

UNIVERSITY OF CALIFORNIA

Santa Barbara

The role of stress in Tibetan tonogenesis:
a study in historical comparative acoustics

A Dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Linguistics

by

Nancy Jill Caplow

Committee in charge:

Professor Carol Genetti, Co-Chair

Professor Matthew Gordon, Co-Chair

Professor Bernard Comrie

Professor Roland Bielmeier

March 2009

The dissertation of Nancy Jill Caplow is approved.

Roland Bielmeier

Roland Bielmeier

Bernard Comrie

Bernard Comrie

Carol Genetti

Carol Genetti, Committee Co-Chair

Matthew Gordon

Matthew Gordon, Committee Co-Chair

August 2008

The role of stress in Tibetan tonogenesis:
a study in historical comparative acoustics

Copyright © 2009

by

Nancy Jill Caplow

DEDICATION

This work is dedicated to the memory of:

My friend Chokyi Sherpa, who welcomed me with warmth and laughter into her home and into the Tokpe Gola community.

Mrs. Gertrude Collins, my French teacher at King Philip Junior High School in West Hartford, Connecticut, who taught me to love the structure of language.

Troya Bogard, whom I thought of so often as I wrote this dissertation.

ACKNOWLEDGEMENTS

It seems that the expertise, support, and friendship that made it possible for me to complete this dissertation came to me from all corners of the globe.

My first thanks go to my dissertation committee members in Santa Barbara, Leipzig, and Berne, all of whom read my work with great thoroughness and provided valuable suggestions for improvement. Carol Genetti worked so patiently to guide me along the way, and Matt Gordon helped me think through many of the key questions of phonetics and phonology, both large scale and small. It was always a pleasure to talk over ideas and questions with Bernard Comrie, whose humor and insight I hope to enjoy for many years to come. Roland Bielmeier took special care to review my every Tibetan word, with its spelling and etymology – despite which, I fear, some errors may remain.

It was also Roland Bielmeier, together with Jackson T.-S. Sun, who created the first opportunity for me to hear Balti Tibetan during the 8th Himalayan Languages Symposium in Berne, Switzerland. This is where my curiosity about Tibetan stress was piqued, and where the idea for this dissertation originated.

Nicolas Tournadre then made it possible for me to pursue this interest by inviting me to join him in Baltistan and Qinghai, introducing me to many valuable contacts and sharing his ideas and his “Pan-dialectal lexicon” with me. He was a wonderful companion as we traveled the Silk Road from one side of Tibet to the other. I would not have managed so well without his language skills and experience!

He has continued to be a great friend and colleague, in France as well as in Asia. I have benefited from and delighted in our conversations about Tibetan language and linguistics, and an infinity of other topics.

I owe a tremendous debt of gratitude, too, to the speakers of Tibetan dialects who shared their language and their time with me. Even as I measured fundamental frequency or plotted points on a graph, I was thinking always of the voices and the smiles of Tsunthar, Samden, Shamba, Konlo, Jigme Lokshe, Samden Tshering, and Lobsang Tender; of Muhammed Ali, Muhammed Hasnain, and Muhammed Raza Tasawor; and of Dandu, Pasang, Pupu, Mingma Dandu, Dawa Thake, Dochi, Ngawang Dochi, and Tashi. And there were many others, in Xining and Rebkong, and in Papung, Donggen, Gupha Pokhari, and Kathmandu.

The collection of data – and indeed, everything – was facilitated by the help of my research assistants and travel companions: Raza Tasawor in Baltistan; Sangye Gyatso in Rebkong (with additional help from young Tashi) and in Kathmandu; and Pasang Sherpa and Dandu Sherpa in Kathmandu and in villages along the Mewa Khola in northeastern Nepal. Dandu especially has become a great friend; I cannot thank him enough for his good company. It was always a pleasure to work and travel together, and I hope we will have opportunities to do so again. As much as his intelligence, energy, competence, initiative, interest, and attention to detail, his good humor and his ever-cheerful and encouraging disposition made all the difference.

There were others who offered good times, friendship, and assistance along the way, and in some cases hosted me graciously in their homes: Phuntsok and Betsy

were great fun and extremely helpful in Xining; the convenience of staying in Janet Rockwell's home in Kathmandu made the collection of final recordings much more comfortable and congenial.

I spent many, many happy hours in Kathmandu visiting with (and working with) Pasang and the late Chokyi Sherpa and their family and friends, and with Dandu Sherpa and his family. During extended stays in the villages of Papung and Donggen I was made welcome and comfortable by Suddha, Namgyal Drolma, Yangkyi, and Tsuang Nupu and their families. And I was honored to be invited to a festive dinner at the home of Raza Tasawor in Skardu.

Other friends in Kathmandu inspired and guided me in different ways, among them Linda S. Smith, Brian Peniston, Pam Ross and Charles Gay, Christine Daniels, and especially Mike Gill of the Fulbright Commission, and Steve LeClerq who helped launch my research on Tibetan dialects in Nepal. And I certainly learned a thing or two from the folks who hung out at the Yoga Hotel, and from the members of the Shiva Slaves motorcycle gang.

Back at UCSB, there were many others who contributed to my success. My first Tibetan teachers B. Alan Wallace and Greg Hillis helped me with various details as I developed this project. Stefan Gries launched me on the analysis of my data using R. I am especially grateful to Rich Harang for the hours he spent with me discussing approaches to statistical analysis and consideration of pitch slope data. Sandy Thompson has been an extraordinary friend and mentor over the years; I believe I have grown to be a better person for knowing her. Pat Clancy's encouragement and

support made all the difference at a crucial time. Mary Rae Staton watched over me, kept tabs on me, and encouraged me at every step.

I would also like to thank Paul Boersma, in Amsterdam, for so patiently responding to my endless queries about writing scripts for the Praat phonetics software. Bert Remijsen – then in Leiden, now in Edinburgh – brought my script-writing to a more advanced level.

Thanks to friends in Santa Barbara for encouragement and fun along the way – including Alyson Prude, Holley-the-Archaeolley Moyes, Nigel and Connie Buxton, Christine and Dave Bourgeois, and Annalee Locke – and to Jane Carlisle, Suzanne Marlow, and Victoria DeMaio who taught me to practice peace, strength, and kindness. And thanks to Professor Emeritus Walter Kohn (Physics) for the entertaining and inspiring conversations we shared as we rode the bus together.

Thanks to my *bahini* Brigitte Huber and her husband Kurt Keller for their friendship and hospitality in Thun, Switzerland, and for that crucial bit of help at the very end.

I would also like to thank my family for their patience and support in so many ways: Sylvia and David Caplow, Jeff Caplow, and Rich and Beth Caplow.

And I would like to express special thanks to my old friend *jbnftnek*, for limitless faith, encouragement and good cheer; and gratitude for the beauty of Santa Barbara and Hendry's Beach, where I did my thinking and learning during long walks in sunshine and in fog.

Finally, I am grateful for the financial support that made my study and research possible. UC Santa Barbara was most generous with grants, fellowships, and teaching positions over the years. My research on Tibetan dialects began in Nepal with a Fulbright scholarship, an opportunity which broadened my view of the world, changed me as a person, and led to many new friendships.

VITA OF NANCY JILL CAPLOW

January 2009

Education

- March 2009 PhD, Linguistics. University of California, Santa Barbara.
Reconstructing stress in Proto-Tibetan: a study in historical comparative acoustics.
Committee members: Carol Genetti (Co-chair), Matthew Gordon (Co-chair), Bernard Comrie, Roland Bielmeier (University of Berne)
- 1994 M.A., Linguistics. Indiana University. Bloomington, Indiana.
- 1982 B.A., Geology. Colgate University. Hamilton, New York.

Non-Degree Summer Programs

- 2005 LOT Summer School (Landelijke Onderzoekschool Taalwetenschap; Netherlands Graduate School of Linguistics). Leiden University.
- 2003 Summer intensive Tibetan language program. Rangjung Yeshe Institute, Kathmandu, Nepal.
- 2001 Linguistic Society of America, Linguistic Institute. UC Santa Barbara.
- 1999 Summer intensive Nepali language program. Cornell University.
- 1995 Linguistic Society of America, Linguistic Institute. University of New Mexico.
- 1994 Summer Language School, French. Middlebury College.
- 1993 Linguistic Society of America, Linguistic Institute. The Ohio State University.

Scholarships and Fellowships

- 2007 Graduate Division Dissertation Fellowship, UC Santa Barbara; tuition and stipend for the spring quarter.
- 2003-2004 Humanities Research Assistant, UC Santa Barbara; tuition and stipend for the academic year.
- 2002-2003 Humanities Research Assistant, UC Santa Barbara; tuition and stipend for the academic year.
- 2003 FLAS (Foreign Language and Area Studies Fellowship), United States Department of Education. Study of Lhasa Tibetan at the Rangjung Yeshe Institute, Kathmandu, Nepal.
- 2000-2002 Fulbright Scholar, United States Department of State / Institute of International Education. Research on the Tokpe Gola Tibetan language of northeastern Nepal.
- 2001 Summer tuition fellowship, Linguistic Institute, Linguistic Society of America.
- 1999-2000 Graduate Division Fee Fellowship, UC Santa Barbara.

Scholarships and Fellowships (cont.)

- 1999 Summer tuition fellowship, Cornell University. Nepali language program.
- 1997-1998 Regents Fellowship, UC Santa Barbara; tuition and stipend for the academic year.
- 1994 Summer tuition grant, Middlebury College. French school.
- 1993 Summer tuition fellowship, Linguistic Institute, Linguistic Society of America.
- 1993-1994 FLAS (Foreign Language and Area Studies Fellowship), United States Department of Education. Study of the Hausa language at Indiana University.

Research Grants

- 2006 Doctoral Student Travel Grant, Academic Senate, UC Santa Barbara. Travel grant for conference attendance.
- 2003 Pre-ABD Research Grant, UC Santa Barbara. Travel grant for conference attendance and archival research.
- 2002 Humanities/Social Sciences Research Grant, UC Santa Barbara. Field research on Tokpe Gola Tibetan; Kathmandu, Nepal.
- 1999 Research Assistant, National Science Foundation Grant. Preliminary field research in Nepal; primary investigator: Carol Genetti (Associate Professor, UC Santa Barbara).
- 1998 Humanities/Social Sciences Research Grant, UC Santa Barbara. Field research on the Western Abenaki language (Algonquian) of New England and Quebec.
- 1998 American Philosophical Society: Phillips Grant for research on Native American Indian languages. Field research on Western Abenaki.

Publications

- Nov 2007 'Directionals in Tokpe Gola Tibetan discourse'. *in* Trends in Linguistics: Linguistics of the Himalayas and Beyond. Roland Bielmeier and Felix Haller, *eds.* pp. 23-46. Berlin: Mouton de Gruyter.
- Nov 2004 'Fundamental frequency and intensity in Tokpe Gola, a tonal dialect of Tibetan'. PhD qualifying paper, UC Santa Barbara. (51 pages).
- Dec 2001 'get in English conversation: a Construction Grammar approach'. PhD qualifying paper, UC Santa Barbara. (88 pages).
- Sept 2000 'The epistemic marking system of émigré Dokpa Tibetan'. MA thesis equivalent, UC Santa Barbara. (85 pages).

Presentations

- Dec 2008 'Tracing Tibetan across space and time'. Department colloquium, University of North Texas; Denton, TX.
- Jan 2008 'The role of stress in Tibetan tonogenesis'. Laboratoire Parole et Langage, CNRS – Université de Provence.
- March 2007 'Tracing Tibetan across space and time'. Invited talk, Macalester College; St. Paul, MN.
- March 2007 'Languages and genes'. Invited talk, Macalester College; St. Paul, MN.
- Jan 2007 'The role of stress in Tibetan tonogenesis'. Dissertation colloquium, UC Santa Barbara; Santa Barbara, CA.
- Jan 2007 'Stress and tone in Tibetan'. Annual Meeting of the Linguistic Society of America; Anaheim, CA. 4-7 January, 2007.
- Sept 2006 'The reconstruction of stress in Proto-Tibetan'. 39th International Conference on Sino-Tibetan Languages and Linguistics; Seattle, WA. 15-17 September, 2006.
- May 2004 'Writing scripts for the Praat phonetics software'. UC Santa Barbara Linguistics Department colloquium.
- Sept 2002 'Directionals in Tokpe Gola Tibetan discourse'; 8th Himalayan Languages Symposium; Berne, Switzerland.
- July 2001 'A field report on phonemes and tones in Tokpe Gola, a Tibetan dialect of northeastern Nepal'; Workshop on Tibeto-Burman Languages; Santa Barbara, CA.
- Nov 2000 'Epistemics and evidentials in émigré Dokpa Tibetan'; Linguistic Society of Nepal; Kathmandu, Nepal.
- June 2000 'The focusing functions of yin and yod in émigré Dokpa Tibetan'; 6th Himalayan Languages Symposium; Milwaukee, WI.

Current position

- 2008-present Visiting Assistant Professor, Department of Linguistics and Technical Communication, University of North Texas, Denton, Texas.

Teaching experience (University of North Texas)

- Fall 2008 Principles of Language Study
Principles of Language Study (Internet course)
Phonology (Graduate course)
- Spring 2009 Principles of Linguistics
Structure of Modern English
Prosody in English and Other Languages of the World (Graduate seminar)

Teaching Experience (UC Santa Barbara)

Teaching Associate:	Fall 1999	Introduction to language and linguistics
Teaching Assistant:	Fall 2005	Global History, Culture, and Ideology
	Winter 2005	English as a Second Language - Pronunciation
	Summer 2002- Summer 2006	Freshman Summer Start Program: 6-week orientation program for incoming freshman
	Fall 2001	TA training course for Linguistics graduate students
	Spring 2000	The story of English (with Arthur Schwartz)
	Winter 2000	Phonology (with Matthew Gordon)
	Spring 1999	Morphology (with Marianne Mithun)
	Winter 1999, Winter 1998	Syntax (with Sandra Thompson)
	Fall 1998	Introduction to language and linguistics

Teaching Awards

2000	“Spotlight on Excellence” Award, presented by UC Santa Barbara undergraduate organization in appreciation for contributions in teaching, mentorship, and inspiration.
1999	Nomination, Graduate Student Association Excellence in Teaching Award, UC Santa Barbara. (nominated by Professor Sandy Thompson)
1998	Nomination, Graduate Student Association Excellence in Teaching Award, UC Santa Barbara. (nominated by Professor Sandy Thompson)

Collaborative research projects / Consulting

2008-2009	Ad hoc editor of Nicolas Tournadre, Gyurme Chodrak, Lhakpa Norbu Sherpa, and Guillaume Oisel’s Sherpa-English and English-Sherpa Dictionary (Forthcoming).
2006	Database editing and modification. Assisted Prof Stefan Gries, UC Santa Barbara Linguistics Department. Max Planck Institute (Leipzig) research project on the acquisition of English verb constructions (ReVerb).
2005-2006	Writing scripts for the Praat phonetics software. Assisted Profs. Carol Lord and Robert Berdan, California State University Long Beach. Research project on improving child literacy skills.
2004	Writing scripts for the Praat phonetics software. Assisted Prof Roger J. Ingham, UC Santa Barbara Department of Speech and Hearing Sciences. Research project on stuttering among adults and children.
2002-2003	Ad hoc editor of Nicolas Tournadre and Sangda Dorje’s Manual of Standard Tibetan: language and civilization (2003).

Linguistic Fieldwork

- Dec 2004 Comparative study of Tibetan dialects. Collected recordings of Rebkong Amdo Tibetan and Tokpe Gola Tibetan in Kathmandu, Nepal.
- Fall 2003 Comparative study of Tibetan dialects. Collected recordings of Balti (Baltistan, Pakistan), Amdo and Khams (Qinghai and Gansu Provinces, China), Lhasa (Tibetan Autonomous Region), and Tokpe Gola (Nepal) for acoustic analysis.
- 2000-2003 (intermittent) Endangered language documentation. Collected recordings of narrative and conversation of the Tokpe Gola language of northeastern Nepal. Worked with language consultants on transcription and translation.
- 1998 Endangered language documentation. Worked with elderly native speakers of Western Abenaki in Quebec and New England. Recorded and transcribed narratives and conversation.

Service activities

- 1998-1999 Organizing Committee, Second Annual Workshop on American Indigenous Languages (WAIL), a student-run conference held at UC Santa Barbara, May 1999.
- 1997-1998 Founding / Organizing Committee, Workshop on American Indigenous Languages (WAIL), UC Santa Barbara, May 1998.
- 1997-1998 Graduate Student Representative to faculty meetings, UC Santa Barbara Department of Linguistics.
- 1997-1998 Coordinator, Native American Indian Languages study group, UC Santa Barbara Department of Linguistics.

Other Professional Experience

- 1982-1984 Geologist. Anaconda Minerals Company; Watts Griffith & McQuat; Anchorage, Alaska. Worked as an exploration geologist in the Alaskan bush: conducted field testing and sample collection, assisted in preparation of scientific reports, lived with a small group of people in a remote area for extended periods of time.
- 1986-1992, 1994-1996 Hydrogeologist / Senior Project Manager. Wagner, Heindel, & Noyes, Inc., Burlington, Vermont. Worked as an environmental consultant: coordinated projects with budgets over \$500,000, designed research studies, directed staff, prepared scientific reports, testified at public hearings.

ABSTRACT

The role of stress in Tibetan tonogenesis:
a study in historical comparative acoustics

by

Nancy Jill Caplow

In this dissertation I demonstrate that stress has played an important role in the development of tone in Tibetan.

Using original field data and acoustic analysis, I compare two non-tonal “Archaic” varieties spoken at the western and eastern edges of the Tibetan language area: Balti (Baltistan, northern Pakistan) and Rebkong Amdo (Qinghai, China). In both dialects, disyllabic nouns, adjectives, and numerals are stressed on the second syllable (σ_2); disyllabic verbs, in contrast, are stressed on σ_1 .

These two speech areas are separated by the vast expanse of the Tibetan Plateau, where considerably different varieties of the language are spoken (the tonal “Innovative” dialects). Ruling out borrowing and coincidence as sources of their shared features, the similarities must be due to inheritance from a common parent. I thus reconstruct a pattern of σ_2 stress for nouns, adjectives, and numerals, and a contrasting pattern of σ_1 stress for verbs.

I also provide a robust statistical analysis of the acoustic correlates of stress in these two dialects. The σ_2 stress observed in nouns is conveyed by “pitch” in Balti, and by both “pitch” and “pitch slope” in Rebkong Amdo. The σ_1 stress observed in verbs is conveyed by both pitch and intensity in both dialects.

Just as stress can be reconstructed for Proto-Tibetan, so, too, can the acoustic correlates of stress. The fact that F0-related parameters (pitch and pitch slope) are the primary cues for σ_2 stress in nouns for both Balti and Amdo indicates that F0 was also prominent on σ_2 of Proto-Tibetan nouns. I use the term “historical comparative acoustics” to refer to this method of reconstructing proto acoustic patterns.

The prominent F0 reconstructed for σ_2 of Proto-Tibetan nouns accounts for constraints on tone patterns long-observed and long-unexplained in the geographically central dialects. There, tone on σ_1 can be either L or H, but tone on σ_2 can only be H. I offer a new explanation for this σ_2 H tone: it is an acoustic reflex of Proto-Tibetan stress. The pattern of F0 prominence has remained robust, but its function has shifted over time, from conveying stress to conveying tone.

TABLE OF CONTENTS

1.	Introduction.....	1
	1.1 Genetic affiliation and typological overview.....	4
	1.2 Dialect variation and distribution.....	5
	1.3 Correlations with Written Tibetan	7
	1.4 Reconstructing Proto-Tibetan	10
2.	Methodology.....	13
	2.1 Recordings and other materials.....	14
	2.1.1 Primary materials.....	14
	2.1.1.1 Subjects	14
	2.1.1.2 Recording procedure	17
	2.1.1.3 Word lists	19
	2.1.2 Supporting materials.....	24
	2.1.3 Transcription and transliteration.....	25
	2.1.4 Stress placement	26
	2.2 Data coding and data collection.....	27
	2.2.1 Coding in the filename.....	28
	2.2.2 Acoustic measurements and coding in textgrids.....	31
	2.2.3 Coding in the data file.....	39
	2.3 Analyzing pitch, intensity, and vowel duration data.....	41
	2.3.1 Iconic representation in calculations and graphs	41
	2.3.1.1 Calculating differences across syllables	41
	2.3.1.2 Graphing a single parameter.....	43
	2.3.1.3 Graphing two parameters	44
	2.3.2 Analyzing the acoustic correlates	47
	2.3.2.1 Pitch and intensity	48
	2.3.2.2 Vowel duration	54
	2.4 Analyzing pitch slope data.....	58
	2.4.1 Calculating differences in pitch slope across syllables.....	61
	2.4.2 Graphing pitch slope alone	66
	2.4.3 Graphing pitch and pitch slope together	71
	2.5 The notion of significance	74
	2.5.1 Statistical significance	74
	2.5.1.1 Interpreting box-and-whisker plots	75
	2.5.1.2 Paired-sample t-tests.....	80
	2.5.1.3 Welch t-tests	82

2.5.2	Perceptual significance	83
2.5.2.1	Perceptual significance for pitch	85
2.5.2.2	Perceptual significance for intensity.....	87
2.5.2.3	Perceptual significance for vowel duration	88
2.5.2.4	Perceptual significance for pitch slope.....	89
2.5.2.5	Summary of perceptual significance levels	92
2.5.3	Contextual significance.....	92
3.	Background on the dialects.....	96
3.1	Previous research	96
3.1.1	Zhongu Tibetan.....	96
3.1.2	Balti Tibetan	98
3.1.3	Amdo Tibetan	108
3.2	Consonant and vowel inventories	109
3.2.1	Balti inventories.....	109
3.2.2	Rebkong Amdo inventories	112
3.3	The syllable template	114
3.3.1	Balti syllable types.....	114
3.3.2	Rebkong Amdo syllable types	116
3.4	Monosyllabic words.....	121
3.4.1	Comparison to an Innovative dialect	122
3.4.2	Pitch in compound formation	129
3.4.3	Pitch in monosyllabic verbs.....	132
3.5	Disyllabic words	134
3.5.1	Compound words.....	135
3.5.2	Borrowed words.....	137
3.5.3	Clitics.....	138
3.5.4	Noun-verb minimal pairs.....	141
3.5.5	Lexical exceptions and anomalies	145
3.6	Words of three or more syllables	146
3.6.1	Longer words in Balti	146
3.6.2	Longer words in Rebkong Amdo.....	147
4.	Acoustic correlates of stress in Balti Tibetan.....	149
4.1	Summary of acoustic correlates of stress in Balti.....	150
4.1.1	Non-verbs	150
4.1.1.1	Nouns.....	150
4.1.1.2	Adjectives.....	152
4.1.1.3	Numerals	153
4.1.2	Verbs.....	154

