). Pitch, tone and intonation in Tanacross. In S. Hargus & K. Rice (Eds.), *Athapaskan prosody* (pp. 249–276). John Benjamins.

Hyman, L. M. (1993). Register tones and tonal geometry. In H. van der Hulst & K. Snider (Eds.), *The phonology of tone: The representation of register* (pp. 75–108). De Gruyter Mouton.

Hyman, L. M. (1999). The interaction between focus and tone in Bantu. In G. Rebuschi & L. Tuller (Eds.), *The Grammar of focus* (pp. 151–178). John Benjamins.

Hyman, L. M., & Monaka, K. C. (2008). Tonal and non-tonal intonation in Shekgalagari. *UC Berkeley Phonology Lab Annual Report*, *4*(4), 269–288.

Hyman, L. M. (2013). How autoesegmental is phonology? *UC Berkeley PhonLab Annual Report*, *9*(9), 32–59.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. MIT Press.

Jannedy, S. (2007). Prosodic focus in Vietnamese. *Interdisciplinary Studies on Information Structure*, *8*, 209–230.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* [Doctoral dissertation, The Ohio State University].

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*(2), 189–226.

Jun, S.-A. (2000). *K-ToBI (Korean ToBI) labelling conventions*. https://linguistics.ucla.edu/people/jun/ktobi/k-tobi.html

Jun, S.-A. (Ed.). (2005a). *Prosodic typology: The phonology of intonation and phrasing*. Oxford University Press.

Jun, S.-A. (2005b). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430–458). Oxford University Press.

Jun, S.-A. (Ed.). (2014a). *Prosodic typology II: The phonology of intonation and phrasing*. Oxford University Press.

Jun, S., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modeling and technology* (pp. 209–242). Kluwer Academic Publishers.

Karlsson, A., House, D., & Svantesson, J.-O. (2012). Intonation adapts to lexical tone: The case of Kammu. *Phonetica*, *69*, 28–47.

Kastenholz, R. (1997). *Sprachgeschichte im West-Mande. Methoden und Rekonstruktionen*. Rüdiger Köppe Verlag.

Keating, P. A. (2006). Phonetic encoding of prosodic structure. In J. Harrington & M. Tabain (Eds.), *Models, phonetic processes, and techniques* (pp. 167–186). Psychology Press.

Keating, P. A., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences.*

Kiss, K. É. (1998). Identificational focus versus information focus. *Language*, *74*(2).

Kohler, K. J. (2000). Linguistic and paralinguistic functions of non-modal voice in connected speech. *Proceedings of the 5th Seminar on Speech Production: Models and Data*, 89–82.

Konoshenko, M., & Sigova, S. (2018). *Interacting tone and intonation: A study of yes-no questions in Guinean Kpelle*.

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*, 162–179.

Krivokapić, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, *40*, 430–442.

Kula, N. C., Hamann, S., Downing, L. J., & Rialland, A. (2016). Intonation in Bemba. In *Intonation in African tone languages* (pp. 321–363). De Gruyter Mouton.

Kügler, F. (2016). Tone and intonation in Akan. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 89–129). De Gruyter Mouton.

Ladd, D. R. (1984). Declination: A review and some hypotheses. *Phonology Yearbook*, *1*, 53–74.

Ladd, D. R. (2001). Intonation. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: An international handbook* (pp. 1380–1390). De Gruyter Mouton.

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press.

Ladefoged, P. (2003). *Phonetic Data Analysis*. Blackwell Publishing.

Le Bris, P., & Prost, A. (1981). *Dictionnaire Bobo-Français: Précedé d'une introduction grammaticale et suivi d'un lexique français-bobo*. SELAF.

Leben, W. R. (2001). *Tonal feet*. Paper presented at the Typology of African Prosodic Systems Workshop, Bielefeld University, Germany. https://www.researchgate.net/publication/229032495_Tonal_feet

Leben, W. R. (2003). Tonal feet as tonal domains. In J. M. Mugane (Ed.), *Linguistic typology and representation of African languages* (pp. 129–138). Africa World Press.

Liberman, M. (1975). *The intonational system of English* [Doctoral dissertation, Massachusetts Institute of Technology].

Lickley, R. J., Schepman, A., & Ladd, D. Robert. (2005). Alignment of "phrase accent" lows in Dutch falling rising questions: Theoretical and methodological implications. *Language and Speech*, *48*(2), 157–193.

Lindau, M. (1986). Testing a model of intonation in a tone language. *The Journal of the Acoustical Society of America*, *80*(3), 757–764.