4.2	Pitch	155
4.2.1	Pitch for speaker BSh_03	155
4.2.1.1	BSh_03 Nouns.....	157
4.2.1.2	BSh_03 Adjectives	160
4.2.1.3	BSh_03 Numerals.....	162
4.2.1.4	BSh_03 Verbs.....	164
4.2.2	Pitch for speaker BM_01	166
4.2.2.1	BM_01 Nouns	167
4.2.2.2	BM_01 Adjectives.....	175
4.2.2.3	BM_01 Verbs	180
4.3	Pitch slope.....	185
4.3.1	Pitch slope for speaker BSh_03	187
4.3.1.1	BSh_03 Nouns.....	189
4.3.1.2	BSh_03 Adjectives	212
4.3.1.3	BSh_03 Numerals.....	218
4.3.1.4	BSh_03 Verbs.....	221
4.3.2	Pitch slope for speaker BM_01.....	230
4.3.2.1	BM_01 Nouns	233
4.3.2.2	BM_01 Adjectives.....	245
4.3.2.3	BM_01 Verbs	249
4.3.3	The interaction of pitch and pitch slope.....	253
4.4	Intensity.....	258
4.4.1	Intensity for speaker BSh_03.....	258
4.4.1.1	BSh_03 Nouns.....	260
4.4.1.2	BSh_03 Adjectives	267
4.4.1.3	BSh_03 Numerals.....	272
4.4.1.4	BSh_03 Verbs.....	274
4.4.2	Intensity for speaker BM_01	276
4.4.2.1	BM_01 Nouns	277
4.4.2.2	BM_01 Adjectives.....	281
4.4.2.3	BM_01 Verbs	285
4.4.3	The interaction of pitch and intensity	288
4.5	Vowel duration.....	296
4.5.1	Vowel duration for speaker BSh_03.....	297
4.5.1.1	BSh_03 Nouns.....	298
4.5.1.2	BSh_03 Adjectives	305
4.5.1.3	BSh_03 Numerals.....	307
4.5.1.4	BSh_03 Verbs.....	308
4.5.2	Vowel duration for speaker BM_01	317
4.5.2.1	BM_01 Verbs	317
4.6	Recapitulation	322

5.	Acoustic correlates of stress in Rebkong Amdo Tibetan.....	324
5.1	Summary of acoustic correlates of stress in Rebkong Amdo	324
5.1.1	Non-verbs	324
5.1.1.1	Nouns.....	325
5.1.1.2	Adjectives.....	325
5.1.1.3	Numerals	326
5.1.2	Verbs.....	327
5.2	Pitch	328
5.2.1	Pitch for speaker AR_04.....	329
5.2.1.1	AR_04 Nouns	331
5.2.1.2	AR_04 Adjectives	334
5.2.1.3	AR_04 Verbs	336
5.2.2	Pitch for speaker AR_05.....	337
5.2.2.1	AR_05 Nouns	339
5.2.2.2	AR_05 Adjectives	352
5.2.2.3	AR_05 Numerals.....	358
5.2.2.4	AR_05 Verbs	362
5.3	Pitch slope.....	365
5.3.1	Pitch slope for speaker AR_04	366
5.3.1.1	AR_04 Nouns	368
5.3.1.2	AR_04 Adjectives	387
5.3.1.3	AR_04 Verbs	395
5.3.2	Pitch slope for speaker AR_05	399
5.3.2.1	AR_05 Nouns	401
5.3.2.2	AR_05 Adjectives	419
5.3.2.3	AR_05 Numerals	425
5.3.2.4	AR_05 Verbs	430
5.3.3	The interaction of pitch and pitch slope.....	437
5.4	Intensity.....	446
5.4.1	Intensity for speaker AR_04.....	447
5.4.1.1	AR_04 Nouns	450
5.4.1.2	AR_04 Adjectives	457
5.4.1.3	AR_04 Verbs	462
5.4.2	Intensity for speaker AR_05.....	464
5.4.2.1	AR_05 Nouns	467
5.4.2.2	AR_05 Adjectives	473
5.4.2.3	AR_05 Numerals.....	478
5.4.2.4	AR_05 Verbs	481
5.4.3	The interaction of pitch and intensity	483
5.5	Vowel duration.....	487
5.5.1	Vowel duration for speaker AR_04.....	488
5.5.1.1	AR_04 Nouns	489
5.5.1.2	AR_04 Adjectives	496
5.5.1.3	AR_04 Verbs.....	498

5.5.2	Vowel duration for speaker AR_05	500
5.5.2.1	AR_05 Nouns	500
5.5.2.2	AR_05 Adjectives	504
5.5.2.3	AR_05 Numerals	506
5.5.2.4	AR_05 Verbs	507
5.6	Recapitulation	510
6.	Summary, implications, and future research.....	512
6.1	Stress in Balti and Rebkong Amdo Tibetan.....	512
6.1.1	Stress patterns	512
6.1.2	Stress correlates	513
6.1.2.1	Stress correlates in non-verbs (σ_2 stress)	513
6.1.2.2	Stress correlates in verbs (σ_1 stress).....	516
6.2	Reconstructing stress in Proto-Tibetan	517
6.2.1	Stress patterns	518
6.2.2	Stress correlates	518
6.3	Background on Tibetan tone and tonogenesis	519
6.3.1	Tone patterns in the Innovative dialects of Tibetan.....	520
6.3.1.1	Monosyllabic words	520
6.3.1.2	Disyllabic words.....	523
6.3.2	Previous research on Tibetan tonogenesis	528
6.3.2.1	Matisoff (1973).....	529
6.3.2.2	Mazaudon (1977).....	530
6.3.2.3	Sun (1997, 2001)	533
6.3.2.4	Bielmeier (1988a).....	536
6.4	A new view of the role of stress in Tibetan tonogenesis	541
6.5	Implications regarding the direction of tone split in Tibetan.....	543
6.6	For future research... ..	544
	REFERENCES	546
	APPENDIX – WORD LIST	552

LIST OF FIGURES

Figure 2.1 Sample Praat textgrid: Rebkong Amdo word [çsam.'ba] <i>zamba</i> ‘bridge’	34
Figure 2.2 Full Praat textgrid: Rebkong Amdo word [çsam.'ba] <i>zamba</i> ‘bridge’	37
Figure 2.3 Syllable position and the number line	42
Figure 2.4 Graphing a single parameter: Pitch	43
Figure 2.5 Graphing two parameters: Pitch difference vs. intensity difference	46
Figure 2.6 Inconclusive intensity data	49
Figure 2.7 Controlling intensity for vowel height.....	50
Figure 2.8 Pitch in a noun with σ_2 open: [ts ^h a.'kʰ] <i>tshwa.khu</i> ‘salt’	53
Figure 2.9 Pitch in a noun with σ_2 closed: AR_04 [h ^h māi.'k ^h ʰŋ] <i>sman.khang</i> ‘hospital’	54
Figure 2.10 Inconclusive vowel duration data	56
Figure 2.11 Controlling for vowel quality and syllable closure type.....	57
Figure 2.12 Pitch slope / σ_1 upward, σ_2 downward: AR_04 [h ^h ʰɔ̄.'ma] <i>glang.ma</i> ‘alpine willow tree’	60
Figure 2.13 Pitch slope values across a disyllabic word	61
Figure 2.14 Pitch slope / σ_1 downward, σ_2 more downward: AR_04 [xtax.'tɕʰɕ] <i>rta.lcag</i> ‘horse whip’	63
Figure 2.15 Pitch slope values for [xtax.'tɕʰɕ] <i>rta.lcag</i> ‘horse whip’	64
Figure 2.16 Pitch slope / σ_1 more downward, σ_2 downward: AR_04 [xkʰ.'pa] <i>skud.pa</i> ‘thread’	65
Figure 2.17 Pitch slope values for [xkʰ.'pa] <i>skud.pa</i> ‘thread’	66
Figure 2.18 Graphing pitch slope measurements	67
Figure 2.19 Pitch slope relationships	68
Figure 2.20 Pitch slope relationships in quadrant (c).....	70
Figure 2.21 Graphing pitch difference vs. pitch slope difference: AR_04 nouns.....	72
Figure 2.22 Graphing pitch difference vs. pitch slope difference: AR_05 nouns.....	73
Figure 2.23 Boxplot illustrating distribution of pitch measurements	75
Figure 2.24 Boxplots without and with overlap.....	79
Figure 2.25 Frequency of pitch slope measurements and pitch slope differences.....	91
Figure 2.26 Pitch slope patterns arising from pitch patterns: BSh_03 [ra.'ma] <i>ra.ma</i> ‘goat’	93
Figure 2.27 Contextually significant pitch slope: AR_05 [k ^h a.'tʰɕ] <i>kha.btags</i> ‘khata, offering scarf’	94
Figure 3.1 [ʰ-] preceding a voiced σ_1 onset: AR_04 [ʰri.'mõ] <i>ri.mo</i> ‘drawing, picture’	120
Figure 3.2 [h ^h -] corresponding to a WT cluster: AR_05 [h ^h zar.'mõ] <i>gzar.mo</i> ‘steep’	121
Figure 3.3 Tokpe Gola (TG_06) / Pitch traces for [t̪a] <i>mda</i> ‘arrow’ and [t̪ā] <i>rta</i> ‘horse’ ...	124
Figure 3.4 Balti (BM_01) / Pitch traces for [nda] <i>mda</i> ‘arrow’ and [ɬta] <i>rta</i> ‘horse’	125
Figure 3.5 Rebkong Amdo (AR_05) / Pitch traces for [nda] <i>mda</i> ‘arrow’ and [xta] <i>rta</i> ‘horse’	126

Figure 3.6 Tokpe Gola (TG_06) / Pitch traces for [me] <i>me</i> ‘fire’ and [mẽ:] <i>sman</i> ‘medicine’	127
Figure 3.7 Balti (BM_01) / Pitch traces for [me] <i>me</i> ‘fire’ and [sman] <i>sman</i> ‘medicine’ ...	128
Figure 3.8 Rebkong Amdo (AR_05) / Pitch traces for [nyi] <i>me</i> ‘fire’ and [h ^ə men] <i>sman</i> ‘medicine’	129
Figure 3.9 Rebkong Amdo (AR_05) / Pitch traces for [ɸna] <i>sna</i> ‘nose’ and [tɕ ^h ɿy] <i>khrag</i> ‘blood’	130
Figure 3.10 Rebkong Amdo (AR_05) / Pitch trace [h ^ə na.tɕɿy] <i>sna.khrag</i> ‘nose blood, bloody nose’	131
Figure 3.11 Rebkong Amdo (AR_04) / Pitch traces for [xso:] <i>gso</i> / <i>gsos</i> ‘eat’ and [xsəl] <i>gsod</i> , <i>bsad</i> ‘kill’	133
Figure 3.12 Balti (BM_01) / Pitch traces for [stor] <i>stor</i> ‘lose’ and [tsəl] <i>btsal</i> ‘look for’ ...	134
Figure 3.13 Balti cliticized noun (BSh_03) / Pitch and intensity for [‘bod.la] <i>bod.la</i> ‘to Tibet’	139
Figure 3.14 BSh_03 / Pitch and intensity for [mbloq.‘pa] ‘ <i>brog.pa</i> ‘nomad, pastoralist’ ...	140
Figure 3.15 Rebkong Amdo cliticized noun / Pitch and intensity for [‘xtsiβ.na] <i>rtsib.na</i> ‘beside, at the side of’	141
Figure 3.16 BSh_03 / Pitch and intensity curves for [‘χmit.pa] <i>mid.pa</i> ‘to devour, to swallow’ and [χmit.‘paχ] ? ‘polo accessory for a horse’	142
Figure 4.1 BSh_03 / Isolation: Pitch contrasts	156
Figure 4.2 BSh_03 / Frame: Pitch contrasts	157
Figure 4.3 BSh_03 / Nouns: Pitch	159
Figure 4.4 BSh_03 / Adjectives: Pitch	161
Figure 4.5 BSh_03 / Numerals: Pitch	163
Figure 4.6 BSh_03 / Verbs: Pitch	165
Figure 4.7 BM_01 / Pitch contrasts	167
Figure 4.8 BM_01 / Nouns: Pitch	168
Figure 4.9 BM_01 / Nouns / Pitch trace: [broq.‘pa] ‘ <i>brog.pa</i> ‘nomad, pastoralist’	170
Figure 4.10 BM_01 / Nouns / Pitch trace: [ŋgo.‘ɳlat] <i>mgo.glad</i> ‘brain, mind’	171
Figure 4.11 BM_01 / Nouns: Pitch vs. σ2 closure	172
Figure 4.12 BM_01 / Nouns / σ2 Open: Pitch vs. vowel height	174
Figure 4.13 BM_01 / Adjectives: Pitch	175
Figure 4.14 BM_01 / Adjectives: Pitch vs. σ2 closure	177
Figure 4.15 BM_01 / Adjectives / Pitch trace: [t ^h äk.‘t ^h äk] <i>mthag.mthag?</i> ‘fat’	178
Figure 4.16 BM_01 / Adjectives / Pitch trace: [tɕ ^h un.‘tse] <i>chung.tshad?</i> ‘small, little’	179
Figure 4.17 BM_01 / Adjectives: Pitch vs. vowel height contrast	180
Figure 4.18 BM_01 / Verbs: Pitch	181
Figure 4.19 BM_01 / Verbs: Pitch contrasts vs. morphological category	184
Figure 4.20 AR_05 / Nouns / Isolation: Pitch slope	186
Figure 4.21 BSh_03 / Isolation: Pitch slope contrasts	188

Figure 4.22 BSh_03 / Frame: Pitch slope contrasts	189
Figure 4.23 BSh_03 / Nouns: Pitch slope	191
Figure 4.24 BSh_03 / Noun / Pitch trace: [kaŋ.'ma] <i>rkang.ma</i> ‘foot, leg’	194
Figure 4.25 BSh_03 / Noun / Pitch trace: [ra.'ma] <i>ra.ma</i> ‘goat’	196
Figure 4.26 BSh_03 / Noun / Pitch trace: [le.'mik] <i>lde.mig</i> ‘key’	197
Figure 4.27 BSh_03 / Noun / Pitch trace: [xlat.'pa] <i>klad.pa</i> / <i>glad.pa</i> ‘brain’	198
Figure 4.28 BSh_03 / Noun / Pitch trace: [smin.'ma] <i>smin.ma</i> ‘eyebrows’	199
Figure 4.29 BSh_03 / Noun / Pitch trace: [tʰik.'tɕʰu] <i>thig.chu</i> ‘water drop, leak’	200
Figure 4.30 BSh_03 / Noun / Pitch trace: [sa.'kʰɛt] <i>sa.khyad</i> ‘farm property, agricultural field’	201
Figure 4.31 BSh_03 / Noun / Pitch trace: [mu.'tik] <i>mu.tig</i> ‘pearl’	203
Figure 4.32 BSh_03 / Noun / Pitch trace: [spyɑŋ.'ku] <i>spyang.ku</i> ‘wolf’	204
Figure 4.33 BSh_03 / Nouns / Isolation: Pitch slope vs. σ_2 closure	205
Figure 4.34 BSh_03 / Noun / Isolation / Pitch trace: [bax.'pʰe] <i>bag.phye</i> ‘wheat flour’	206
Figure 4.35 BSh_03 / Noun / Frame / Pitch trace: [mbax.'pʰe] <i>bag.phye</i> ‘wheat flour’	207
Figure 4.36 BSh_03 / Nouns / Frame: Pitch slope	208
Figure 4.37 BSh_03 / Noun / Frame / Pitch trace: [ri.'dax] <i>ri.dwags</i> ‘ibex’	209
Figure 4.38 BSh_03 / Nouns / Frame: Pitch slope vs. σ_2 closure	210
Figure 4.39 BSh_03 / Nouns: Pitch difference vs. pitch slope difference	212
Figure 4.40 BSh_03 / Adjectives: Pitch slope	214
Figure 4.41 BSh_03 / Adjective / Isolation / Pitch trace: [mã.'mo] <i>mang.mo</i> ‘many’	215
Figure 4.42 BSh_03 / Adjective / Frame / Pitch trace: [mã.'mo] <i>mang.mo</i> ‘many’	216
Figure 4.43 BSh_03 / Adjectives: Pitch difference vs. pitch slope difference	217
Figure 4.44 BSh_03 / Numerals: Pitch slope	218
Figure 4.45 BSh_03 / Numerals / Pitch trace: [tɕub.'dun] <i>bcu.bdun</i> ‘seventeen’	219
Figure 4.46 BSh_03 / Numerals: Pitch difference vs. pitch slope difference	220
Figure 4.47 Pitch trace for BSh_03 ['χsøn.ma] <i>gson.ma</i> ‘to survive, to be nourished’ (isolation)	222
Figure 4.48 Pitch trace for BSh_03 ['xsøn.ma] <i>gson.ma</i> ‘to survive, to be nourished’ (frame)	223
Figure 4.49 Pitch trace for BSh_03 [tʰaq.pa] <i>'thag.pa</i> ‘to grind’ (isolation)	224
Figure 4.50 BSh_03 / Verbs: Pitch slope	225
Figure 4.51 BSh_03 / Verb / Isolation / Pitch trace: ['mm.ma] <i>sbyin.ma</i> ? ‘to give’	226
Figure 4.52 BSh_03 / Verb / Frame / Pitch trace: ['mm.ma] <i>sbyin.ma</i> ? ‘to give’	227
Figure 4.53 BSh_03 / Verb / Pitch trace: [tʰɿs.pa] <i>'khris.pa</i> ‘to entangle’	228
Figure 4.54 BSh_03 / Verbs: Pitch difference vs. pitch slope difference	230
Figure 4.55 BM_01 / Isolation: Pitch slope contrasts	231
Figure 4.56 BM_01 / Nouns: Pitch slope	233

Figure 4.57 BM_01 / Noun / Isolation / Pitch trace: [smor.'ðo] <i>mu.rdo</i> ‘border’	234
Figure 4.58 BM_01 / Noun / Isolation / Pitch trace: [ʃka.'lo] <i>skra.lo</i> ‘woman’s hair’	235
Figure 4.59 BM_01 / Nouns: Pitch slope vs. σ 2 closure	236
Figure 4.60 BM_01 / Noun / Isolation / Pitch trace: [na.'tiŋ] ?? ‘hat’	237
Figure 4.61 BM_01 / Noun / Isolation / Pitch trace: [ɦla.'kʰaŋ] <i>lha.khang</i> ‘temple’	238
Figure 4.62 BM_01 / Noun / Isolation / Pitch trace: [ŋgo.'rlat] <i>mgo.glad</i> ‘brain, mind’	239
Figure 4.63 BM_01 / Noun / Isolation / Pitch trace: [ɦa.'tekʰ] ?? ‘frog’	240
Figure 4.64 BM_01 / Nouns: Pitch slope vs. σ 2 closure, compared to Rebkong Amdo	242
Figure 4.65 BM_01 / Nouns: Pitch difference vs. pitch slope difference, by σ 2 closure ...	244
Figure 4.66 BM_01 / Adjectives / Isolation: Pitch slope	245
Figure 4.67 BM_01 / Adjectives / Isolation / Pitch trace: [tɕʰat.tɕʰat] <i>chad.chad</i> ‘short, few’	246
Figure 4.68 BM_01 / Adjectives / Isolation / Pitch trace: [tʰɣn.'mo] <i>mthon.mo</i> ‘high’	247
Figure 4.69 BM_01 / Adjectives: Pitch slope vs. σ 2 closure	248
Figure 4.70 BM_01 / Adjectives: Pitch difference vs. pitch slope difference	249
Figure 4.71 BM_01 / Verbs: Pitch slope	250
Figure 4.72 BM_01 / Verbs / Pitch trace: ['min.ma] <i>sbyin.ma</i> ? ‘to give’	251
Figure 4.73 BM_01 / Verbs / Pitch trace: ['skʰat.zer] <i>skad.zer</i> ‘to call, to shout’	252
Figure 4.74 BM_01 / Verbs: Pitch difference vs. pitch slope difference	253
Figure 4.75 BSh_03 / Isolation: Pitch differences vs. pitch slope differences	255
Figure 4.76 BSh_03 / Frame: Pitch differences vs. pitch slope differences	256
Figure 4.77 BM_01 / Isolation: Pitch differences vs. pitch slope differences	257
Figure 4.78 BSh_03 / Isolation: Intensity contrasts	259
Figure 4.79 BSh_03 / Frame: Intensity contrasts	260
Figure 4.80 BSh_03 / Nouns: Intensity	261
Figure 4.81 BSh_03 / Nouns / Isolation: Intensity vs. vowel height contrast	262
Figure 4.82 BSh_03 / Nouns / Isolation: Intensity difference vs. vowel height contrast ...	263
Figure 4.83 BSh_03 / Nouns / Frame: Intensity vs. vowel height contrast	265
Figure 4.84 BSh_03 / Nouns / Frame: Intensity difference vs. vowel height contrast	266
Figure 4.85 BSh_03 / Adjectives: Intensity	267
Figure 4.86 BSh_03 / Adjectives / Isolation: Intensity vs. vowel height contrast	268
Figure 4.87 BSh_03 / Adjectives / Isolation: Intensity difference vs. vowel height contrast	269
Figure 4.88 BSh_03 / Adjectives / Frame: Intensity vs. vowel height contrast	270
Figure 4.89 BSh_03 / Numerals: Intensity	272
Figure 4.90 BSh_03 / Numerals: Intensity vs. vowel height contrast	273
Figure 4.91 BSh_03 / Verbs: Intensity	275
Figure 4.92 BM_01 / Isolation: Intensity contrasts	277
Figure 4.93 BM_01 / Nouns: Intensity	278
Figure 4.94 BM_01 / Nouns: Intensity vs. vowel height contrast	279
Figure 4.95 BM_01 / Nouns / Isolation: Intensity difference vs. vowel height contrast	280

Figure 4.96 BM_01 / Adjectives: Intensity.....	282
Figure 4.97 BM_01 / Adjectives / Isolation: Intensity vs. vowel height contrast.....	283
Figure 4.98 BM_01 / Adjectives / Isolation: Intensity difference vs. vowel height contrast	284
Figure 4.99 BM_01 / Verbs: Intensity.....	286
Figure 4.100 BM_01 / Verbs: Intensity contrasts vs. morphological category.....	287
Figure 4.101 BSh_03 / Isolation: Pitch difference vs. intensity difference.....	289
Figure 4.102 BSh_03 / Isolation / Same vowel height: Pitch difference vs. intensity difference.....	291
Figure 4.103 BSh_03 / Frame: Pitch difference vs. intensity difference.....	292
Figure 4.104 BSh_03 / Frame / Same vowel height: Pitch difference vs. intensity difference	293
Figure 4.105 BM_01 / Isolation: Pitch difference vs. intensity difference.....	295
Figure 4.106 BM_01 / Isolation / Same vowel height: Pitch difference vs. intensity difference.....	296
Figure 4.107 BSh_03 / Nouns: Gross vowel duration (words with diphthongs and compensatory lengthening excluded).....	300
Figure 4.108 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration vs. syllable closure contrast duration (words with diphthongs and compensatory lengthening excluded) ...	302
Figure 4.109 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration differences vs. syllable closure contrast.....	303
Figure 4.110 BSh_03 / Adjectives: Gross vowel duration (words with compensatory lengthening circled).....	306
Figure 4.111 BSh_03 / Numerals: Vowel duration.....	308
Figure 4.112 BSh_03 / Verbs: Gross vowel duration.....	309
Figure 4.113 BSh_03 / Verbs / Same nucleus / Isolation: Vowel duration vs. syllable closure contrast.....	311
Figure 4.114 BSh_03 / Verbs / Same nucleus / Frame: Vowel duration vs. syllable closure contrast.....	315
Figure 4.115 BM_01 / Verbs: Gross vowel duration contrasts.....	318
Figure 5.1 AR_04 / Isolation: Pitch contrasts.....	330
Figure 5.2 AR_04 / Frame: Pitch contrasts.....	331
Figure 5.3 AR_04 / Nouns: Pitch.....	333
Figure 5.4 AR_04 / Adjectives: Pitch.....	335
Figure 5.5 AR_04 / Verbs: Pitch.....	337
Figure 5.6 AR_05 / Isolation: Pitch contrasts.....	338
Figure 5.7 AR_05 / Frame: Pitch contrasts.....	339
Figure 5.8 AR_05 / Nouns: Pitch.....	341
Figure 5.9 AR_05 / Nouns: Pitch vs. σ 2 closure.....	342
Figure 5.10 AR_05 / Nouns / Isolation / Pitch trace: [na.'pa] <i>nad.pa</i> 'sick person, patient'	343
Figure 5.11 AR_05 / Nouns / Isolation / Pitch trace: [xtsa.'th Δ] <i>rtswa.thang</i> 'pasture, grazing area'.....	344
Figure 5.12 AR_05 / Nouns / Frame / Pitch trace: [na.'pa] <i>nad.pa</i> 'sick person, patient' ...	345
Figure 5.13 AR_05 / Nouns / Frame / Pitch trace: [s ^h o.'nim] <i>bsod.nams</i> 'merit, luck'	346