Ma, J. K.-Y., Ciocca, V., & Whitehill, T. L. (2006). Effect of intonation on Cantonese lexical tones. *The Journal of the Acoustical Society of America*, *120*(6), 3978–3987.

Maddieson, I. (1996). Phonetic universals. *UCLA Working Papers in Linguistics*, *92,* 160–178.

Makasso, E.-M., Hamlaoui, F., & Lee, S. J. (2016). Aspects of the intonational phonology of Bàsàá. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 167–193). De Gruyter Mouton.

Matić, D., & Wedgwood, D. (2013). The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics*, *49*, 127–163.

McCarthy, J. J., & Prince, A. R. (1993). Generalized alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993* (pp. 79–153). Springer.

McPherson, L. (2017). Tone features revisited: Evidence from Seenku (Mande, Burkina Faso). In D. L. Payne, S. Pacchiarotti, & M. Bosire (Eds.), *Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics* (pp. 5–22). Language Science Press.

McPherson, L. (2020). *A Grammar of Seenku*. De Gruyter Mouton.

Mendoza-Denton, N. (2011). The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a Chicano/a gang persona. *Journal of Linguistic Anthropology*, *21*(2), 261–280.

Michaud, A. (2008). Tones and intonation: Some current challenges. In R. Sock, S. Fuchs & Y. Laprie (Eds.), *ISSP 2008 Proceedings* (pp. 13–18). INRIA. https://issp2008.loria.fr/Proceedings/PDF/invited3.pdf

Michaud, A., & Vaissière, J. (2015). Tone and intonation: Introductory notes and practical recommendations. *KALIPHO (Kieler Arbeiten Zur Linguistik Und Phonetik)*, *3*, 43–80.

Morse, A. M. L. (1976). *A sketch of the phonology and morphology of Bobo (Upper Volta)* [Doctoral dissertation, Columbia University].

Mountford, K. W. (1983). *Bambara declarative sentence intonation* [Doctoral dissertation, Indiana University].

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation—Is it mediated by accentuation? *Journal of Phonetics*, *44*, 47–61.

Myers, S. (1996). Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African Linguistics*, *25*(1), 29–59.

Nakai, S., Kunnari, S., Turk, T., Alice E., Suomi, K., & Ylitalo, R. (2009). Utterance-final lengthening and quantity in Northern Finnish. *Journal of Phonetics*, *37*(1), 29–45.

Nakai, S., Turk, A. E., Suomi, K., Granlund, S., Ylitalo, R., & Kunnari, S. (2012). Quantity constraints on the temporal implementation of phrasal prosody in Northern Finnish. *Journal of Phonetics*, *40*(6), 796–807.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Foris Publications.

Nikitina, T. (2018). Focus marking and differential argument marking: The emergence of bidirectional case marking in Wan. In E. Adamou, K. Haude, & M. Vanhove (Eds.), *Information Structure in lesser-described langauges: Studies in prosody and syntax* (pp. 195–216). John Benjamins.

Ogden, R. (2004). Non-modal voice quality and turn-taking in Finnish. In E. Couper-Kuhlen & C. E. Ford (Eds.), *Sound Patterns in Interaction: Cross-linguistic Studies from Conversation* (pp. 29–62). John Benjamins.

Patin, C. (2016). Tone and intonation in Shingazidja. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 285–320). De Gruyter Mouton.

Peng, S.-H., Chang, M. K. M., Tseng, C., Huang, T., Lee, O. J., & Beckman, M. E. (2005). Towards a pan-Mandarin system for prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 230–270). Oxford University Press.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* [Doctoral dissertation, Massachusetts Institute of Technology].

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonation in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in Communication* (pp. 271–311). MIT Press.

Podesva, R. J. (2013). Gender and the social meaning of non-modal phonation types. *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, 427–448.

Prokhorov, K. (2014). Focalization particles in Bambara. *Mandenkan*, *52*, 60–72.

Rialland, A. (2009). The African lax question prosody: Its realisation and geographical distribution. *Lingua*, *119*, 928–949.

Rialland, A., & Aborobongui, M. E. (2016). How intonations interact with tones in Embosi (Bantu C25), a two-tone language without downdrift. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 195–222). De Gruyter Mouton.

Rialland, A., & Robert, S. (2001). The intonational system of Wolof. *Linguistics*, *39*(5), 893–939.

Rice, K. D. (1987). On defining the intonational phrase: Evidence from Slave. *Phonology Yearbook*, *4*, 37–59.

Rooth, M. (1992). A theory of focus interpretation. *Natural Langauge Semantics*, *1*, 75–116.

Sanou, J.-F. (1993). *La langue bobo de Tondogosso (Bobo-Dioulasso) (Haute Volta)* [Doctoral dissertation, Université Paris V].