Figure 5.14	AR_05/ Nouns / σ 2 Open: Pitch.....	347
Figure 5.15	AR_05 / Nouns / σ 2 Open / Isolation: Pitch vs. vowel height.....	348
Figure 5.16	AR_05 / Nouns / σ 2 Open / Isolation: Pitch differences vs. vowel height	349
Figure 5.17	AR_05 / Nouns / σ 2 Open / Frame: Pitch vs. vowel height contrast.....	351
Figure 5.18	AR_05 / Adjectives: Pitch	354
Figure 5.19	AR_05 / Adjectives / Isolation / Pitch trace: [ʒɛv.'lɛv] <i>leb.leb</i> ‘flat’	355
Figure 5.20	AR_05 / Adjectives / σ 2 Open / Isolation: Pitch vs. vowel height contrast.....	357
Figure 5.21	AR_05 / Numerals: Pitch.....	359
Figure 5.22	AR_05/ Numerals / Isolation: Pitch vs. vowel height contrast.....	360
Figure 5.23	AR_05 / Verbs: Pitch.....	363
Figure 5.24	AR_04 / Isolation / Pitch slope contrasts.....	366
Figure 5.25	AR_04 / Frame / Pitch slope contrasts	367
Figure 5.26	AR_04 / Nouns / Isolation: Pitch slope	368
Figure 5.27	AR_04 / Nouns / Isolation: Pitch difference vs. pitch slope difference.....	370
Figure 5.28	AR_04 / Noun / Pitch trace: [xka.'tuɯx] <i>skad.lugs</i> ‘dialect’	372
Figure 5.29	AR_04 / Noun / Pitch trace: [tʃ.'ma] <i>dri.ma</i> ‘smell, scent, odor’.....	373
Figure 5.30	AR_04 / Noun / Pitch trace: [xtaɕ.'tɕɯɕ] <i>rta.lcag</i> ‘horse whip’	375
Figure 5.31	AR_04 / Noun / Pitch trace: [ɕtʃɯ.'kɯ] <i>sprul.sku</i> ‘incarnation, incarnate lama’.....	376
Figure 5.32	AR_04 / Nouns / Isolation: Pitch slope vs. σ 2 closure	377
Figure 5.33	AR_04 / Nouns / Frame: Pitch slope.....	379
Figure 5.34	AR_04 / Nouns / Frame: Pitch slope vs. σ 2 closure	380
Figure 5.35	AR_04 / Noun / Frame / Pitch trace: [xka.'tʃx] <i>skad.lugs</i> ‘dialect’	381
Figure 5.36	AR_04 / Nouns / Frame: Pitch difference vs. pitch slope difference	383
Figure 5.37	AR_04 / Noun / Frame / Pitch trace: [s ^h a.'tɕa] <i>sa.cha</i> ‘land, property’	384
Figure 5.38	AR_04 / Noun / Frame / Pitch trace: [tsɔi.'ma] <i>btson.ma</i> ‘prisoner’	385
Figure 5.39	AR_04 / Noun / Frame / Pitch trace: [°dim.'bɯm] <i>rdo. 'bum</i> ‘sacred pile of 100,000 stones’.....	386
Figure 5.40	AR_04 / Adjectives / Isolation: Pitch slope contrasts	387
Figure 5.41	AR_04 / Adjectives / Frame: Pitch slope contrasts	389
Figure 5.42	AR_04 / Adjectives: Pitch difference vs. pitch slope difference.....	391
Figure 5.43	AR_04 / Adjectives / Pitch trace: [tɕ ^h e.'po] <i>chen.po</i> ‘big’	392
Figure 5.44	AR_04 / Adjectives / Frame / Pitch trace: [tɕ ^h e.'po] <i>chen.po</i> ‘big’	393
Figure 5.45	AR_04 / Adjectives / Pitch trace: [s ^h o.ma] <i>so.ma</i> ‘new’	394
Figure 5.46	AR_04 / Adjectives / Frame / Pitch trace: [s ^h o.ma] <i>so.ma</i> ‘new’	395
Figure 5.47	AR_04 / Verbs: Pitch slope	396
Figure 5.48	AR_04 / Verbs / Frame / Pitch trace: ['ʒɛ̃.tɕɯv] <i>lan.brgyab</i> ‘to answer’	397
Figure 5.49	AR_04 / Verbs / Frame / Pitch trace: ['nɒm.biʋ] <i>gnam.babs</i> ‘to rain’	398
Figure 5.50	AR_04 / Verbs: Pitch difference vs. pitch slope difference.....	399
Figure 5.51	AR_05 / Isolation / Pitch slope contrasts.....	400
Figure 5.52	AR_05 / Frame / Pitch slope contrasts	401

Figure 5.53	AR_05 / Nouns: Pitch slope	403
Figure 5.54	AR_05 / Nouns / Isolation: Pitch slope vs. σ 2 closure	405
Figure 5.55	AR_05 / Nouns / Isolation / Pitch trace: [xsam.'ba] <i>zam.ba</i> 'bridge'	406
Figure 5.56	AR_05 / Nouns / Isolation / Pitch trace: [na.'ma] <i>mna'.ma</i> 'bride'	407
Figure 5.57	AR_05 / Nouns / Isolation / Pitch trace: [k ^h a.'tʂʂ] <i>kha.btags</i> 'khata, offering scarf'	408
Figure 5.58	AR_05 / Nouns / Isolation / Pitch trace: [h ^a raŋ.'t ^h ʊŋ] <i>ring.thung</i> 'length'	410
Figure 5.59	AR_05 / Nouns / Isolation: Pitch difference vs. pitch slope difference	411
Figure 5.60	AR_05 / Nouns / Frame: Pitch slope vs. morphological composition	414
Figure 5.61	AR_05 / Nouns / Frame: Pitch slope vs. σ 2 closure	414
Figure 5.62	AR_05 / Nouns / Frame / Pitch trace: [χɔim.'bo] <i>dpon.po</i> 'official, chief'	415
Figure 5.63	AR_05 / Nouns / Frame / Pitch trace: [na.'do] <i>sna.tha? sna.tho? sna.rdul?</i> 'snuff'	416
Figure 5.64	AR_05 / Nouns / Frame / Pitch trace: [tʂɛn.'ts ^h ʌŋ] <i>khyim.tshang</i> 'household, family'	417
Figure 5.65	AR_05 / Nouns / Frame: Pitch difference vs. pitch slope difference	418
Figure 5.66	AR_05 / Adjectives / Isolation: Pitch slope contrasts	420
Figure 5.67	AR_05 / Adjectives / Isolation / Pitch trace: [s ^h o.'ma] <i>so.ma</i> 'new'	421
Figure 5.68	AR_05 / Adjectives / Frame / Pitch trace: [s ^h o.'ma] <i>so.ma</i> 'new'	422
Figure 5.69	AR_05 / Adjectives / Frame: Pitch slope contrasts	423
Figure 5.70	AR_05 / Adjectives: Pitch difference vs. pitch slope difference	424
Figure 5.71	AR_05 / Numerals: Pitch slope contrasts	426
Figure 5.72	AR_05 / Numerals / Isolation / Pitch trace: [zʂb.'tʂʂ] <i>bzhi.bcu</i> 'forty'	427
Figure 5.73	AR_05 / Numerals / Isolation / Pitch trace: [kʂb.'tʂʂ] <i>dgu.bcu</i> 'ninety'	428
Figure 5.74	AR_05 / Numerals: Pitch difference vs. pitch slope difference	429
Figure 5.75	AR_05 / Verbs: Pitch slope contrasts	431
Figure 5.76	AR_05 / Verbs / Isolation / Pitch trace: ['nda.hen] <i>mda'. 'phen</i> 'to shoot an arrow'	433
Figure 5.77	AR_05 / Verbs / Isolation / Pitch trace: ['k ^h ʌŋ.bʌp] <i>gangs.babs?</i> 'to snow' ...	434
Figure 5.78	AR_05 / Verbs / Frame / Pitch trace: ['nda.hen] <i>mda'. 'phen</i> 'to shoot an arrow'	435
Figure 5.79	AR_05 / Verbs / Frame / Pitch trace: ['k ^h ʌŋ.bʌp] <i>gangs.babs?</i> 'to snow'	436
Figure 5.80	AR_05 / Verbs: Pitch difference vs. pitch slope difference	437
Figure 5.81	AR_04 / Isolation: Pitch difference vs. pitch slope difference	438
Figure 5.82	AR_04 / Frame: Pitch difference vs. pitch slope difference	440
Figure 5.83	AR_05 / Isolation: Pitch difference vs. pitch slope difference	441
Figure 5.84	AR_05 / Isolation / Control groups: Pitch difference vs. pitch slope difference	442
Figure 5.85	AR_05 / Frame: Pitch difference vs. pitch slope difference	444
Figure 5.86	AR_05 / Frame / Control groups: Pitch difference vs. pitch slope difference	445

Figure 5.87 AR_04 / Isolation: Intensity contrasts	448
Figure 5.88 AR_04 / Frame: Intensity contrasts	450
Figure 5.89 AR_04 / Nouns: Intensity	451
Figure 5.90 AR_04 / Nouns / Isolation: Intensity vs. vowel height contrast	452
Figure 5.91 AR_04 / Nouns / Isolation: Intensity differences vs. vowel height contrasts .	453
Figure 5.92 AR_04 / Nouns / Frame: Intensity vs. vowel height contrast	455
Figure 5.93 AR_04 / Nouns / Frame: Intensity differences vs. vowel height contrasts.....	456
Figure 5.94 AR_04 / Adjectives: Intensity	458
Figure 5.95 AR_04 / Adjectives / Isolation: Intensity vs. vowel height contrast.....	459
Figure 5.96 AR_04 / Adjectives / Frame: Intensity vs. vowel height contrast	460
Figure 5.97 AR_04 / Adjectives: Intensity with vowels of same height	461
Figure 5.98 AR_04 / Verbs: Intensity	463
Figure 5.99 AR_05 / Isolation: Intensity contrasts	465
Figure 5.100 AR_05 / Frame: Intensity contrasts	467
Figure 5.101 AR_05 / Nouns: Intensity	468
Figure 5.102 AR_05 / Nouns / Isolation: Intensity vs. vowel height contrast	469
Figure 5.103 AR_05 / Nouns / Isolation: Intensity differences vs. vowel height contrasts	470
Figure 5.104 AR_05 / Nouns / Frame: Intensity differences vs. vowel height contrasts....	472
Figure 5.105 AR_05 / Adjectives: Intensity	474
Figure 5.106 AR_05 / Adjectives / Isolation: Intensity vs. vowel height contrast.....	475
Figure 5.107 AR_05 / Adjectives / Frame: Intensity vs. vowel height contrast	477
Figure 5.108 AR_05 / Numerals: Intensity	479
Figure 5.109 AR_05 / Numerals: Intensity vs. vowel height contrast	480
Figure 5.110 AR_05 / Verbs: Intensity	482
Figure 5.111 AR_04 / Isolation: Pitch difference vs. intensity difference	484
Figure 5.112 AR_04 / Frame: Pitch difference vs. intensity difference	485
Figure 5.113 AR_05 / Isolation: Pitch difference vs. intensity difference	486
Figure 5.114 AR_05 / Frame: Pitch difference vs. intensity difference	487
Figure 5.115 AR_04 / Nouns: Gross vowel duration.....	490
Figure 5.116 AR_04 / Nouns / Isolation: Vowel duration vs. σ 2 closure	492
Figure 5.117 AR_04 / Nouns / Frame: Vowel duration vs. syllable closure	493
Figure 5.118 AR_04 / Nouns / Frame / Control group: Vowel duration contrasts	495
Figure 5.119 AR_04 / Adjectives: Gross vowel duration vs. syllable closure	497
Figure 5.120 AR_04 / Verbs: Gross vowel duration contrasts	499
Figure 5.121 AR_05 / Nouns: Gross vowel duration contrasts	501
Figure 5.122 AR_05 / Nouns / Frame / Control group: Vowel duration contrasts	503
Figure 5.123 AR_05 / Adjectives: Gross vowel duration contrasts	505
Figure 5.124 AR_05 / Numerals / Isolation: Gross vowel duration contrasts	507
Figure 5.125 AR_05 / Verbs: Gross vowel duration contrasts	508

LIST OF TABLES

Table 2.1 Lexical category codes.....	29
Table 2.2 Morphological structure codes.....	30
Table 2.3 Acoustic measurements collected using Praat	33
Table 2.4 Setting values used in the Praat script.....	33
Table 2.5 Syllable properties coded.....	40
Table 2.6 Word properties coded.....	40
Table 2.7 BSh_03 / Isolation / Pitch data: results of paired-sample t-tests (two-tailed).....	82
Table 2.8 AR_05 / Nouns / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)	83
Table 2.9 Perceptually significant pitch slopes: Rebkong Amdo nouns / isolation.....	90
Table 2.10 Perceptual significance levels	92
Table 3.1 Sprigg’s list of non-stress-bearing suffixes (2002: 5).....	101
Table 3.2 Balti [o.o] nouns and adjectives.....	105
Table 3.3 Balti: inventory of simplex consonants	110
Table 3.4 Balti: Inventory of consonant clusters (preliminary).....	111
Table 3.5 Balti: vowel inventory	111
Table 3.6 Rebkong Amdo: inventory of simplex consonants.....	112
Table 3.7 Rebkong Amdo: Inventory of consonant clusters.....	113
Table 3.8 Rebkong Amdo: vowel inventory.....	113
Table 3.9 Balti / Possible syllable types	115
Table 3.10 Balti / Examples of open syllables.....	115
Table 3.11 Balti / Examples of closed syllables	116
Table 3.12 Rebkong Amdo / Possible syllable types.....	117
Table 3.13 Rebkong Amdo / Examples of open syllables	118
Table 3.14 Rebkong Amdo / Examples of closed syllables.....	119
Table 3.15 ‘arrow’, ‘horse’: Non-tonal vs. tonal dialects	122
Table 3.16 ‘fire’, ‘medicine’: Non-tonal vs. tonal dialects	126
Table 3.17 Balti: Stress in borrowed nouns	138
Table 3.18 Balti noun-verb minimal pairs and near-minimal pairs	143
Table 3.19 Balti / Words of more than two syllables.....	147
Table 3.20 Rebkong Amdo / Words of more than two syllables	148
Table 4.1 BSh_03 / Nouns / Pitch: results of paired-sample t-tests (two-tailed).....	160
Table 4.2 BSh_03 / Adjectives / Pitch: results of paired-sample t-tests (two-tailed)	162
Table 4.3 BSh_03 / Numerals / Pitch: results of paired-sample t-tests (two-tailed).....	163
Table 4.4 BSh_03 / Verbs / Citation / Pitch: results of paired-sample t-tests (two-tailed)..	166
Table 4.5 BM_01 / Nouns / Pitch: Results of paired-sample t-tests (two-tailed)	173
Table 4.6 BM_01 / Nouns / σ 2 Open /Pitch: Results of paired-sample t-tests (two-tailed)	174
Table 4.7 BM_01 / Adjectives / Pitch: results of paired-sample t-tests (two-tailed).....	176
Table 4.8 BM_01 / Verbs: t-tests comparing stress correlates vs. morphological category (Welch)	182
Table 4.9 BM_01 / Verbs / Pitch: Results of paired-sample t-tests (two-tailed).....	185

Table 4.10 BSh_03 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)	202
Table 4.11 BSh_03 / Nouns / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed)	210
Table 4.12 BSh_03 / Adjectives / Pitch slope: results of paired-sample t-tests (two-tailed).....	217
Table 4.13 BSh_03 / Numerals / Pitch slope: results of paired-sample t-tests (two-tailed)	220
Table 4.14 BSh_03 / Verbs / Isolation / Pitch slope: results of paired-sample t-tests (two-tailed)	229
Table 4.15 BM_01 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)	241
Table 4.16 BM_01 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed).....	248
Table 4.17 BM_01 / Verbs / Pitch slope: Results of paired-sample t-tests (two-tailed).....	252
Table 4.18 BSh_03 / Nouns / Isolation: Intensity by height: paired-sample t-tests (two-tailed)	264
Table 4.19 BSh_03 / Nouns / Frame / Intensity by height: Paired-sample t-tests (two-tailed)	266
Table 4.20 BSh_03 / Adjectives / Isolation / Intensity by height: Paired-sample t-tests (two-tailed)	269
Table 4.21 BSh_03 / Adjectives / Frame / Intensity by height: Paired-sample t-tests (two-tailed)	271
Table 4.22 BSh_03 / Numerals / Intensity: results of paired-sample t-tests (two-tailed)	274
Table 4.23 BSh_03 / Verbs / Citation / Intensity: results of paired-sample t-tests (two-tailed)	276
Table 4.24 BM_01 / Nouns / Isolation / Intensity by height: paired t-tests (two-tailed)	281
Table 4.25 BM_01 / Adjectives / Intensity by height: paired-sample t-tests (two-tailed) ...	285
Table 4.26 Intensity on Balti verbs: results of paired-sample t-tests (two-tailed)	288
Table 4.27 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration by syllable closure type: paired-sample t-tests (two-tailed)	304
Table 4.28 BSh_03 / Nouns / Same nucleus / Same closure / Frame: Vowel duration ratios	304
Table 4.29 BSh_03 / Verbs / [closed.open] / Same nucleus / Isolation / Vowel duration: results of paired-sample t-tests (two-tailed).....	312
Table 4.30 BSh_03 / Verbs / Isolation / [closed.open] / Same nucleus: Vowel duration ratios	313
Table 4.31 BSh_03 / Verbs / [closed.open] / Same nucleus / Frame / Vowel duration: results of paired-sample t-tests (two-tailed)	315
Table 4.32 BSh_03 / Verbs / Frame / [closed.open] / Same nucleus: Vowel duration ratios	316
Table 4.33 BM_01 / Verbs / Vowel duration: Results of paired-sample t-tests (two-tailed)	320
Table 4.34 BM_01 / Verbs / N+Vblzr: Vowel duration ratios	320
Table 4.35 BM_01 / Verbs / Citation / [closed.open]: Vowel duration ratios	321
Table 4.36 Acoustic correlates of stress in Balti.....	323
Table 5.1 AR_04 / Nouns / Pitch: Results of paired-sample t-tests (two-tailed)	334
Table 5.2 AR_04 / Adjectives / Pitch: Results of paired-sample t-tests (two-tailed)	336

Table 5.3 AR_05 / Nouns / σ 2 Open / Isolation: Pitch by vowel height contrast: paired-sample t-tests (two-tailed).....	350
Table 5.4 AR_05 / Nouns / σ 2 Open / Frame: Pitch by vowel height contrast: paired-sample t-tests (two-tailed).....	352
Table 5.5 AR_05 / Adjectives / Frame / Pitch: Results of paired-sample t-tests (two-tailed).....	353
Table 5.6 AR_05 / Adjectives / σ 2 Open / Isolation / Pitch: Results of paired-sample t-tests (two-tailed).....	358
Table 5.7 AR_05 / Numerals / Pitch by vowel height contrast: Paired-sample t-tests (two-tailed).....	361
Table 5.8 AR_05 / Verbs / Pitch: Results of paired-sample t-tests (two-tailed).....	365
Table 5.9 AR_04 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed).....	369
Table 5.10 AR_04 / Nouns / Pitch slope: Results of paired-sample t-tests (two-tailed).....	380
Table 5.11 AR_04 / Nouns / Potential stress correlates: [xka.'tux] <i>skad.lugs</i> 'dialect'	382
Table 5.12 AR_04 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed).....	388
Table 5.13 AR_04 / Adjectives / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed).....	389
Table 5.14 AR_05 / Nouns / Pitch slope: Results of paired-sample t-tests (two-tailed).....	404
Table 5.15 AR_05 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed).....	420
Table 5.16 AR_05 / Adjectives / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed).....	423
Table 5.17 AR_05 / Numerals / Pitch slope: Results of paired-sample t-tests (two-tailed).....	428
Table 5.18 AR_05 / Verbs / Pitch slope: Results of paired-sample t-tests (two-tailed)	432
Table 5.19 AR_04 / Nouns / Isolation / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)	454
Table 5.20 AR_04 / Nouns / Frame / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)	457
Table 5.21 AR_04 / Adjectives with same vowel height / Intensity: Results of paired-sample t-tests (two-tailed).....	462
Table 5.22 AR_05 / Nouns / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)	471
Table 5.23 AR_05 / Nouns / Frame / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)	472
Table 5.24 AR_05 / Adjectives / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed).....	476
Table 5.25 AR_05 / Adjectives / Frame: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed).....	478
Table 5.26 AR_05 / Numerals / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed).....	481
Table 5.27 AR_05 / Verbs / Intensity: Results of paired-sample t-tests (two-tailed)	483
Table 5.28 AR_04 / Nouns / Frame / Vowel duration / Control group: Results of paired-sample t-tests (two-tailed).....	495

Table 5.29 AR_04 / Verbs / Frame / [closed.open] / Same nucleus: Vowel duration ratios	500
Table 5.30 AR_05 / Nouns / Frame / Vowel duration / Control group: Results of paired-sample t-tests (two-tailed).....	504
Table 5.31 AR_05 / Verbs / Isolation: Vowel duration differences and ratios.....	509
Table 5.32 AR_05 / Verbs / Frame: Vowel duration differences and ratios.....	510
Table 5.33 Acoustic correlates of stress in Rebkong Amdo	511
Table 6.1 Meaningful acoustic correlates of stress for nouns.....	514
Table 6.2 Meaningful acoustic correlates of stress for adjectives	515
Table 6.3 Meaningful acoustic correlates of stress for numerals.....	516
Table 6.4 Meaningful acoustic correlates of stress for verbs.....	517
Table 6.5 Tokpe Gola Tibetan: Tone patterns in monosyllabic non-verbs.....	521
Table 6.6 Tokpe Gola Tibetan: Tone patterns in disyllabic non-verbs.....	524
Table 6.7 Tokpe Gola Tibetan: Tone patterns in disyllabic verbs	525

1. Introduction

The vast Tibetan language area – which extends across parts of Pakistan, India, Nepal, China, and Bhutan – provides an extraordinary living laboratory for the study of language variation and change. This dissertation focuses on variation and change in two prosodic features: stress and tone.