Schreiber, H. (2008). Copula constructions in Mande—An overview. *Mandenkan*, *44*, 63–78.

Schwarszchild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, *7*, 141–177.

Schwarz, A. (2009). Tonal focus reflections in Buli and some Gur relatives. *Lingua*, *119*, 950–972.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology yearbook*, *3*, 371–405.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 550–559). Blackwell Publishing.

Selkirk, E. (2008). Contrastive focus, givenness and the unmarked status of "discourse-new." *Acta Linguistica Hungarica*, *5*(3–4), 331–336.

Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (2nd Edition, pp. 435–484). Blackwell Publishing.

Selkirk, E., & Tateishi, K. (1991). Syntax and downstep in Japanese. In C. Georgopoulos & R. Ishihara (Eds.), *Interdisciplinary approaches to language* (pp. 519–543). Springer.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–246.

Sherwood, K. (2016). Intonational phrase boundaries in Southern Bobo Madaré. *Speech Prosody 2016*, 89–93.

Simons, G.F., & Fennig, C.D. (2018). Bobo Madaré, Southern. In *Ethnologue: Languages of the world* (21st Edition). SIL International. https://www.ethnologue.com/language/bwq

Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Féry, C., & Krifa, M. (2006). *Questionnare on information structure: Reference manual*. Universitätsverlag Potsdam.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, *37*(1), 21–43.

Tiendrébéogo, B. (1998). *Rapport Sociolinguistique sur la Langue Bobo Madaré*. SIL International/ANTBA.

Truckenbrodt, H. (1995). *Phonological phrases: Their relation to syntax, focus, and prominence*. [Doctoral dissertation, Massachusetts Institute of Technology].

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*, 445–472.

Vydrin, V. (2002). Some hasty notes on the ways of evolution of Mande tonal systems. In R. Nicolaï & P. Zima (Eds.), *Lexical and structural diffusion: Interplay of internal and external factors of language development in the West African Sahel.* (pp. 243–264). Publications de la Faculté des Lettres, Arts et Sciences Humaines de Nice.

Vydrin, V. (2009). On the problem of the Proto-Mande homeland. *Journal of Language Relationship/Вопросы Языкового Родства*, *1*, 107–142.

Vydrin, V., & Diané, M. (2016). L'interrogation en maninka de Guinée. *Mandenkan*, *56*, 95–117.

Vydrina, A. (2017). *A corpus-based description of Kakabe, a Western Mande language: Prosody and grammar* [Doctoral dissertation, Institut National des Langues et Civilisations Orientales].

Wagner, M. (2005). *Prosody and recursion* [Doctoral dissertation, Massachusetts Institute of Technology].

Wagner, P., Origlia, A., Avesani, C., Christodoulides, G., Cutugno, F., D'Imperio, M., Escudero, D., Mancebo, B. G. F., Lacheret, A., Ludusan, B., Moniz, H., Ní Chasaide, A., Niebuhr, O., Rousier-Vercruyssen, L., Simon, A.-C., Šimko, J., Tesser, F., & Vainio, M. (2015). Different parts of the same elephant: A roadmap to disentangle and connect different perspectives on prosodic prominence. In M. Wolter, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow.

Welmers, W. E. (1971). Niger-Congo, Mande. In T. K. Sebeok (Ed.), *Linguistics in Sub-Saharan Africa* (pp. 113–140). De Gruyter Mouton.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Wong, W. Y. P., Chan, M. K. M., Beckman, M. E., & Jun, S.-A. (2005). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In *Prosodic typology: The phonology of intonation and phrasing* (pp. 271–300). Oxford University Press.

Xu, Y., & Wang, M. (2009). Organizing syllables into groups—evidence from F0 and duration patterns in Mandarin. *Journal of Phonetics*, *37*(4), 502–520.

Yang, Y., & Wang, B. (2002). Acoustic correlates of hierarchical prosodic boundary in Mandarin. *Speech Prosody 2002/*

Yip, M. (2002). *Tone*. Cambridge University Press.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, *85*(3), 315–337.

Zerbian, S. (2007). Investigating prosodic focus marking in Northern Sotho. In E. O. Aboh, K. Hartmann, & M. Zimmerman (Eds.), *Focus Strategies: Evidence from African Languages*. De Gruyter Mouton.

Zimmerman, M. (2007). Contrastive focus. In C. Féry, G. Fanselow, & M. Krifka (Eds.), *Interdisciplinary Studies on Information Structure* (Vol. 6, pp. 147–159).

Zimmerman, M., & Onea, E. (2011). Focus marking and focus interpretation. *Lingua*, *121*, 1651–1670.