Using original field data and acoustic analysis, I compare the stress patterns of two non-tonal dialects of Tibetan: Balti, spoken in Baltistan, northern Pakistan (at the western extreme of the Tibetan language area); and Rebkong Amdo, spoken in Qinghai Province, China (at the eastern extreme of the Tibetan language area). For both dialects, I demonstrate that disyllabic non-verbs (nouns, adjectives, and numerals) are stressed on the second syllable (σ_2), and that this pattern is conveyed primarily by the fundamental-frequency-related parameters pitch and pitch slope. In contrast, I demonstrate that disyllabic verbs are stressed on the first syllable (σ_1), and that this pattern is conveyed acoustically by both pitch and intensity.

Ruling out borrowing and coincidence as sources of these common characteristics, I conclude that the stress patterns of Balti and Rebkong Amdo are derived from their common parent, Proto-Tibetan. That is, I reconstruct a pattern of σ_2 stress for disyllabic Proto-Tibetan non-verbs, and a pattern of σ_1 stress for disyllabic Proto-Tibetan verbs.

Furthermore, through a method I refer to as “historical comparative acoustics”¹, I reconstruct the acoustic correlates of these proto stress patterns: the σ_2 stress of Proto-Tibetan non-verbs was conveyed by pitch and pitch slope, while the σ_1 stress of Proto-Tibetan verbs was conveyed by pitch and intensity.

This reconstruction begs the question of what happened to stress in those Tibetan dialects which innovated tone as a lexically contrastive feature. Do these dialects, too, show evidence of – or relicts of – second-syllable stress? Did consistent patterns of stress evolve into consistent patterns of tone?

I begin to address these questions by considering disyllabic words in Tokpe Gola Tibetan, a tonal dialect spoken in northeastern Nepal. I demonstrate that the tone patterns observed on disyllabic words are acoustically consistent with the historical stress patterns. That is, I suggest that stress has played an important role in constraining the possible tone patterns, and that there has been a shift in the function of acoustic resources over time.

These findings have implications regarding the direction of tone split in Tibetan. They also suggest a number of questions that will be interesting to investigate in future: (a) Could the stress patterns I identify here be productively analyzed as pitch accent? (b) Is there evidence of stress in the tonal dialects, and has a shift in stress occurred over time? (c) What is the domain of tone in Tibetan? and (d) Is it appropriate to draw a rigid distinction between stress as syntagmatic, and tone as paradigmatic?

¹ I thank Bernard Comrie for suggesting this term.

In the present chapter I provide some general background on the Tibetan language, and describe dialect variation and distribution.

Chapter 2 is a detailed account of the methodology I used in this study. This includes information about the collection of recorded materials; techniques used in the measurement of acoustic parameters; data coding; and data analysis and interpretation.

Chapter 3 provides background information relevant to my study. After summarizing previous research, I present my own observations on the phonetics and phonology of Balti and Rebkong Amdo, including consonant and vowel inventories, syllable types, and prosodic patterns in words of different lengths.

Chapters 4 and 5 are at the heart of this work. They comprise the detailed analysis of the acoustic correlates of stress for non-verbs and verbs in Balti and Rebkong Amdo, respectively.

In Chapter 6 I summarize the conclusions of the previous two chapters. Based on these findings, I offer a reconstruction of stress in Proto-Tibetan and present a hypothesis regarding the role of stress in Tibetan tonogenesis. (This entails consideration of previous research on Tibetan tonogenesis; this was not included in Chapter 3, because it is more relevant here.) I briefly introduce several issues which merit further thought and investigation: the direction of tone split in Tibetan; stress shift in Tibetan; and the questions of whether a pitch accent analysis of Tibetan tone might be advantageous, and whether tone might sometimes be regarded as syntagmatic.

1.1 Genetic affiliation and typological overview

The Tibetan language constitutes the Central Bodish branch of the Bodic division of the Tibeto-Burman family tree (Shafer 1974, Bradley 1997, van Driem 2001: 828ff). Tibeto-Burman itself is a branch of the larger Sino-Tibetan language family.

Bielmeier (CDTD, n.d.) identifies Balti (northern Pakistan) as one of the Western Archaic dialects of Tibetan. (As opposed to the Western Innovative dialects.) Rebkong Amdo (Qinghai, China) is one of the Conservative Amdo Nomad dialects of the Eastern Amdo dialects group. Tokpe Gola belongs to Bielmeier's Central Tibetan group – more specifically, it seems, among the Northern Nepalese Border Area dialects of eastern Nepal (along with Lhomi and Halung/Walung). Tournadre (p.c. 2006) considers Tokpe Gola to be one of the Tö (*stod*) dialects of Central Tibetan. (These classifications are based partly on linguistic criteria and partly on geographic criteria.)

Tibetan is a fairly strict SOV language. It is suffixing and postpositional. Numerals, quantifiers, and determiners follow the noun within a noun phrase; adjectives also usually follow the noun, but may sometimes precede it. Core arguments are marked by an ergative case-marking system. Frequently occurring oblique cases are genitive, instrumental, dative-locative, ablative, and comitative. Many case markers also function as enclitic clause-connectors. Tibetan is also

characterized by a rich system of evidential / epistemic marking, and some dialects employ an honorific register.

It seems that most words in Tibetan are monosyllabic or disyllabic. Denwood (1999: 88) notes that monosyllabic words often denote "... many of the common referents of daily conversation", including words for people, body parts, nature and the elements, agriculture, and household objects. Trisyllabic and quadrisyllabic words also occur, but are most often composed of disyllabic and monosyllabic elements joined together.

As discussed in detail below, dialects range from non-tonal to tonal. They exhibit various degrees of vowel harmony.

1.2 Dialect variation and distribution

Tibetan exhibits a tremendous degree of internal variation. Based on geographic and linguistic criteria, Bielmeier (n.d.; p.c. July 2005) identifies five main dialect groups which encompass at least 129 varieties. According to Tournadre (2005, 2008), Tibetan is comprised of 25 major groups which are quite distinct, and, within these groups, more than 220 different dialects.

In some cases, dialects are quite similar to one another, differing only in terms of minor phonological details or lexical items. In other cases the differences result in mutual incomprehensibility.

Based on their phonology, Tibetan is commonly divided into two broad categories: the "Archaic" dialects and the "Innovative" dialects (Jaeschke 1871, cited

in Bielmeier 1988b; Róna-Tas 1966). Balti and Rebkong Amdo are both Archaic dialects, while Tokpe Gola is an Innovative dialect.

As originally defined by Róna-Tas (1966: 21), the Archaic dialects “...do not have pitch as a phonematic suprasegmental feature, and have preserved in a more or less complete form the preradical system of Old Tibetan.”² In addition to preserving these onset consonant clusters, they also preserve coda consonants and coda consonant clusters.

The Archaic dialects are found at the western and eastern margins of the Tibetan language area – in northern Pakistan and Ladakh, to the west, and in parts of Qinghai, Gansu, and Sichuan provinces of China, to the east. When the Tibetan empire was at its zenith in the 8th century, garrisons and settlements were established in these areas, and Tibetan was the language of power and prestige. As the empire collapsed and contracted late in the 9th century, remnants of the language were left stranded in these abandoned border regions, isolated from each other and from the dramatic linguistic changes which radiated through the vast geographic center. Thus archaic features of Tibetan were preserved at the periphery of the language area.

The “Innovative” dialects are spoken throughout the broad central region of the Tibetan linguistic area, and physically separate the western and eastern Archaic

² “Old Tibetan” refers to the language spoken at the time of the earliest preserved written records (Beyer 1992: 10 fn 4; 19) – i.e., the 8th century manuscripts discovered at Dun Huang and Khotan in Chinese Turkestan (Xinjiang Autonomous Region). Beyer uses the term “Classical Tibetan” to refer to the entire body of written Tibetan texts produced up through modern times (except for religious texts translated from Sanskrit) (Beyer 1992: 36-37); Goldstein distinguishes more recent newspapers and publications as “Modern Literary Tibetan” (Goldstein 1991). “Proto-Tibetan” refers to a form of the spoken language which pre-dates written records, and which is reconstructed through comparison of the modern spoken dialects and available written materials.

dialects areas. In contrast to the Archaic dialects, the Innovative dialects make use of phonemic tonal contrasts, and often have larger vowel inventories and contrastive nasalization and/or vowel length; syllables have simple onsets, and simple or no codas.

The Innovative dialects sound so radically different from the Archaic dialects that one might think, at first, that the two groups have nothing in common and are completely unrelated.³ However, by comparing the dialects to each other and to Written Tibetan, linguists have been able to establish shared correspondences and a common historical parent.

1.3 Correlations with Written Tibetan

According to tradition, as the Tibetan empire was expanding in the 7th century, King Songtsen Gampo commissioned the scholar Thonmi Sambhota to develop a writing system for the language.⁴ The alphabet he developed is based on an Indic script of that period. Tibetan orthography was then standardized in the 8th and 9th centuries, under King Ralpacan (Beyer 1992: 29). This standardization was intended to facilitate the translation of Buddhist religious texts, but one effect was that it became possible for all literate persons to communicate with one another, regardless of their dialect differences. Tibetan spelling has undergone only minor

³ The first time my Tokpe Gola Tibetan language consultant met with and listened to speakers of Rebkong Amdo Tibetan, he was completely astounded by the difference. They were able to communicate through the medium of Standard or Diasporic Tibetan.

⁴ This was an important endeavor at that time, crucial to the administration of the now-sprawling empire. Early documents include court orders and records, accounts of important events and treaties, tax records, inventories, and communiqués with remote garrisons.

changes in the 1200 years since that time. The earliest Tibetan writings – 8th century manuscripts recovered from the caves of Dun Huang in 1906, and a treaty inscribed on a pillar in Lhasa – can still be read and understood today. This is quite remarkable, when one considers how much English spelling has changed in just the past few centuries.

Written Tibetan (WT) plays an important role in linguistic investigation and analysis. Representatives of the Archaic and Innovative groups – such as the three dialects I consider in this dissertation – differ so significantly in terms of syllable structure and lexical prosody that it can be quite difficult to recognize cognates across the modern spoken varieties. Synchronic correlations can often be most easily identified by first examining diachronic correlations – that is, through the intermediary of Written Tibetan.

For all of the modern spoken dialects – regardless of whether they have ever been written or not – there is a direct correlation between pronunciation and the standardized spelling. For the Archaic dialects, the consonants and vowels of the conservative orthography correspond almost directly to consonant and vowel segments in the spoken language. Thus a word like ‘accomplished’ བསྟན་པས་ (transliterated *bsgrubs*), would be pronounced [zgrubs] in an Archaic dialect, with onset cluster and coda largely preserved.⁵

⁵ As pointed out to me by Roland Bielmeier (p.c. 2008), there are cases in my data in which a modern spoken form seems to preserve elements which are even more archaic than the form suggested by Written Tibetan. For example, in Balti the word [smor.'ðo] mu.rdo ‘border’

In the Innovative dialects, these very same orthographic consonants and vowels correspond instead to suprasegmental features such as tone register and contour, as well as to onset voicing, umlaut, and nasalization. Here, the word ‘accomplished’ *bsgrubs* would prototypically be pronounced [d̥ùp], with a low falling tone. This particular CCCC- onset cluster is pronounced as the single voiced retroflex [d]. The low tone also corresponds to this onset cluster, while the falling contour corresponds to the coda cluster.

In general, then, WT “initial” and “pre-initial” letters are preserved as syllable onsets and onset clusters in the Archaic dialects, while WT “finals” and “post-finals” are preserved as syllable codas. In the Innovative dialects, on the other hand, WT initials and pre-initials are devoiced and/or simplified, and correspond to high or low register tone; WT finals and post-finals are reduced or have disappeared altogether, corresponding to level, falling, or rising tone, as well as to vowel lengthening, nasalization, and fronting.

Further prosodic differences are illustrated by considering disyllabic words. The word for ‘star’ སྐར་མ་ *skar.ma* is pronounced [xkar.'ma] with strong second-syllable stress in Rebkong Amdo, an Archaic dialect, but [kā.mā] with a high tone on both syllables in Tokpe Gola, an Innovative dialect. The prominent prosodic feature in the Archaic dialects is stress, while in the Innovative dialects it is tone.

preserves an [sm-] onset which in WT is only *m-*. Similarly, in Rebkong Amdo [xsam.'ba] *zam.ba* ‘bridge’, we hear an initial consonant cluster where there is none in WT.

Thus a word with a particular Written Tibetan spelling may correspond to radically different pronunciations. Though the correspondences differ from one dialect to the next, Written Tibetan links them all through time and space. Beginning with the work of Jaeschke (1871, [1881] 1958), there are innumerable studies which focus on or include a comparison of features of different spoken dialects to the corresponding Written Tibetan forms. Among them are Yang (1974), Haller (1999), Makley et al. (1999) and Huber (2005).

1.4 Reconstructing Proto-Tibetan

These systematic correspondences contribute to the conclusion that all of the modern spoken Tibetan dialects derive from a common parent. By comparing the sounds of diverse varieties to each other, and to the standardized Written Tibetan forms, direct correlations can be identified even across mutually incomprehensible dialect groups. Furthermore, through such comparisons, linguists have been able to reconstruct not only individual words in Proto-Tibetan, the common source language (e.g., Bielmeier 2002), but entire evolutionary paths, such as the process by which tone arose as a contrastive feature in the Innovative dialects (e.g., Sprigg 1972; Matisoff 1973; Kjellin 1975; Mazaudon 1977; Hari 1980; Bielmeier 1988a, b; Bufan 1995).

The geographic distribution of the phonologically conservative Archaic dialects – like Balti and Rebkong Amdo – is significant to this reconstruction. As noted, these varieties are spoken at the western and eastern edges of the Tibetan

linguistic area, separated by the huge expanse of the Tibetan Plateau, where the considerably different Innovative dialects are spoken. Given the physical and linguistic separation between these two speech areas, borrowing is ruled out as a source of the similarities they share. Coincidence is ruled out as a factor as well, given the comprehensive and systematic nature of these similarities. Thus linguists have concluded that the shared features must be due to inheritance from a common source language.

(This methodology of comparing Archaic dialects from across the Tibetan linguistic area in order to reconstruct Proto-Tibetan directly informs the present work. I use parallel reasoning – comparing the features of Balti and Rebkong Amdo – to reconstruct the stress patterns and acoustic features of Proto-Tibetan.)

By thus comparing modern spoken dialects and examining Written Tibetan materials, Proto-Tibetan has been reconstructed as lacking tonal contrasts and possessing complex syllable onsets and codas (Benedict 1972; Sprigg 1972, 1980; Bielmeier 1982, 1985b, 2002; Makley et al 1999; among others). The complex clusters and the lack of phonemic tone observed in Balti and Rebkong Amdo preserve these features. This preservation is attributable to the geographic and political isolation of these areas following the fall of the Tibetan empire in the 9th century.

In the intervening central area, however, the Tibetan language underwent dramatic changes. These innovations probably originated near the seat of power and prestige around Lhasa and the nearby Yarlung valley, and diffused radially from there. In the dialects of this area, initial and final consonant clusters were drastically

reduced, or eliminated altogether. The phonologization of pitch is regarded as having occurred concomitantly: once-distinct words would have come to sound the same, had not tonal contrasts arisen to assume the functional contrastive load. This interrelationship between consonants and tone is reflected in the correlations we observe between the characters of Written Tibetan and the suprasegmental features of the Innovative dialects.

In part because tonogenesis in Tibetan has been an area of such fascination for linguists – with a focus on the relationship between consonants and tone – the role and properties of stress have been largely overlooked. This dissertation helps to fill that gap. I provide here a detailed description of stress in two Archaic dialects, and offer a hypothesis regarding the role of stress in Tibetan tonogenesis.

2. Methodology

The basic methodology of this project was to measure the acoustic properties of disyllabic words in Balti and Rebkong Amdo Tibetan, to determine whether these acoustic properties are more prominent in one syllable or the other, and to compare these patterns of prominence to patterns of perceived stress. These three steps lead to the identification of the acoustic correlates of stress in each dialect. In this chapter I describe the details of this approach. In Chapter 6, the acoustic characteristics of Balti and Rebkong Amdo will be compared, yielding a reconstruction of some of the acoustic properties of their common parent, Proto-Tibetan.

In section 2.1 below I discuss the collection of raw language materials, including a description of word lists, recording techniques, elicited data, and the transcription and transliteration conventions I have adopted. I also present background information on my language consultants. In section 2.2 I summarize the techniques used in the measurement of acoustic parameters, and the decisions made in coding various properties of the words analyzed (e.g., lexical category, morphological structure, etc.). In section 2.3 I describe, in detail, how pitch, intensity, and vowel duration data was analyzed; in section 2.4 I describe how pitch slope data was analyzed. Finally, in section 2.5 I discuss what it means for a contrast in one of the acoustic parameters to be statistically, perceptually, and contextually significant. All three of these conditions must be met in order for a correspondence with the stress pattern to be considered a non-trivial correlate of – or cue for – stress.

2.1 Recordings and other materials

The raw language materials on which my dissertation is based fall into two categories. “Primary materials” are the target words I recorded specifically for acoustic analysis. “Supporting materials” includes various other recordings and field notes.

2.1.1 Primary materials

In this section I describe the materials which were recorded specifically for acoustic and quantitative analysis. I used a high-quality head-worn microphone and focused on disyllabic words, particularly nouns and adjectives. All of my primary Rebkong Amdo materials were collected in Kathmandu, Nepal in December 2004. I recorded one speaker of Balti Tibetan in northern Pakistan in August 2003, and the other in Seattle, Washington in September 2006.

2.1.1.1 Subjects

I recorded six speakers of Rebkong Amdo Tibetan while working in Kathmandu, Nepal in December 2004. For all of the speakers I worked with, the target dialect is their mother tongue. For this study, my acoustic and quantitative analysis is based on the recordings of two speakers of the dialect. I chose the speakers whose recordings combine the best acoustic quality, the most complete word list, and the greatest speaker confidence and clarity.

The Rebkong Amdo speakers I recorded were all refugees passing briefly through the *sne len khang*, the UNHCR (United Nations High Commission for

Refugees) camp on the outskirts of Kathmandu. As I was not permitted to enter the camp, it was only possible to work with the Amdo speakers through the assistance of my Amdo research assistant, Sangye Gyatso. Sangye is himself from Rebkong County, and I had worked with him previously collecting recordings of Amdo in Rebkong village in Qinghai Province. Since Sangye was working for the UNHCR interviewing refugees, it was possible for him to visit the camp with little difficulty. He interviewed and selected speakers to work with me; they were permitted to leave the camp under Sangye's accompaniment. At the time of recording they had been in the refugee camp from one week to two months; they were gone, en route to India, within another few weeks.

All of the Rebkong Amdo speakers I worked with were natives of Rebkong County (WT: *reb.gong* or *reb.kong*; Ch: Tongren), Malho Prefecture (WT: *rma.lho*; Ch: Huangnan), Qinghai Province. Within Rebkong County they were born in several different townships and, if the same township, then in different villages. However, this geographic range does not seem to correspond to any obvious dialect variations. More important are the three speech varieties recognized in the area: *rong.skad* 'village speech' spoken in the village centers or agricultural areas; *'brog.skad* 'nomad speech' spoken among the pastoralist nomads; and *rong.ma.'brog* 'neither village nor nomad' (i.e., an indistinct mixed variety; also known as *sa.ma.'brog* in Standard Tibetan). Most of the subjects I recorded described themselves as speakers of *rong.ma.'brog* [roj.man.ɖo]. Both of the subjects whose recordings I used for acoustic analysis in this study were speakers of this variety.

My first Rebkong Amdo consultant (AR_04) was 30 years old at the time of recording. He was born in the village of *btson.mo.yar.ster* in *btson.mo* Township. He completed primary school through class five, and from the age of 12 was educated at a local monastery where he learned to read and write Tibetan. The *rong.ma. 'brog* variety of Rebkong Amdo is the only language he speaks. He had never traveled from his home until the pilgrimage which brought him through Kathmandu. He had arrived at the UNHCR refugee camp only three weeks before the time we worked together.

The second Rebkong Amdo language consultant whose recording is considered here (AR_05) was born in *ldong.nge* village in *mdo.ba* township. He was 24 years old at the time of recording. He had no formal education until the age of 13, when he began his studies at a local monastery where he learned to read and write Tibetan. He speaks a little bit of Chinese in addition to his native *rong.ma. 'brog*. Other than two years of study at Labrang monastery, he had never left his home until the pilgrimage which brought him to the UNHCR camp six weeks before we met.

My recordings of Balti Tibetan were made under different circumstances. I recorded my first language consultant (BM_01) in the village of Machulo, in Baltistan, Pakistan in August 2003. He was 50 years old at the time and had worked as a trekking guide. He had little – if any – formal education, and was not literate.

It is somewhat unusual to encounter natives of Baltistan in the United States, and I was fortunate to be able record my second Balti language consultant (BSh_03) in Seattle, WA in September 2006, when we were both attending the 39th International Conference on Sino-Tibetan Languages and Linguistics. He was born in

the Shigar area of Baltistan; his mother tongue is the variety of Balti spoken there. When he was a boy, his father found work in Islamabad, so he went there to live and attend school, returning to his native village for several months each year. He completed a university degree, and then returned to Baltistan for several years to work in the promotion of Balti language and culture. He now lives in New York city. He is fluent and literate in Urdu and English, and has begun to study the Tibetan writing system.⁶

2.1.1.2 Recording procedure

Recordings of Rebkong Amdo were made using a Sharp MD-722 minidisk recorder and a Shure SM10A monaural head-worn cardioid condenser microphone with a Shure A96F line-matching transformer. The microphone was generally maintained at a distance of ½ to 1 inch from the subject’s mouth; a windscreen was used to reduce explosive breath sounds. Though the transformer was new, some corrosion was discovered at its juncture with the cable leading to the minidisk recorder. This corrosion appears to have introduced intermittent background static into some recordings; this did not in any way affect analysis of the data.

Each word was recorded twice in isolation, and then twice within the frame “In our language we say X” (sometimes varied to “In my homeland we say X” or “My people say X”), which places the target word in utterance-medial position (e.g., literally: “In our language we X say”). In this context, the target word was often

⁶ He did not learn the Tibetan writing system during his formal education, as – until quite recently – Balti has been written using only the Urdu script, which is not ideally suited to the task.

produced with sentence-level stress. Finally, each subject spontaneously composed a very short original sentence containing the target word. This was helpful in confirming the lexical category of the target word and in providing an opportunity to check for errors – e.g., if the subject misunderstood what word he had been asked to produce, this became clear in his spontaneous composition. Also, in most of these spontaneous sentences the word was produced without emphatic stress.

The subjects varied in terms of their comfort with the task, which sometimes affected speaking rate, loudness, and clarity. This was more often the case with the Amdo speakers, since they were generally much more shy and less familiar with technology. Since I recorded several Amdo speakers on a given day, while we worked with the first subject the other(s) would observe and listen, and then were usually more relaxed and I was able to get better recordings.

I followed the same protocol when working with the second speaker of Balti (BSh_03) in Seattle, WA in September 2006, except that I used an Edirol R-09 digital recorder rather than a minidisk recorder.

When I worked with the first speaker of Balti (BM_01) in Machulo, Pakistan, I recorded an elicitation session conducted by my colleague Nicolas Tournadre, whose research objectives were different than mine. Target words were produced only in isolation. Some words were produced two or more times; some were produced only once. I have no recordings of the target words in a fixed frame, and only a few in natural, spontaneously-produced sentences. In recording this material I again used a Sharp MD-722 minidisk recorder, but with an Audio Technica monaural head-worn

cardioid condenser microphone. The microphone was generally maintained at a distance of 1 to 2 inches from the speaker's mouth, though it had a tendency to slip around on the speaker's head.

All of these recordings (except for the one made in Seattle) were conducted with the assistance of other native speakers of Balti or Rebkong Amdo, who also speak some English as well as Urdu, Tibetan, and/or Nepali. Subjects were prompted with a word, definition, or phrase in one of these languages (or in the target dialect, when necessary), and the subject then produced the target form in his dialect.

The recordings made using the Sharp MD-722 mini-disk recorder were transferred to CD's by connecting a Tascam MD-350 Professional Minidisk Deck to a Tascam CD-RW 700 CD Rewritable Recorder, thus preserving the quality of the digital recording. The material I collected using the Edirol R-09 digital recorder was transferred directly in digital format to a personal computer through a USB cable. The digital material was then available for acoustic analysis.

2.1.1.3 Word lists

A list of the disyllabic words analyzed for this study can be found in the Appendix. The Written Tibetan form is provided for each word, and the IPA transcription is provided for the forms as produced by each of the four speakers, in isolation and frame settings.

As noted above, when I recorded my first Balti consultant in northern Pakistan (BM_01), I piggy-backed on an elicitation session conducted by Nicolas Tournadre. Thus the words I recorded from this speaker were all from the pan-dialectal 'Basic

Lexicon' he had developed for his ongoing documentation and comparison of language varieties across the entire Tibetan linguistic area. This list (Tournadre p.c. 2003) consists of more than 650 words representing all lexical categories, including many disyllabic nouns, adjectives, and verbs which are of direct relevance to my work.

When I worked with my Rebkong Amdo consultants (AR_04 and AR_05), I used my own list of approximately 500 lexical items, which was dominated by nouns and adjectives (and which included many words which were not disyllabic). Since I plan to use this same word list in future research projects, I took into consideration several factors which are of special relevance to tonal varieties of Tibetan but which were not so important for the present study. In the tonal dialects – even those which have never been written – many scholars have noted that there is a systematic correlation between the letters of Written Tibetan and the perceived tone patterns. Thus in order to achieve a balanced sample, as much as possible I selected disyllabic words to represent each letter of the Tibetan alphabet as an onset of σ_1 and of σ_2 , both as an isolated root and as part of consonant clusters. Syllable rhymes also represented different Written Tibetan spelling patterns: open syllable; closed syllable with a simple suffix; closed syllable with suffix + post-suffix. Controlling for these parameters will permit consideration of correlations between onsets and register tone, and between codas and contour tone in future studies. Of more immediate relevance, it also permits consideration of the potential relevance of onsets to syllable

prominence, and the potential contribution of codas to syllable weight, both of which can play a role in stress assignment.

Disyllabic nouns and adjectives in the list were also chosen to represent various morphological structures (i.e., monomorphemic, compound, and reduplicated forms), as well as the different tone patterns which are common in the tonal dialects (i.e., HH and LH, for nouns and adjectives).⁷

The list also includes many monosyllabic words – in particular, pairs of monosyllabic words that can be combined to form compounds: e.g., *rta* ‘horse’ + *sga* ‘saddle’ → *rta.sga* ‘horse saddle’. Disyllabic compounds whose elements could be reversed – such as *mig.chu* (‘eye’ + ‘water’) ‘tear’ vs. *chu.mig* (‘water’ + ‘eye’) ‘spring’, and *bcu.gsum* (‘ten’ + ‘three’) ‘thirteen’ vs. *gsum.bcu* (‘three’ + ‘ten’) ‘thirty’. Such cases provide a means of testing the correspondence between syllable position and stress.

A number of tri- and quadri-syllabic words were added to collect preliminary material for future work, and words from other lexical categories or morphological composition were included to illustrate contrasting patterns of stress or tone. For instance, noun + clitic constructions exhibit $\sigma 1$ stress (in non-tonal dialects) or a HL tone pattern (in tonal dialects). Such forms serve as “the exceptions which prove the rule”, since disyllabic nouns otherwise do not exhibit either $\sigma 1$ stress or low tone on $\sigma 2$. Thus in addition to the nouns and adjectives which were my focus when I

⁷ Of course, tone is not relevant to Balti or Rebkong Amdo, but including such words will facilitate comparative work in the future.

developed the list, also included are disyllabic numerals, verbs, and adverbs, as well as disyllabic noun + clitic and noun + verb constructions.

I sought to populate the list with lexical items common throughout the vast Tibetan linguistic area, despite dialectal, environmental, cultural, and religious differences. This allowed me to elicit many of the same words from both dialects, and will enable me to expand the project to other dialects in future. I began the list in collaboration with my Balti language consultant in Skardu, Baltistan, Muhammad Raza Tasawor. From the glossary of an intermediate level Chinese-[Lhasa] Tibetan-English pedagogical text which I had with me in Pakistan (Thup bstan tbang po et al., 2002), we selected appropriate words which he identified as also part of the Balti lexicon. I also drew on word lists which I had developed during my previous work on Tokpe Gola Tibetan – a tonal dialect of northeastern Nepal which I have studied – and included new ideas offered by Dandu Sherpa, my primary Tokpe Gola consultant.

I also borrowed from several very helpful sources, selecting words from the ‘Noun’ volume of Roland Bielmeier’s *Comparative Dictionary of Tibetan Dialects* (Bielmeier 2002), from the ‘Basic Lexicon’ developed by Nicolas Tournadre (p.c. Tournadre 2003), and from Roerich’s grammar and lexicon of Rebkong Amdo (Roerich 1958). Finally, I included words which were used in the works of Dawson (1980), Bielmeier (1988a), and Meredith (1990), in order to check the stress patterns they reported for various dialects of Tibetan, and in the work of Volkart (2003) to illustrate the various types of compounds which occur in Tibetan.

I elicited from this same list of about 500 lexical items when I recorded the second speaker of Balti (BSh_03) in Seattle, WA, though some of the monosyllabic words were skipped and other disyllabic words – especially verbs – were added.

Not all of the words in these lists were recorded: In some cases I was not able to effectively describe or define the target word for the consultant; in other cases a consultant did not have a form of the target in his idiolect. During several sessions, in the interest of time, attention was focused on disyllabic forms which had already been elicited successfully from other speakers of that dialect. And not all of the words recorded were considered further: Words consisting of more or less than two syllables (i.e., mono-, tri-, and quadrisyllabic forms) were set aside for future research projects.

Most significantly, not all of the disyllabic words recorded could be analyzed: There were many cases in which the phonetic composition of a word did not lend itself to accurate identification of segmental boundaries. Especially problematic, for instance, were words with (a) syllables containing glides, such as [Cyi-] (occurring in both dialects) or [-wa] (occurring in both dialects, but especially common in Amdo, as in [moq.'wa] *mag.pa* ‘bridegroom, son-in-law’); (b) syllables in which a velar nasal was reduced to a velar approximant: [ŋ] → [ɰ]. This occurred frequently in Balti, in words such as [k^hãɰ.'ma] *kang.pa* ‘house’ and [ʔɛdũɰ.'ma] *gdung.ma* ‘beam’, and in Amdo in words such as [ʔŋa.'mõɰ] *rnga.mong* ‘camel’; (c) syllables in which a vowel was highly fricated before a fricative coda, as in σ2 of Amdo [ʔrɣ.'dɣɣ] *ri.dwags* ‘wild animal; herbivore’; and (d) syllables with syllabic sonorant nuclei. Since the

span of the nuclear vowel could not be defined for such syllables, there was no way to collect precise measurements of the vowel's duration, average pitch, etc. For each dialect, there were scores of disyllabic words I recorded which could not be analyzed because they could not be reliably segmented. These cases will provide useful guidelines in developing word lists for future projects.

Of course, words which were mumbled or partially obscured by background noise were also set aside.

Thus the list of disyllabic words which were analyzed, provided in the Appendix, includes all words from both target lists which could possibly be recorded and segmented. With all speakers, it invariably happened that numerous words similar to or contrasting with those on the target lists arose during our discussions; these, of course, are included as well.

2.1.2 Supporting materials

“Supporting materials” refers to everything other than the recordings collected specifically for acoustic and quantitative analysis of the target words. This includes miscellaneous recordings and field notes gathered in different places and under varying conditions. These materials were useful in developing the phonetic and phonological descriptions of the two dialects which are presented in Chapters 4 and 5. The stress patterns they exhibit are consistent with those quantified through acoustic measurements.

For Balti Tibetan, these supporting materials consist of field notes with observations and close phonetic transcriptions of my 500-item word list, as elicited

from a third male speaker of the Skardu variety (speaker BSk_02). After eliciting and transcribing the word list, I subsequently recorded all of the words, but – as discussed in Chapter 5 – I was concerned that the speaker had suddenly adopted an unnatural intonation under the more formal conditions of recording, and so was reluctant to consider this data.

I collected preliminary material on Rebkong Amdo while conducting fieldwork in Qinghai Province, China in September 2003. I elicited my 500-item word list from a male speaker who was living in the city of Xining. This session was closely transcribed, and was also recorded. I later worked with one male speaker and one female speaker in the village of Rebkong, eliciting a set of words provided to me by Jackson T.S. Sun (p.c. 2003). This set was designed for pan-dialectal use in identifying diachronic correspondences between WT spelling (i.e., various onsets and codas) and features such as [de-]voicing, the phonemic status of tone, the preservation of codas, the development of labial initials, etc. Again, this elicitation session was recorded and closely transcribed.

In all of this supporting material, stress patterns were the same as those quantified here.

2.1.3 Transcription and transliteration

I transcribed the data from the two dialects according to the 2005 IPA standards, with a few exceptions. For the palatal approximant I use [y] rather than [j]. This is consistent with the use of *y* in the Wylie transliteration of Written Tibetan

(Wylie 1957), and thus eliminates the confusion which might arise if I were to use [j] in transcription and *y* in transliteration to represent the same sound. It also maintains a clear distinction from the Wylie voiced palatal affricate *j* ([dz̥] or [dʒ̥]) and from the IPA voiced palatal stop [j], both of which are also used in this project. For the palatal nasal, I use [ny] rather than the IPA [ɲ], a choice which again is consistent with the Wylie transliteration.

My transliteration of Written Tibetan into the Latin alphabet generally follows the Extended Wylie Transliteration Scheme developed by the University of Virginia's Tibetan and Himalayan Digital Library (THDL)⁸, based on Wylie (1957), except that I prefer to use a period "." rather than a blank space to represent boundaries between syllables. This leaves the blank space available for use as a word boundary, which is not indicated at all in the Tibetan orthography.

2.1.4 Stress placement

I make frequent reference in this dissertation to my own perceptions of stress. Stress was very easy to hear, and the patterns were the same in both dialects: nouns, adjectives, and numerals (the non-verbs) are stressed on σ_2 , while verbs are stressed on σ_1 . These patterns have been observed and reported by previous scholars, looking at Balti and other varieties of Amdo separately. What is new in my work is that I connect the dots: by thinking of stress in the two dialects together, rather than

⁸ <http://iris.lib.virginia.edu/tibet/>

separately, it is quite clear that the patterns are the same, and that the most plausible reason for this is that they were inherited from a common source.

I also attempted to find out if my language consultants were consciously aware of stress. During my work with Rebkong Amdo speakers in Kathmandu, I tried “clapping” and “tapping” tests on several occasions. I asked the speakers to clap their hands or tap the table in synchrony with production of a word, each syllable accorded one clap or tap, with a louder, more emphatic impact corresponding to the stressed syllable. These experiments did not yield the hoped-for results. It seems I did not succeed in explaining the notion of stress or prominence, and my consultants instead understood this to be an exercise in counting the syllables of a word – with one clap per syllable – which they were able to do quite readily.

I did not have the opportunity to conduct such an experiment when recording my first Balti consultant, BM_01, in Pakistan. Speaker BSh_03 seemed to be consciously aware of stress contrasts, since he was able to offer me noun-verb minimal pairs, such as [xlat.'pa] *klad.pa* /*glad.pa* ‘brain’ and ['xlat.pa] **glad.pa*⁹ ‘to be tired’.

2.2 Data coding and data collection

Data coding and data collection were completed in three stages, which are discussed in the three sections below. Non-variable information associated with a word – such as its lexical category and morphological structure – was coded in the

⁹ “*” indicates a reconstructed Proto-Tibetan form, provided to me by Roland Bielmeier (p.c. 2008).

name of the soundfile, as described in section 2.2.1. In 2.2.2 I discuss the use of Praat “textgrids” to encode the potentially variable information specific to each utterance – segments, perceived stress, syllable structure, etc. In this section I also describe the techniques used to collect measurements of the four acoustic parameters which are likely correlates of stress: pitch, intensity, pitch slope, and vowel duration. Finally, in section 2.2.3 I discuss the coding of categorical information such as syllable type (open *vs.* closed) or vowel height (low *vs.* mid *vs.* high). This stage of coding was completed in an Excel spreadsheet or .txt file, and was based on the token-specific information embedded in the textgrids.

2.2.1 Coding in the filename

Each word recorded and analyzed was assigned a unique filename, summarizing and storing the kind of information that does not vary from one token to the next. This information could later be “unpacked” using a simple equation in an Excel file.

Typical filenames are illustrated in (1):

(2.1) AR__01_04_01_02_0029_star.wav
AR__01_04_01_02_0029_star.TextGrid

They have the following structure:

(2.2) dialect_project_speaker_lexical.category_morphology_item#_gloss

Three character spaces were allotted for the dialect codes, allowing for future expansion of my study. The codes I used in this project were as follows:

- (2.3) BM_ Balti – as spoken in the village of Machulo
 BSh Balti – as spoken in the Shigar area
 AR_ Amdo – Rebkong variety

The project code is a key referencing which word list I recorded and what kind of recording equipment was used. The speaker code is self-explanatory. I allotted two character spaces for each of these codes. All of the filenames in this project thus begin with one of the following fixed prefixes:

- (2.4) BM__02_01_
 BSh__03_03_
 AR__01_04_
 AR__01_05_

The remaining elements of the filename encode other non-variable information associated with the target word. In this study I consider nouns, adjectives, numerals, and verbs, but words from other lexical categories were recorded as well. Lexical category codes are shown in Table 2.1 below.

Table 2.1 Lexical category codes

Code	Lexical category
01	noun
02	adjective
03	numeral
04	adverb
05	verb
06	clause
07	question word
08	verb – imperative
09	complex nominal or noun phrase
10	pronoun (personal, deictic)
11	auxiliary / copula / evidential
12	kinship term
13	verb phrase

Morphological structure codes are shown in Table 2.2. This list includes all of the forms I encountered during recording. Again, only a few are relevant to this analysis; for nouns, adjectives, and numerals: (02) monomorphemic, (03) compound, (04) reduplicated, (05) borrowed, (00) undetermined; for verbs: (12) N+Vblzr, (16) verb - citation form.

Table 2.2 Morphological structure codes

Code	Morphological structure / origin
00	unspecified / undetermined / irrelevant
01	monosyllabic
02	monomorphemic
03	compound
04	reduplicated (repetitive, distributive)
05	borrowed
06	negated
07	N + adjective
08	nominalized
09	linked by genitive
10	cliticized form
11	interrogative
12	N + verb / verbalizer
13	construction
14	verb – past form
15	verb – non-past form
16	verb – citation form

The final two elements of the filename are the item number and gloss of the word recorded.

Thus AR__01_04_01_02_0029_star refers to item 0029 ‘star’, a monomorphemic noun recorded from Rebkong Amdo speaker 04, drawn from a particular word list and recorded using a particular combination of recording equipment and microphone.

The filename information was inextricably linked to acoustic measurements and other coding information in a .txt file, as discussed in the sections below.

2.2.2 Acoustic measurements and coding in textgrids

In this study, four potential acoustic correlates of stress were evaluated: “pitch”, “pitch slope”, intensity, and vowel duration. This set of parameters differs from conventional acoustic studies in two respects.

First, it is not common to identify “pitch slope” as an indicator of stress. In this project, though, on hearing Rebkong Amdo nouns produced in isolation, one is immediately and unmistakably struck by the dramatic drop in pitch which occurs in the second (stressed) syllable. Quantifying this drop in pitch is thus essential to an evaluation of stress in Rebkong Amdo.

Second, technically speaking, “pitch” and “pitch slope” are not “acoustic correlates”. Pitch is actually an auditory or perceptual parameter. The corresponding measurable acoustic parameter is fundamental frequency, which reflects the rate of vibration of the vocal folds of the speaker, and the rate of vibration of the eardrum of the hearer. Here, though, I regard pitch and pitch slope as two expressions of one acoustic resource: fundamental frequency. “Pitch” is the average F0 over an interval, while “pitch slope” is the change in F0 over an interval. As demonstrated in Chapters 4 and 5, pitch and pitch slope sometimes reinforce one another as correlates of stress, and sometimes complement one another as correlates of stress.

Measurements of these four parameters were collected using the Praat phonetics software (Boersma 2001, Boersma and Weenink 2008).¹⁰ The measurement process was expedited and semi-automated using a custom script which I wrote within Praat. For vowel duration, pitch, and intensity, measurement was fully automated, so it was just as easy to collect several data points for these parameters as to collect one. For pitch slope, the script automatically opened the textgrid, allowing the user to inspect the pitch contour. The mid 50% of the vowel was highlighted automatically, and the script would calculate the slope across this span by default. Alternatively, the user could manually select a span with a representative pitch slope, avoiding discontinuities in the pitch contour or anomalous adjacency effects of onset and coda consonants. Slope was calculated by subtracting the pitch at the start of the interval from the pitch at the end of the interval, and then dividing by the elapsed time. The resultant value, in Hz/msec, was then divided by 10 to yield a result in Hz/100msec.

Table 2.3 below summarizes the ten parameters measured – and the span(s) over which each parameter was measured – for each syllable of each target word. The pitch and intensity measurements collected over the mid 50% of the vowel – which I regard as “stable” since they are free of the influence of adjoining consonants – are the ones I used in statistical analysis and on which the conclusions of this study are

¹⁰ This program is updated frequently. The versions I used for this project were 4.4.22 through 4.5.11, accessed from <http://www.fon.hum.uva.nl/praat/> between May 2006 and January 2008. According to http://www.fon.hum.uva.nl/praat/manual/What_s_new_.html, none of the modifications implemented in recent years have affected the algorithms used to measure the acoustic parameters of interest.

based. (In general, there was very little difference – only a few Hz, or a dB or two – between these stable values and the values measured over the full span of the vowel.)

Table 2.3 Acoustic measurements collected using Praat ¹¹

Acoustic feature	Measurement span			
	Average over full nuclear vowel	Average over mid 50% of nuclear vowel	Value at midpoint of nuclear vowel	Maximum value
Vowel duration (msec)	✓	NA	NA	NA
Pitch (Hz)	✓	✓	✓	✓
Intensity (dB)	✓	✓	✓	✓
Pitch slope (Hz/100 msec)		✓	NA	NA

NA: not applicable

The setting values I selected in Praat for each acoustic parameter are summarized in Table 2.4 below.

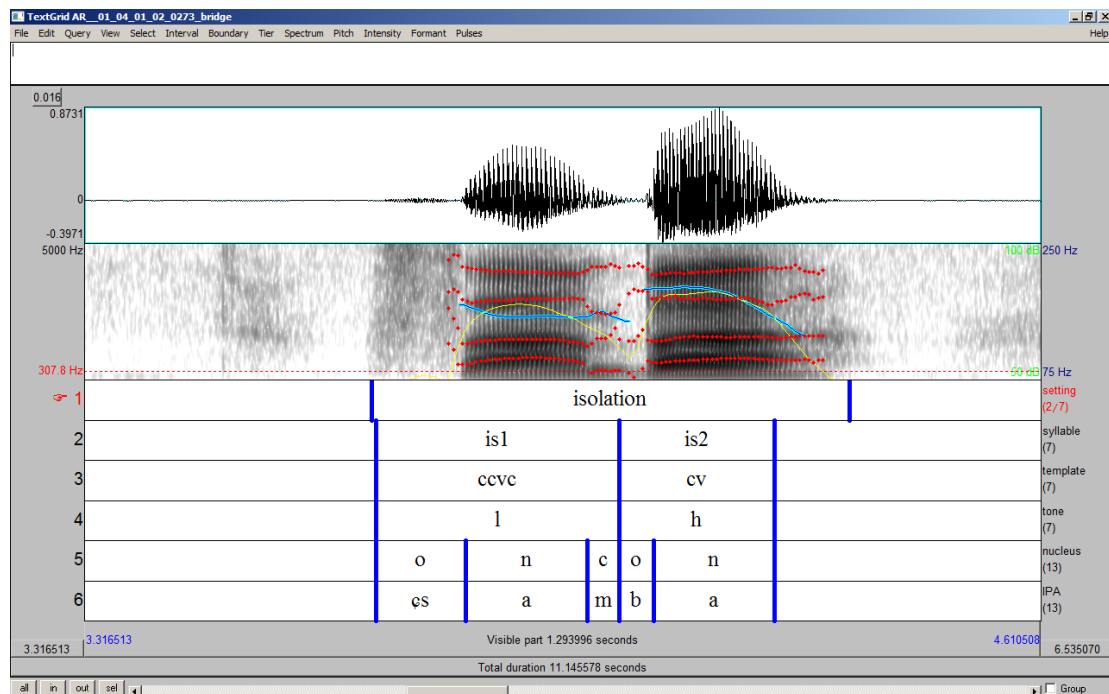
Table 2.4 Setting values used in the Praat script

Acoustic feature	Setting	Value
Pitch	Time step (sec)	0.0 (= auto)
	Pitch floor (Hz)	75
	Pitch ceiling (Hz)	350
Intensity	Minimum pitch (Hz)	75
	Time step (sec)	0.0 (= auto)
	Averaging method	energy
	Subtract mean	yes
Formants (burg)	Max number of formants	5
	Maximum formant (Hz)	5000

¹¹ In addition to the ten measurements shown in the table, the script also measured pitch and intensity at four other discrete points in time: the beginning of the vowel; the beginning of the stable 50% of the vowel; the end of the stable 50% of the vowel; and the end of the vowel. The absolute time at each of these points was recorded, as was the absolute time at which maximum pitch and intensity occurred.

The script requires as input a sound file (.wav or other format) and a corresponding “textgrid”, an artifact which is created in Praat. Textgrids provide a means of permanently storing the segmental content, syllable structure, and other information which is unique to each particular token of the target word. The structure and labeling of an input textgrid required for operation of this script is illustrated by the example in Figure 2.1 below.

Figure 2.1 Sample Praat textgrid: Rebkong Amdo word [çsam.'ba] *zamba* ‘bridge’



The measurement script operates by scanning Tier 5 of a textgrid for any intervals labeled ‘n’; these intervals correspond to the vowel nucleus of a syllable. For each ‘n’ interval, the ten acoustic parameters listed in Table 2.3 are measured, temporal information is recorded, and then labeling information is collected from the

tiers above and below the ‘n’, and from onset (‘o’) and coda (‘c’) intervals to the left and right, when an onset or coda exists.

Tier 1 provides information on the setting of the token being analyzed. ‘isolation’ indicates a pausal form; ‘frame’ indicates a form produced in the frame such as “In our language we X say”; ‘sentence’ indicates a form occurring within the speaker’s spontaneously-composed short sentence. The labels on Tier 2 encode the syllable being analyzed: ‘is1’ stands for “isolation: syllable 1”; ‘fs2’ stands for “frame: syllable 2”, and so forth. The syllable template is given on Tier 3.

Tier 4 is labeled to indicate syllable tone: ‘l’ for low tone, ‘h’ for high tone. This parameter is not relevant to the present study; Balti Tibetan and Rebkong Amdo Tibetan are not tonal dialects. The purpose of this coding scheme is to facilitate a future investigation of the correspondence between consonants and tone across dialects.¹²

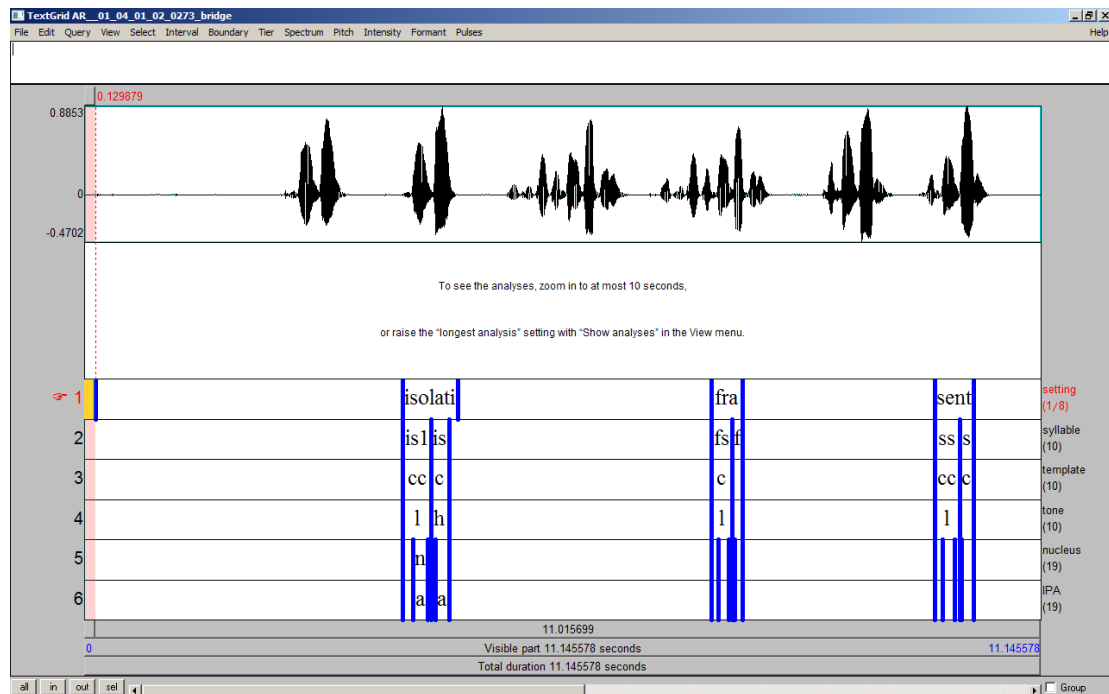
¹² While Balti Tibetan and Rebkong Amdo Tibetan are non-tonal, the labels on this tier reflect the tone that syllable *would* have in a dialect such as Tokpe Gola Tibetan, a tonal dialect spoken in northeastern Nepal on which I have conducted extensive field work. The tone patterns of Tokpe Gola are similar to those reported for other Innovative dialects such as Lhasa Tibetan, Kyirong Tibetan, and others. Storing the “hypothetical tone” values for Balti and Rebkong Amdo together with other syllable information will make it easier, in future work, to identify correlations between, for instance, consonant clusters and tone patterns. These hypothetical tones were generally determined by comparing Balti and Rebkong Amdo words to the Tokpe Gola cognate. For instance, the words for ‘willow tree’ in Balti and Rebkong Amdo are [t̪a.χ.ma] and [xt̪aŋ.ma], respectively, with second-syllable stress. They are coded on Tier 4 with ‘h’ on both syllables, following the high-high tone pattern in the Tokpe Gola word [t̪a.ŋ.ma]. Where there was no clear Tokpe Gola cognate, the hypothetical tone was determined by finding the word in a dictionary such as Goldstein (1984), Norberg-Hodge and Paldan (1991), Goldstein (2001), or Sprigg (2002) and deducing the tone pattern from the Written Tibetan spelling. For nouns, adjectives, and numerals, of course, it was only necessary to apply this technique to the first syllable, since the tone of the second syllable is always high in Tokpe Gola and similar dialects. For verbs, whose tone patterns are more variable, I generally did not assign a hypothetical tone to the second syllable since I did not know what it might be.

Finally, the IPA symbols for the nuclear vowel and for onset and coda consonants (when present) are provided on Tier 6.

Once all of the acoustic measurements and labeling information have been gathered for a particular 'n', the script automatically writes all of the data collected – together with the name of the sound file – to its own line in a specified .txt file, which can be readily opened in Microsoft Excel. The script then moves on to search for the next 'n' in the textgrid, and repeats the process.

Figure 2.1 shows only a portion of a full textgrid. The full textgrid from which this selection was extracted is shown in Figure 2.2. In addition to the isolation form, the frame and sentence forms have also been segmented and coded. (Since I did not collect measurements from the sentence forms, I usually did not bother to label them on the textgrid.) In most cases, the speaker produced the isolation and frame forms twice; only the second iteration was marked for analysis. When the script is run, it collects measurements and labels from all forms marked in the textgrid, as long as their vowel nuclei are coded 'n'. In some cases, though, when the speaker produced the frame form in a rush, the 'n' interval was too short for Praat to define the necessary analysis window. Thus there are more measurements for isolation forms than for frame forms.

Figure 2.2 Full Praat textgrid: Rebkong Amdo word [çsam.'ba] *zamba* ‘bridge’



The script was designed to operate on batches of sound file / textgrid pairs. If a number of such pairs are present in Praat’s ‘Object window’, when the script collects data for the last ‘n’ interval of a particular word, it will immediately move on to the next word in the list, and will continue collecting data and sending it to the specified .txt file until the entire list has been processed. Another script I wrote instantly opens batches of sound file / textgrid pairs from a file folder to the Object window, so a large number of words can be processed very quickly.

With the measurement process thus automated, the only time-consuming task was the creation, segmentation, and labeling of the textgrids. I was able to semi-automate this step through the use of an additional Praat script, and used another script to expedite the process of double-checking textgrids by automatically opening

them and zooming to the portions of interest. Using scripts to minimize the time devoted to all of these mechanical steps meant that attention could be focused instead on areas requiring a human eye and interpretation.

As noted above, all of the acoustic measurements listed in Table 2.1 are collected within the span defined by the ‘n’ intervals on Tier 5 of the textgrid. Thus the collection of meaningful data is entirely dependent on accurately identifying the boundaries of the nuclear vowel. Segmentation was based on waveform, spectrogram, and formant displays. As a matter of standardization, boundaries were drawn at the nadir of the trough of a completed waveform, closest to the point where waveform, formant, and spectrogram information more or less coincided. The beginning and end of the F2 formant were considered in defining vowel boundaries.

In the case of words ending in an open syllable, these cues do not provide a consistent means of defining the end of the final vowel. In this setting, vowels display considerable variation in the fading out of waveforms, voice bars, and formant structure; some vowels exhibit strong or faint final aspiration. I thus incorporated into my Praat script a heuristic suggested to me by Ian Maddieson (p.c. 2004): the terminus of a vowel in an open final syllable is identified as the point at which intensity declines to a level 10 dB less than the vowel’s peak intensity (about $\frac{7}{8}$ of the peak intensity). In practice, this solution turned out to select an endpoint which was generally consistent with the signature of the waveform, spectrogram, formant, and/or pitch and intensity traces. When tested on words with a closed final syllable, it

often selected a vowel endpoint that coincided with the actual nucleus - coda boundary.

2.2.3 Coding in the data file

All of the information described above – the non-variable information coded in the filename, the utterance-specific information coded in the textgrid, and the acoustic measurements – was sent by Praat to a separate .txt / Excel file for each speaker of each dialect. The final stage of data preparation was completed within these files, and consisted of coding the measurements and observations for categorical factors relevant to the analysis of stress and tone.

The parameters for which syllables were coded are shown in Table 2.5 below. In this table, factor (a) provides a means of testing and controlling for the inherent variation of fundamental frequency and intensity as a function of vowel height (Lehiste 1970). Factors (b) and (c) are relevant to the analysis of vowel duration: clearly, it is essential to distinguish vowels in which compensatory lengthening has occurred from vowels which are relatively long as an expression of stress. Syllable closure type is also relevant to the analysis of vowel duration, since vowels are longest in open syllables, long in syllables closed with a voiced coda, and shortest in syllables closed with a voiceless coda (Peterson and Lehiste 1960; Lehiste 1970).

Table 2.5 Syllable properties coded

	Factor	Levels	Basis for coding
a.	vowel height	low, high, mid	IPA transcription in the textgrid
b.	vowel length	short, long, diphthong	long vowels are cases of compensatory lengthening and nasalization
c.	closure type	open syllable closed with a voiced coda closed with a voiceless coda	IPA transcription in the textgrid

Table 2.6 shows the factors which were coded at the word level, illustrating contrasts across the two syllables. Factors (a) and (b) are relevant to the analysis of fundamental frequency and intensity, while (c) is relevant to the analysis of vowel duration.

Table 2.6 Word properties coded

	Factor	Levels	Basis for coding
a.	vowel quality contrast	same.vowels different.vowels	IPA transcription in the textgrid
b.	vowel height contrast	lower.higher higher.lower same.height	vowel height coded for each syllable
c.	closure type contrast	open.open closed.closed open.closed closed.open	syllable closure type coded for each syllable

With coding completed, the .txt files were loaded into the “R” (2007) application, for statistical analysis of correlations with stress.

2.3 Analyzing pitch, intensity, and vowel duration data

Since all of the words considered in this study are disyllabic, pitch, intensity, and vowel duration can only be either greater on σ_1 , greater on σ_2 , or about the same on both syllables. This σ_1 - σ_2 syllable comparison is the foundation of all elements of the analysis – mathematical, graphical, and statistical. In this section I describe the techniques and methods I used in analyzing the acoustic data collected and in identifying correlations with the non-variable, variable, and categorical parameters described in section 2.2.

In section 2.3.1 below, I present the conventions I adopted in an effort to make calculations and figures easy to interpret. In section 2.3.2 I provide detailed examples of the analysis of these acoustic parameters.

2.3.1 Iconic representation in calculations and graphs

Since both English and Tibetan are written from left to right, it is not unnatural to think of σ_1 as occurring to the left, and σ_2 as occurring to the right. As much as possible, I arranged calculations and graphs according to this σ_1 - σ_2 / left-right metaphor, with the objective of facilitating data interpretation.

2.3.1.1 Calculating differences across syllables

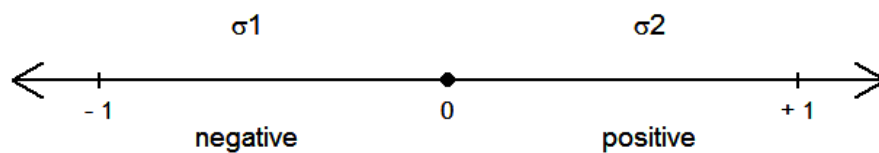
Pitch, intensity, and vowel duration are analyzed by comparing measurements across the two syllables of a word through simple subtraction. I always subtract the value measured in σ_1 from the value measured in σ_2 . For instance, in assessing pitch difference across syllables I use the equation $\text{pitch}_{\sigma_2} - \text{pitch}_{\sigma_1}$, rather than vice versa.

Thus in the Balti verb [‘min.ma] ‘to give’, where the pitch on σ_1 is 144 Hz and the pitch on σ_2 is 87 Hz, I calculate the pitch difference across syllables *not* as $(144 - 87) = 57$ Hz, but as $(87-144) = -57$ Hz. The resultant difference is thus a negative number.

There are [at least] two ways in which one can think about this negative number. Most obviously, the calculated difference represents the change in pitch across the word: pitch *decreases* by 57 Hz. (A positive number would thus indicate an *increase* in pitch).

Perhaps more useful is an analogy to the conventional orientation of a number line. As illustrated in Figure 2.3, negative numbers are on the left side of the number line, and σ_1 is on the left side of a word; positive numbers are on the right side of the number line, and σ_2 is on the right side of a word.

Figure 2.3 Syllable position and the number line



Thus it is easy to remember that a negative pitch difference means that pitch is higher on the σ_1 , and a positive pitch difference means that pitch is higher on σ_2 .

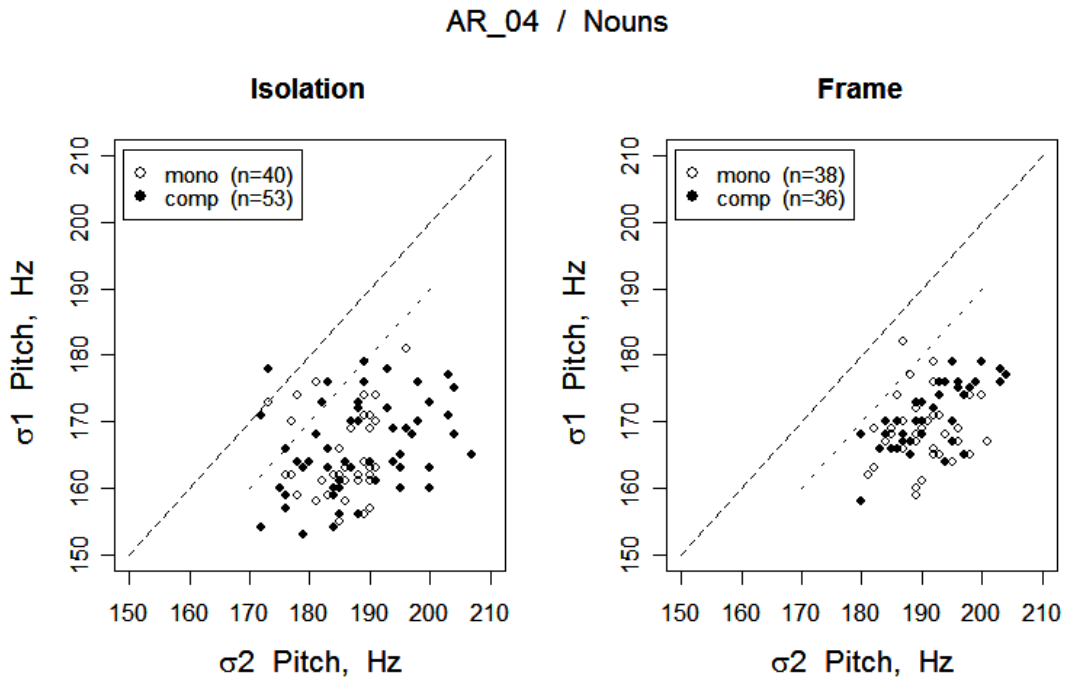
This same principle pertains to the analyses of intensity data. For vowel duration, a

negative difference indicates a longer vowel in σ_1 , and a positive difference indicates a longer vowel in σ_2 .

2.3.1.2 Graphing a single parameter

Graphs are ubiquitous in Chapters 4 and 5. One type of graph which I use frequently provides a direct comparison of the acoustic measurements collected from the two syllables of a subset of words. For example, Figure 2.4 below shows pitch measurements for nouns produced by Amdo speaker AR_04. (The nouns are plotted in terms of their morphological structure – monomorphemic *vs.* compound – which turns out not to be relevant to the distribution.)

Figure 2.4 Graphing a single parameter: Pitch



In these graphs, the x-axis shows the value for σ_2 , and the y-axis shows the value for σ_1 . The dashed line represents all theoretical cases in which the pitch measured in σ_1 is equal to the pitch measured in σ_2 . The few points which plot to the left of the dashed line have a higher pitch on σ_1 . Again, this should be easy to interpret, because in our writing system σ_1 is to the left. All points which plot to the right of the dashed line have a higher pitch on σ_2 .

What we see in the graphs above is that – with only a handful of exceptions – pitch is always higher on σ_2 of AR_04 nouns. This acoustic pattern is consistent with my perception of σ_2 stress in Rebkong Amdo nouns. In fact, in most words the difference in pitch across syllables is >10 Hz, as indicated by the position of the points with respect to the dotted “Pitch difference = 10 Hz” reference line. Given the consistency of the pattern and the magnitude of the difference across syllables, I consider pitch to be a “robust” correlate of σ_2 stress.

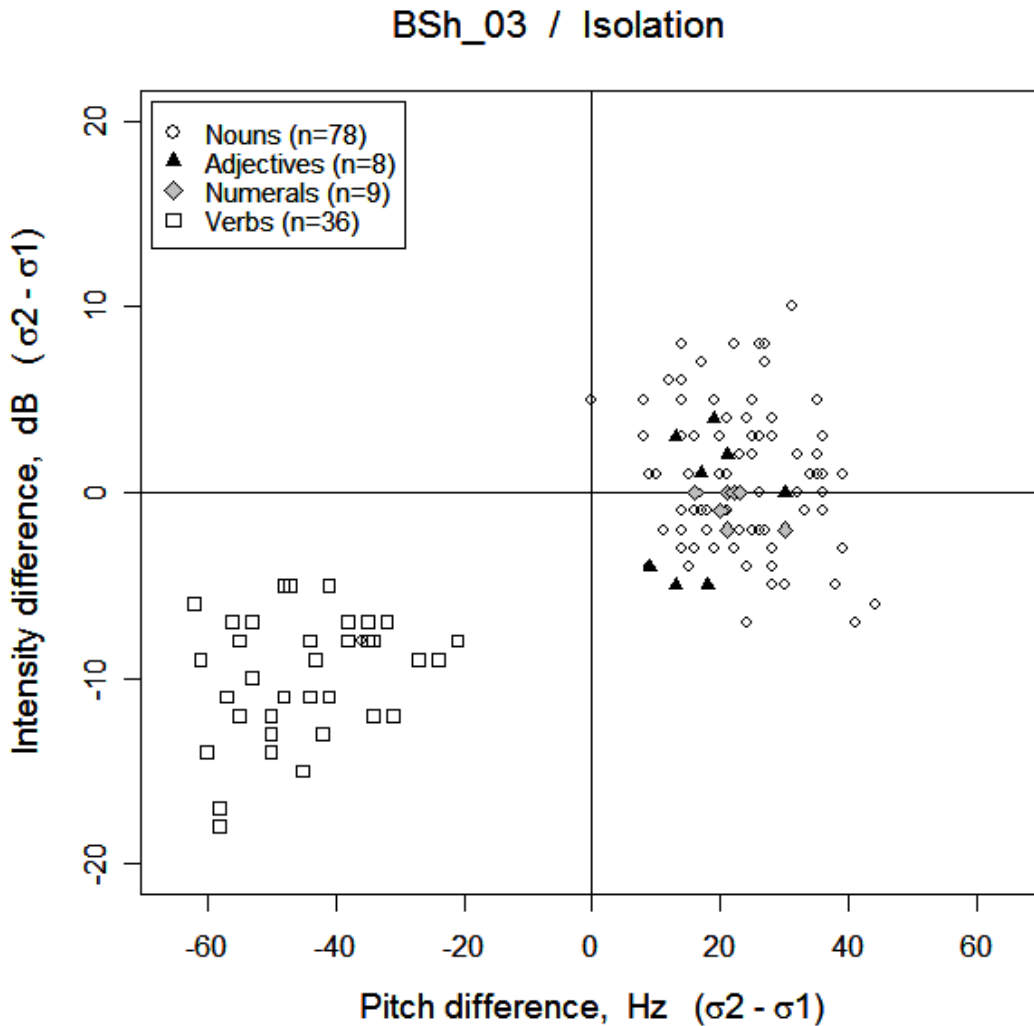
Note that I use the same scale for both x and y axes for isolation and frame forms, so the distribution in the two settings can be readily compared.

2.3.1.3 Graphing two parameters

Other graphs were designed to illustrate the relationship between two acoustic parameters. In Figure 2.5 below, for example, pitch and intensity are compared for all of the words analyzed for Balti speaker BSh_03. The x-axis shows the difference in pitch across the two syllables of each word. As described in 2.3.1.1, the equation used for the comparison was $\text{pitch}_{\sigma_2} - \text{pitch}_{\sigma_1}$, and not vice versa. All points which fall to the left of the vertical “Pitch difference = 0” line – those with a negative pitch

difference – represent words which have a higher pitch on σ_1 ; all words which plot to the right of the vertical line – with a positive pitch difference – have a higher pitch on σ_2 . It is easy to see that all the verbs have a higher pitch on σ_1 , and all the non-verbs (nouns, adjectives, and numerals) have a higher pitch on σ_2 . For both verbs and non-verbs, pitch is higher on the stressed syllable, corresponding with the perceived locus of stress. The correlation is robust, in both cases.

Figure 2.5 Graphing two parameters: Pitch difference vs. intensity difference



The y-axis in this graph shows the difference in intensity across syllables: $\text{intensity}_{\sigma_2} - \text{intensity}_{\sigma_1}$. All words which plot below the horizontal “Intensity difference = 0” line – those with a negative intensity difference – have a higher intensity on σ_1 ; words which plot above the horizontal line – those with a positive intensity difference – have a higher intensity on σ_2 .

It is easy to see in this case that all verbs have a higher intensity on σ_1 . Thus there is a correlation between intensity and the σ_1 stress I perceived on BSh_03 verbs.

The non-verbs are more or less evenly distributed below and above the “Intensity difference = 0” line. This shows only that intensity does not always correspond to the σ_2 stress pattern observed in these lexical categories. Nothing more specific can be concluded without consideration of the intrinsic effects of vowel height, as described in 2.3.2.1 below. Nonetheless, this graph is extremely useful in showing how acoustic correlates interact to distinguish lexical categories.

2.3.2 Analyzing the acoustic correlates

In this section I provide examples of the analysis of pitch, intensity, and vowel duration. I present a few specific cases in order to illustrate the issues and confounding factors which arise most frequently in the analysis. My objective in providing such detail here is to avoid repetition in the analysis chapters which follow. If the reader loses track, there, of why a certain course was chosen, she can usually look back and find an explanation here.

The objective of this analysis is not simply to determine whether there is a correlation between acoustic parameters and stress, but, more specifically, to determine whether there is a *meaningful* correlation between acoustic parameters and stress. If a correlation occurs as a result of other factors, that is not of real interest. Thus it is essential to identify and control for factors which may have an incidental effect on pitch, intensity, and vowel duration. The factors which come into play are

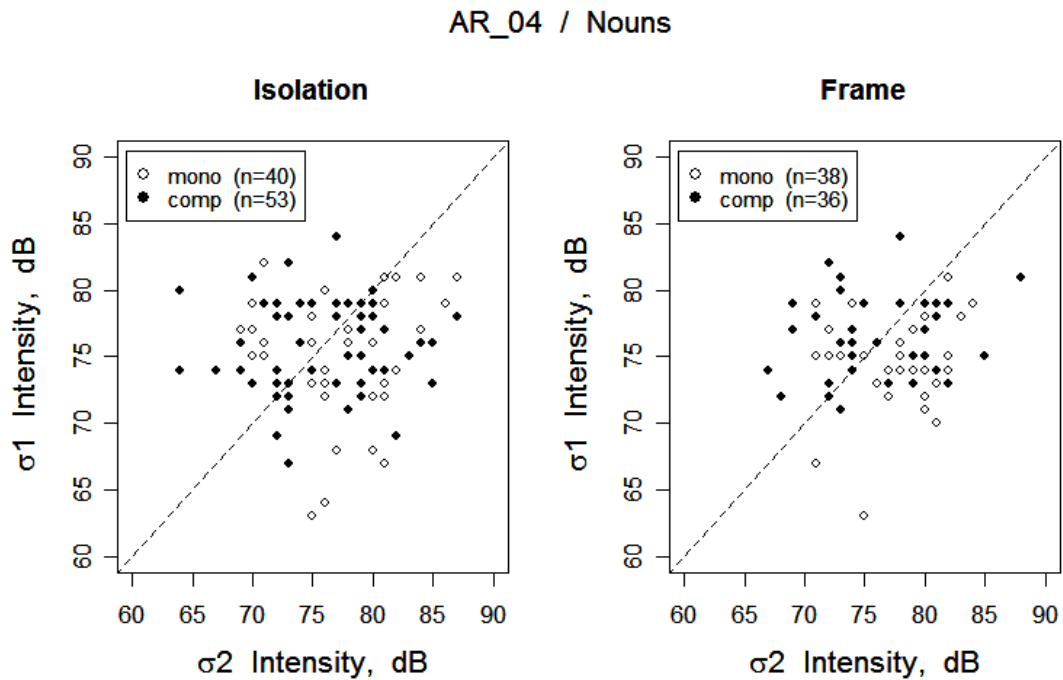
vowel height / vowel quality, syllable closure, compensatory lengthening, and position in the utterance.

2.3.2.1 Pitch and intensity

Analysis of pitch and intensity data began with a simple comparison of values across syllables, using the equations and graphs described above. In some cases this yielded conclusive results. For instance, Figure 2.4 above shows unequivocally that pitch is higher on σ_2 of AR_04 nouns, and is thus a consistent correlate of the perceived σ_2 stress. Likewise, Figure 2.5 above shows unequivocally that both pitch and intensity are higher on σ_1 of BSh_03 verbs, and are thus consistent correlates of the perceived σ_1 stress. (Confirmation of the statistical significance of these graphical patterns is described in section 2.5.1.)

In other cases, the results at this stage were inconclusive. In Figure 2.6 below, about half the AR_04 nouns have a higher intensity on σ_1 , and about half have a higher intensity on σ_2 , for both isolation and frame forms.

Figure 2.6 Inconclusive intensity data

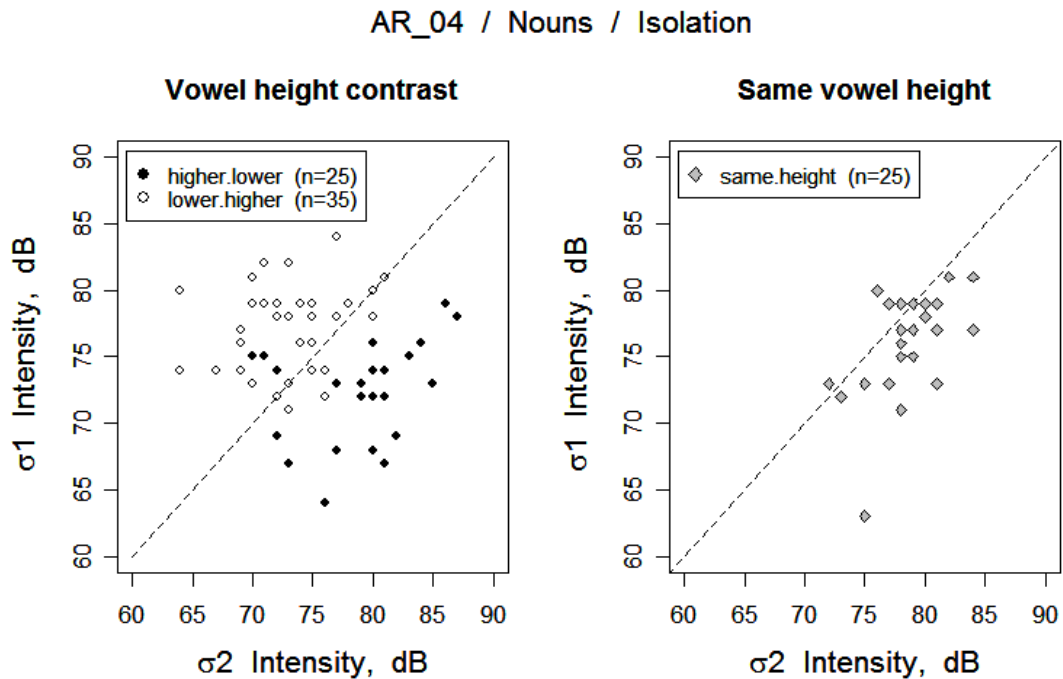


All we can conclude from the pattern here is that intensity does not always correspond to the σ_2 stress pattern observed on Amdo nouns. (This was also the case for BSh_03 non-verbs in Figure 2.5.) We cannot say, at this point, that intensity is definitively *not* a correlate of stress, because other factors may come into play.

As reported by Lehiste (1970: 68, 120), both fundamental frequency and intensity show intrinsic variation as a function of vowel height: Fundamental frequency is intrinsically higher on high vowels and lower on low vowels. The opposite is observed for intensity: intensity is intrinsically higher on low vowels and lower on high vowels. When a plot of *all* the nouns produced by a speaker is not conclusive, as in Figure 2.6, the analysis can then be narrowed by controlling for vowel height.

Figure 2.7 below shows the isolation forms which were plotted in Figure 2.6, but here plotted in terms of the vowel height contrast across syllables, coded as described in Table 2.6.¹³ On the left side, nouns with a contrast in vowel height behave almost exactly as predicted following Lehiste (1970): intensity is higher on whichever syllable has a lower vowel. This provides no definitive information about a potential correlation between intensity and stress.

Figure 2.7 Controlling intensity for vowel height



On the right side of Figure 2.7, nouns with vowels of the same height in both syllables most often have a higher intensity on σ_2 . For this control group, intensity does, indeed, correlate with stress. If it did not – if intensity were random within this

¹³ Note that the total number of nouns plotted in Figure 2.7 is less than the number of isolation forms plotted in Figure 2.6. This occurs because words with a diphthong in either syllable were excluded from consideration when the vowel height contrast was coded.

subset – then the points would be evenly distributed to either side of the dashed line. (The statistical significance of the observed correlation is then examined using a paired-sample t-test, as described in section 2.5.1.2.)

The conclusion that I draw from these patterns is that intensity shows a “limited” correlation with stress in AR_04 nouns produced in isolation.

It is true that the [same.height] nouns are evidence that the speaker manipulates intensity to be higher on σ_2 , the stressed syllable. But this effect is not as strong as the intrinsic correlation between intensity and vowel height. Thus in the [lower.higher] nouns, the speaker’s efforts are obscured. (And in the [higher.lower] nouns they are emphasized.) The net result is that the correlation between intensity and stress occurs only in a limited and controlled subset of the sample.

This net result is what I focus on when I describe intensity as a “limited” correlate of stress. My wording reflects my consideration of whether the listener can rely on intensity as an indication of the locus of stress: intensity is of limited usefulness in this regard.

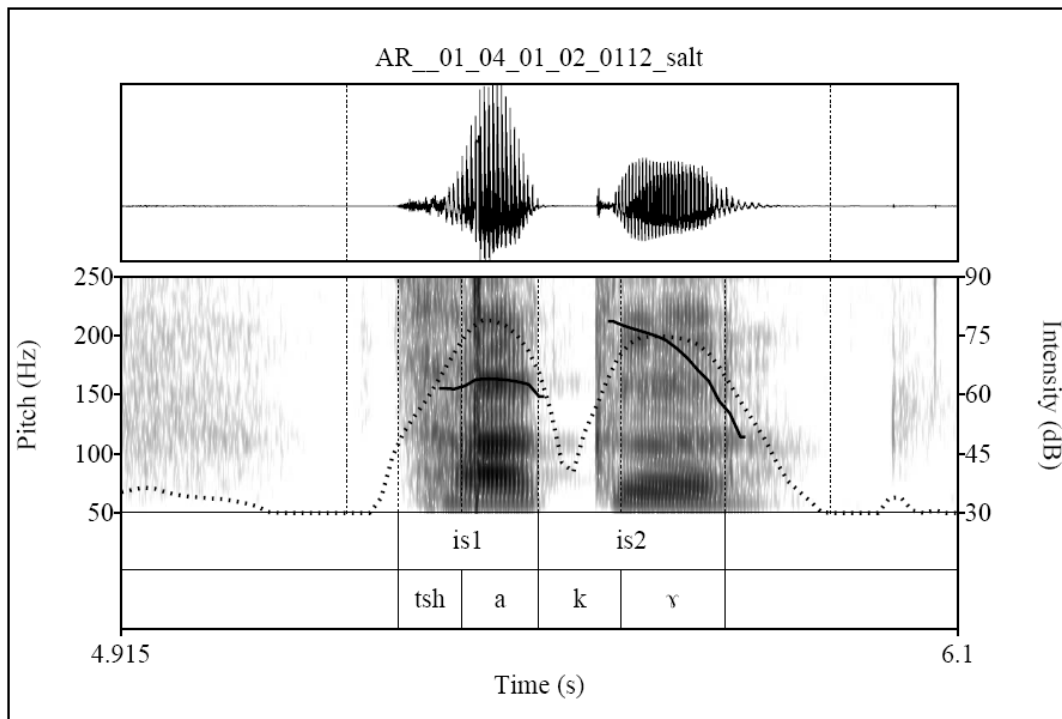
A different way to describe the same patterns would be to say that, yes, intensity is a correlate of stress, though it is sometimes over-ridden by the effects of intrinsic vowel height. I feel that the wording I use is more conservative.

A similar approach was used in the analysis of pitch data, when appropriate. In some cases, such as that illustrated in Figure 2.4, any intrinsic variation in pitch as a function of vowel height is moot because it is never sufficient to yield a higher pitch

on σ_1 than on σ_2 . In such cases I consider the correlation between pitch and stress to be “robust”.

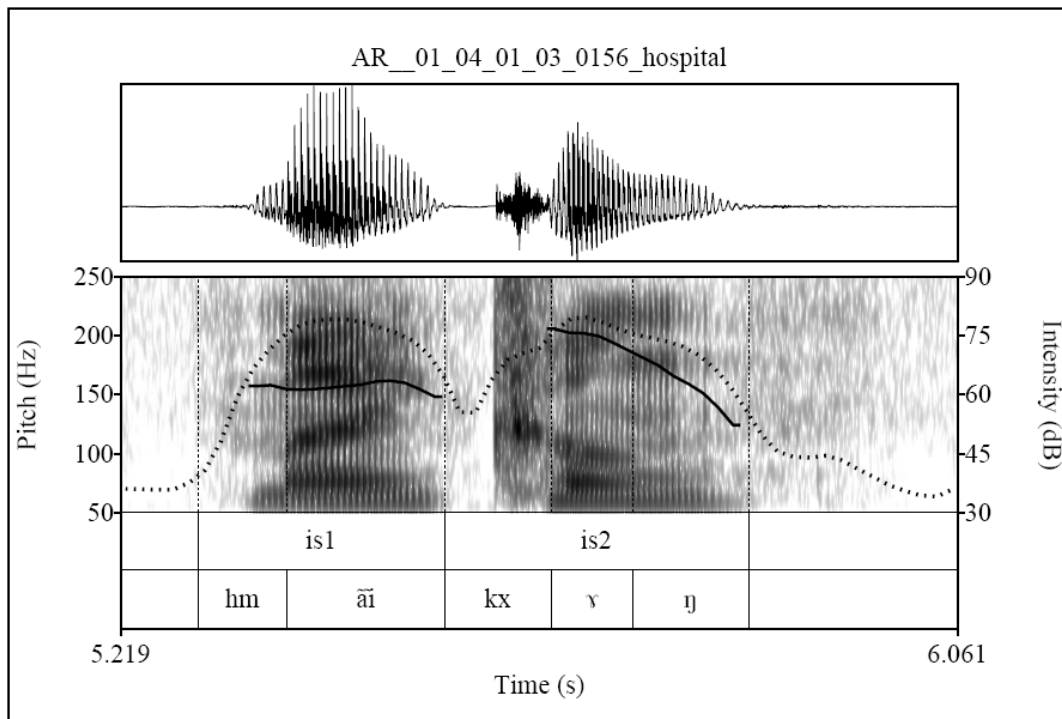
Another factor which has an incidental effect on pitch in non-verbs in Rebkong Amdo is whether σ_2 is open or closed. In Rebkong Amdo, pitch slope is a robust cue for σ_2 stress in the isolation forms of non-verbs, as demonstrated in section 5.3. When σ_2 is open, the slope is borne entirely by the vowel, as is the case in [ts^ha.'kɣ] *tshwa.khu* ‘salt’ shown in Figure 2.8 below. Here, the average pitch over the medial 50% of the vowel in σ_2 is 190 Hz. This value is compared to the average pitch over the medial 50% of the vowel in σ_1 , which is 163 Hz. The pitch difference across syllables is calculated as $190 - 163 = 27$ Hz, which is representative of what the listener hears.

Figure 2.8 Pitch in a noun with σ_2 open: [ts^ha.'kɤ] *tshwa.khu* ‘salt’



When σ_2 is closed, the pitch slope is distributed across the full rhyme, as shown for [h^hmāi.'k^hɤŋ] *sman.khang* ‘hospital’ in Figure 2.9 below. The average pitch measured for the nuclear vowel is thus artificially elevated (200 Hz) compared to the average pitch over the rhyme as a whole (178 Hz). The ideal solution to this problem would have been to measure pitch over the rhyme, in closed syllables; as a *post-hoc* alternative, the effects of syllable closure can be controlled for by focusing analysis on words with σ_2 open. For speaker AR_04, the contrast in pitch across syllables is robust regardless of syllable closure; for speaker AR_05 syllable closure is a more significant factor, as discussed in detail in section 5.3.2.

Figure 2.9 Pitch in a noun with $\sigma 2$ closed: AR_04 [^hmāi.'k^hɣŋ] *smān.khāng* ‘hospital’



2.3.2.2 Vowel duration

The duration of vowels is influenced by a number of factors, and – unlike pitch and intensity – it is never possible to derive meaningful results for a group of words as a whole without taking these factors into consideration from the start.

Most important: in isolation forms of non-verbs – which are stressed on $\sigma 2$ – the vowel in $\sigma 2$ is subject to lengthening simply on the basis of being the final syllable in the utterance. There is thus no possible way to distinguish between lengthening which is associated with final position, and lengthening which is associated with stress. This means that vowel duration cannot be assessed at all in the isolation forms of nouns, adjectives, and numerals.

Second, the role of vowel duration as a correlate of stress cannot be determined if comparisons are confounded by compensatory lengthening or by the presence of diphthongs in one syllable or the other. To obtain meaningful results, words influenced by these factors must be excluded from consideration.

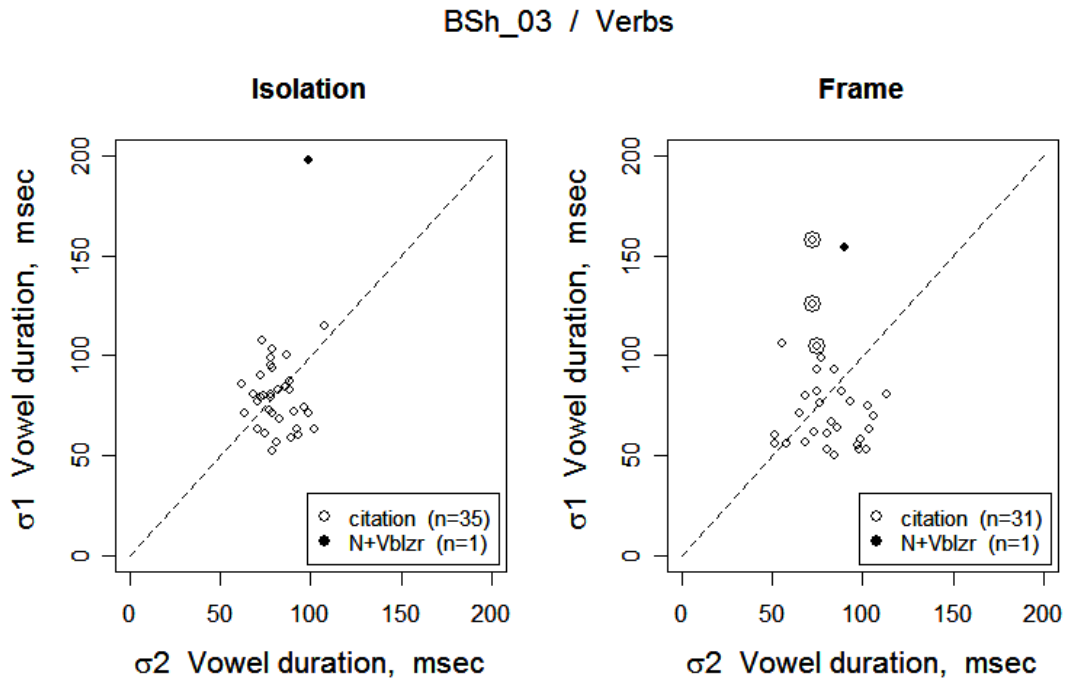
Vowel duration is also influenced by properties of the vowels themselves and of the syllables in which they occur. Lehiste (1970: 18ff) summarizes several studies demonstrating that vowel duration varies as a function of vowel height (low vowels tend to be longer than high vowels) and as a function of the voicing of a following coda (vowels tend to be longer before a voiced consonant than before a voiceless consonant – this factor is never relevant in the present study). In addition, vowels tend to be longer in open syllables than in closed syllables. A contrast in any of these features across syllables may have a strong influence on relative vowel duration. Tibetan is certainly vulnerable to such effects since, in so many words, the second syllable is open and has [a] as its nucleus, while the first syllable may vary widely in both respects.

The most effective way to eliminate these incidental factors is to focus on a subset of nouns which is controlled for vowel quality and for syllable closure type (according to the coding parameters listed in Table 2.6), and to take into consideration the setting (isolation *vs.* frame) as well.

The case of verbs in Balti illustrates some of these issues. As shown in Figure 2.10 below, for both isolation and frame forms produced by speaker BSh_03, it appears that the vowel is longer in σ_1 for about half the verbs, and longer in σ_2 for

the other half. This suggests that there is no correlation between vowel duration and the perceived $\sigma 1$ stress pattern.

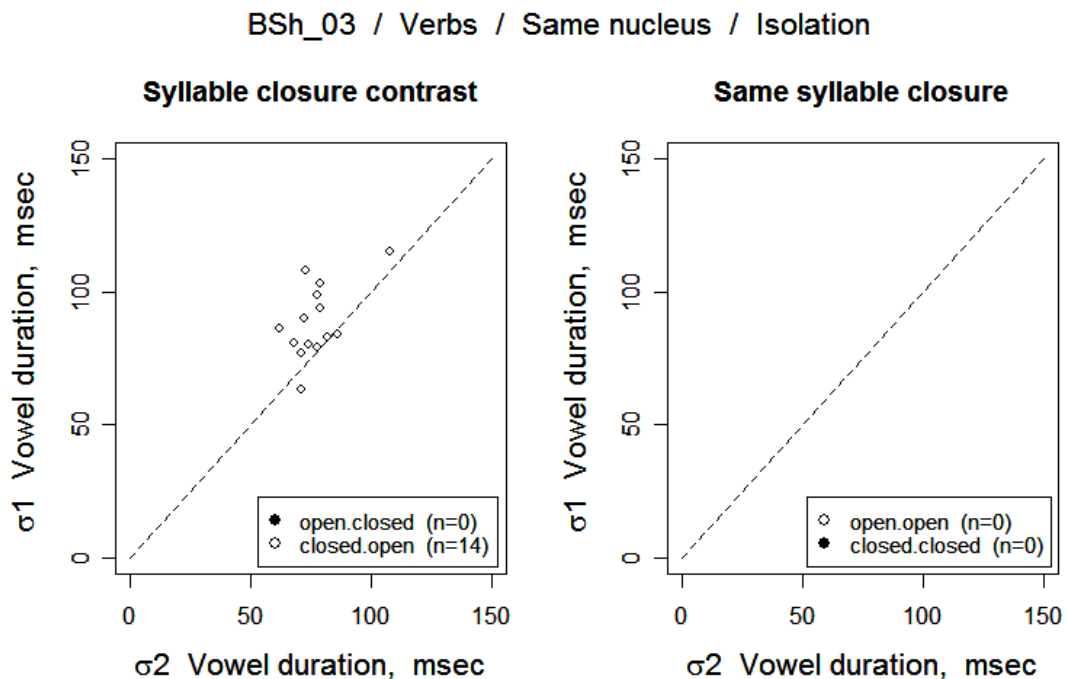
Figure 2.10 Inconclusive vowel duration data (circles indicate compensatory lengthening on $\sigma 1$)



In fact, though, BSh_03 verbs *do* show a correlation between vowel duration and stress placement (that is, the isolation forms do), but this is obscured by some of the factors discussed above. First, we can exclude from further consideration the lone verb with a Noun+Verbalizer structure ([*'za:n.za*] *zan.za* 'to eat food'. With only one example, no conclusions can be formed about this morphological group, and it clearly behaves differently than the citation forms of verbs: the vowel in $\sigma 1$ is unusually long.

Since the citation forms always have [-pa] as the second syllable, any intrinsic variation in duration as a function of vowel quality / vowel height can be controlled for by focusing only on words which also have [a] in the first syllable, such as [ˈblaq.pa] 'breg.pa' 'to shave' and [ˈɰzar.pa] 'dzar.ba / gzar.ba / bzar.ba' 'to drip'. The isolation forms of such [a.a] verbs are plotted in Figure 2.11 below, coded for syllable closure types. The graph on the right has no points at all, because there are no [a.a] verbs with the same closure type in both syllables. In the graph on the left there are no [open.closed] forms. (Tautologically, as the citation forms all end in [-pa].)

Figure 2.11 Controlling for vowel quality and syllable closure type



Since all fourteen of the [a.a] verbs have a [closed.open] syllable structure, there are two factors which favor a longer vowel in σ_2 , and which would lead us to

expect these verbs to plot to the right of the dashed line. First, vowels are generally longer in open syllables than in closed syllables. Second, in isolation, the vowel in σ_2 is in utterance-final position and likely to be lengthened on that basis, as well.

Instead, though, nearly all of the [a.a] verbs plot to the left of the dashed line, indicating that the vowel is longer on σ_1 . The most plausible explanation for this is that vowel duration is preferentially lengthened on this stressed syllable.

What we can conclude here is that vowel duration for isolation forms of BSh_03 verbs shows a “limited” correlation with σ_1 stress. Vowel duration cannot be considered a robust or exceptionless cue for stress here, since the pattern is only evident when it is not obscured by other factors, such as a contrast in vowel quality across syllables.

The analyses of vowel duration data for Balti (Chapter 4) and Rebkong Amdo (Chapter 5) take into consideration the factors discussed here: setting (isolation vs. frame), position in the utterance, syllable closure, and vowel quality / vowel height. Once these features are controlled for, vowel duration can also be assessed by calculating the ratio of measurements across syllables. This is discussed further in section 2.5.2.3.

2.4 Analyzing pitch slope data

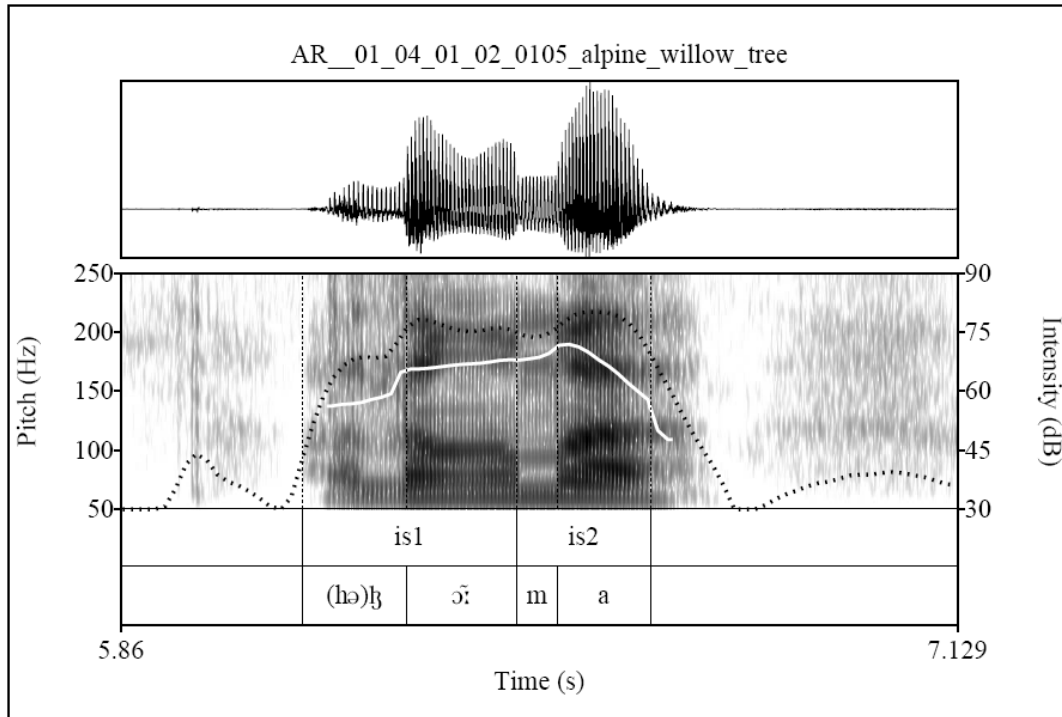
Pitch slope differs from the other acoustic parameters in several respects. The core difference is that pitch slope is not a scalar quantity, but is instead a vector with

direction as well as magnitude. That is, pitch may slope either upward or downward, and the slope may be gentle or steep. As a result, a different graphing protocol is necessary – which unfortunately is a little more difficult to interpret.

As described in section 2.2.2 above, the pitch slope in a vowel is determined by subtracting the pitch measured at the start of an interval from the pitch measured at the end of the interval, and then dividing by the elapsed time. This yields a slope ($\Delta y / \Delta x$) in units of Hz/sec – i.e., Hz/1000msec. Dividing this value by 10 converts the result to units of Hz/100msec. This is conceptually a more useful measure, since 100 msec is a reasonable duration for a vowel in natural speech.

An example from Rebkong Amdo – where pitch slope is a robust correlate of $\sigma 2$ stress – is the word [h^aʒ̥ɔ̃ː.ma] *glang.ma* ‘alpine willow tree’, illustrated in Figure 2.12 below. In $\sigma 1$, pitch slope was measured as 4.73 Hz/100msec, which is a very gentle upward slope. (Since the actual duration of the vowel is 168 msec, the total increase in pitch over the span of the vowel is 8 Hz.) For $\sigma 2$, pitch slope was measured as -40 Hz/100msec, a steep downward slope. (With a duration of 141 msec, the total change in pitch over the span of the vowel is -56 Hz.)

Figure 2.12 Pitch slope / σ_1 upward, σ_2 downward: AR_04 [h^hʒɔ̃.'ma] *glang.ma* ‘alpine willow tree’



In this case, pitch in σ_1 slopes upward, while pitch in σ_2 slopes downward. Of course, there are other possibilities: downward in σ_1 and upward in σ_2 , or downward in both syllables. Theoretically, pitch could slope upward in both syllables, but there are no such cases in Rebkong Amdo – for either speaker, in any lexical category – and only a few such cases in Balti, with very gentle upward slopes – so gentle that the slope is nearly flat.

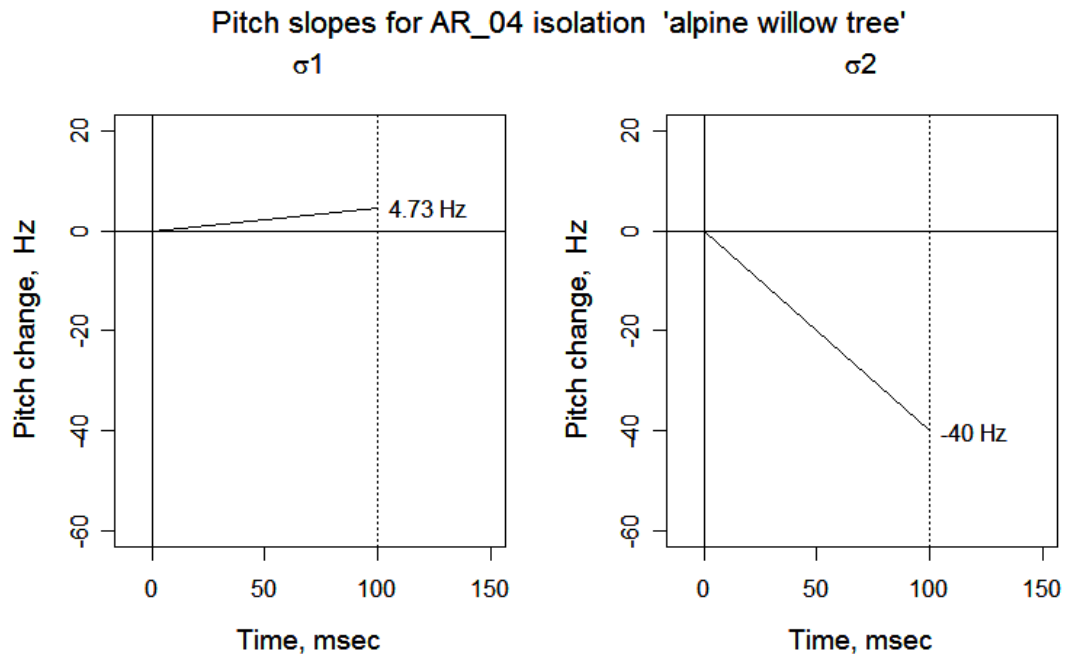
Note that in discussing pitch slope I avoid terminology which is commonly used to describe contour tones. First of all, Balti and Rebkong Amdo Tibetan are not tone languages, so terms like “rising” pitch or “falling” pitch could be confusing. Instead, I say that “pitch slopes upward” or “pitch slopes downward”. Second, rising

and falling tones in tone languages are not usually quantified. Here, by using the term “slope” – with its mathematical connotation – rather than a term like “contour”, I emphasize that pitch slope measurements quantify the rate of change in pitch.

2.4.1 Calculating differences in pitch slope across syllables

As with pitch, intensity, and vowel duration, the contrast in pitch slope across syllables is calculated by subtracting the measurement in σ_1 from the measurement in σ_2 . In the case of the isolation form of [h³³ʂɔ̃.'ma] *glang.ma* ‘alpine willow tree’ as produced by speaker AR_04 – whose pitch trace was shown above – the difference in slope is $(-40 \text{ Hz}/100\text{msec}) - (4.73 \text{ Hz}/100\text{msec}) = -44.73 \text{ Hz}/100\text{msec}$, rounded to $-45 \text{ Hz}/100\text{msec}$. This relationship is illustrated graphically in Figure 2.13 below.

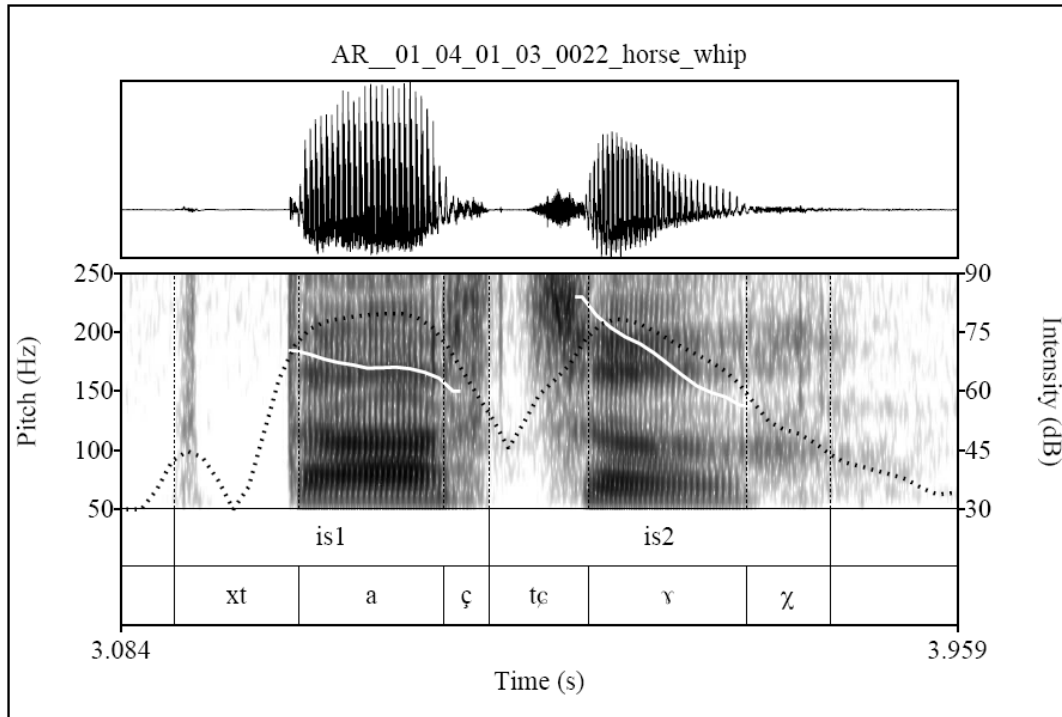
Figure 2.13 Pitch slope values across a disyllabic word



Whenever the difference in pitch slopes is a negative number – as is the case here – the relationship between slopes can be described by saying that the vector representing σ_2 is “below” the vector representing σ_1 . Or we can say that the slope in σ_2 is “more downward” than the slope in σ_1 . Crucially, a negative difference in pitch slope values thus means that slope is more prominent in σ_2 . (Note that this is exactly the opposite of how pitch, intensity, and vowel duration data was interpreted; there, a negative difference meant that the parameter was more prominent in σ_1 .)

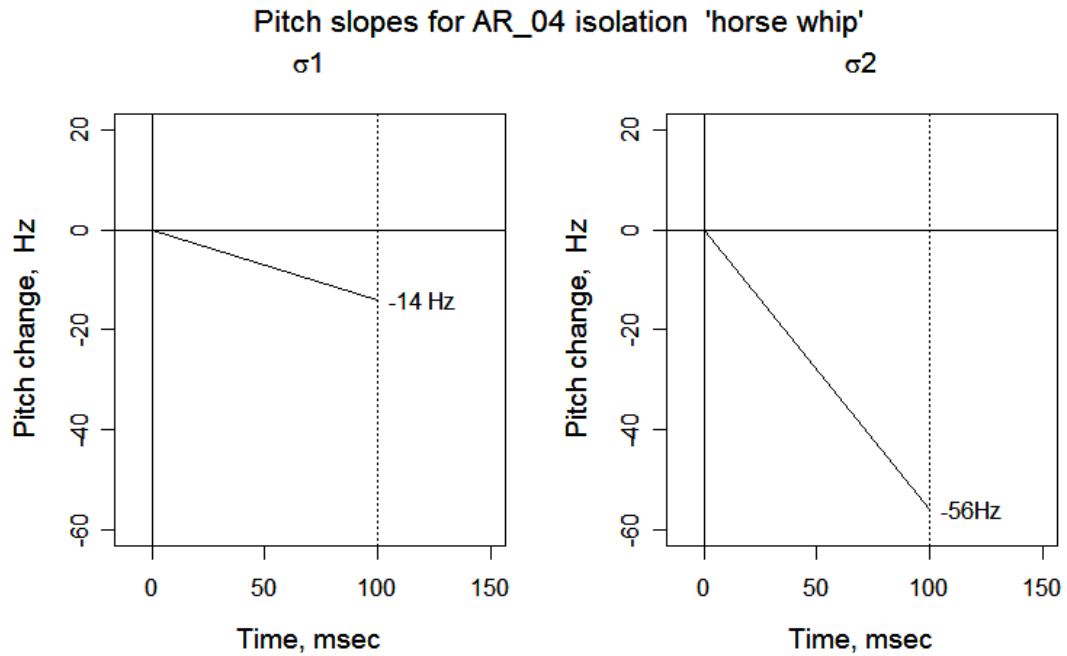
Such a relationship between vectors does not always require an upward slope in σ_1 . For instance, in [xtax.'tɕɾɣ] *rta.lcag* ‘horse whip’, shown below, pitch slopes downward in both syllables – in σ_1 at -14 Hz/100msec, and in σ_2 at -56 Hz/100msec (which is very steep, and very noticeable). The difference in pitch slopes is $(-56) - (-14) = -42$ Hz/100msec.

Figure 2.14 Pitch slope / σ_1 downward, σ_2 more downward: AR_04 [xtax.'tɕɣχ] rta.lcag 'horse whip'



We can again say here that the vector representing σ_2 is “below” the vector representing σ_1 , or that pitch slopes “downward more steeply” in σ_2 than in σ_1 . This relationship is illustrated graphically in Figure 2.15 below.

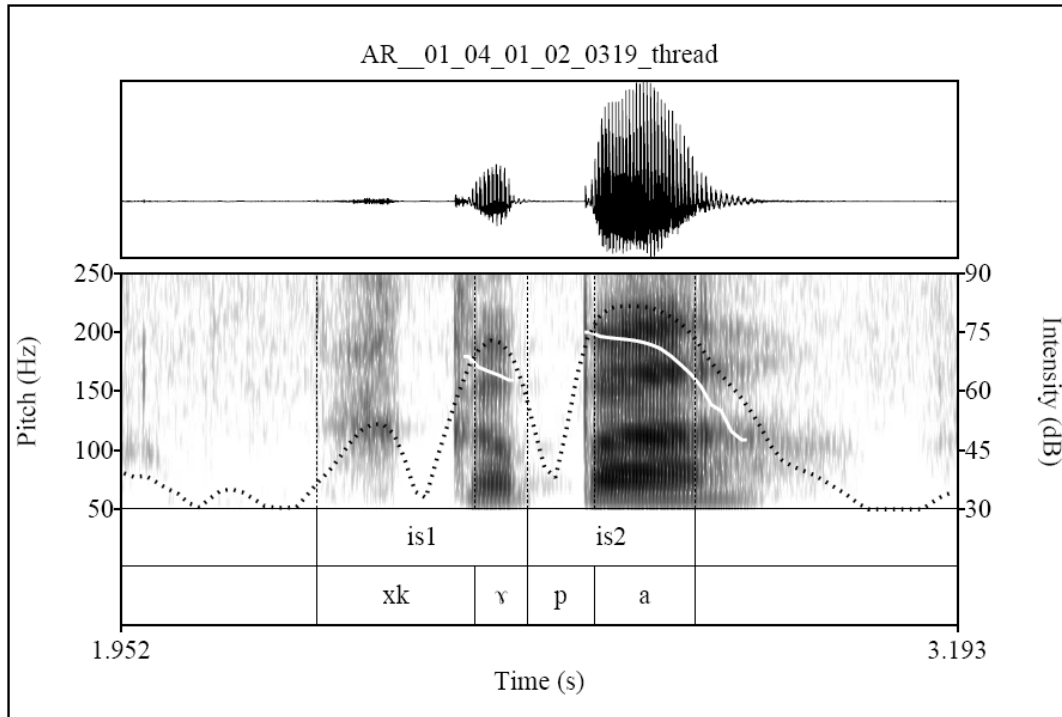
Figure 2.15 Pitch slope values for [xtax.'tɕɿχ] *rta.lcag* ‘horse whip’



Cases like the two shown above – where pitch slope is “more downward” in σ_2 than in σ_1 – are the most common types in my data. In a small number of words, pitch slope is “more downward” in σ_1 than in σ_2 . These differences are usually small, as we see in the pitch trace for [xkɿ.'pa] *skud.pa* ‘thread’ shown Figure 2.16 below.

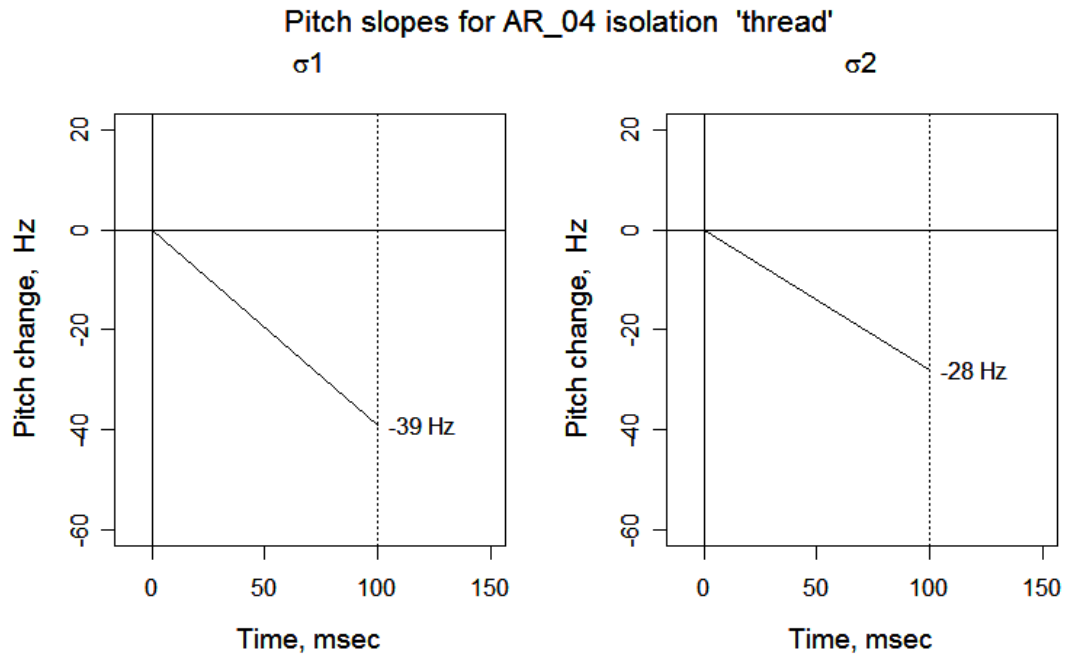
Here, slope was measured as -39 Hz/100msec in σ_1 , and -28 Hz/100msec in σ_2 , for a difference of $(-28) - (-39) = 11$ Hz/100msec.

Figure 2.16 Pitch slope / σ_1 more downward, σ_2 downward: AR_04 [xk γ .'pa] *skud.pa* 'thread'



This positive difference in pitch slope can be described by saying that the vector representing σ_2 is “above” the vector representing σ_1 , or by saying that pitch slope is “more downward” in σ_1 than in σ_2 .

Figure 2.17 Pitch slope values for [xkʁ.'pa] *skud.pa* 'thread'



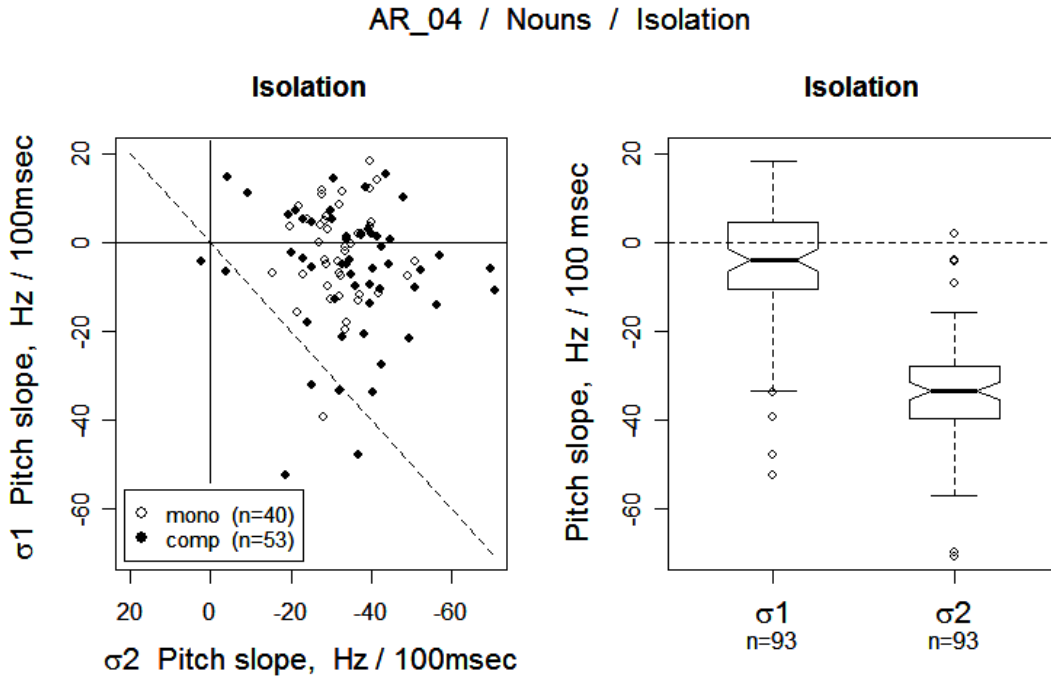
2.4.2 Graphing pitch slope alone

Because pitch slope is a vector – with direction as well as magnitude – when the measurements are plotted, the resultant graphs are more complex than those for the scalar entities pitch, intensity, and vowel duration.

As an example, pitch slope measurements for the isolation forms of nouns produced by speaker AR_04 are shown in Figure 2.18 below. In the distribution plot on the left, measurements for $\sigma 2$ are represented on the x-axis, and measurements for $\sigma 1$ are represented on the y-axis; this was also the protocol for pitch, intensity, and vowel duration. What is different here – compared to the graph of pitch data in Figure 2.4, for instance – is that both negative and positive values need to be shown for both

syllables (for downward and upward slopes, respectively), and so the graph includes the origin, (0,0).

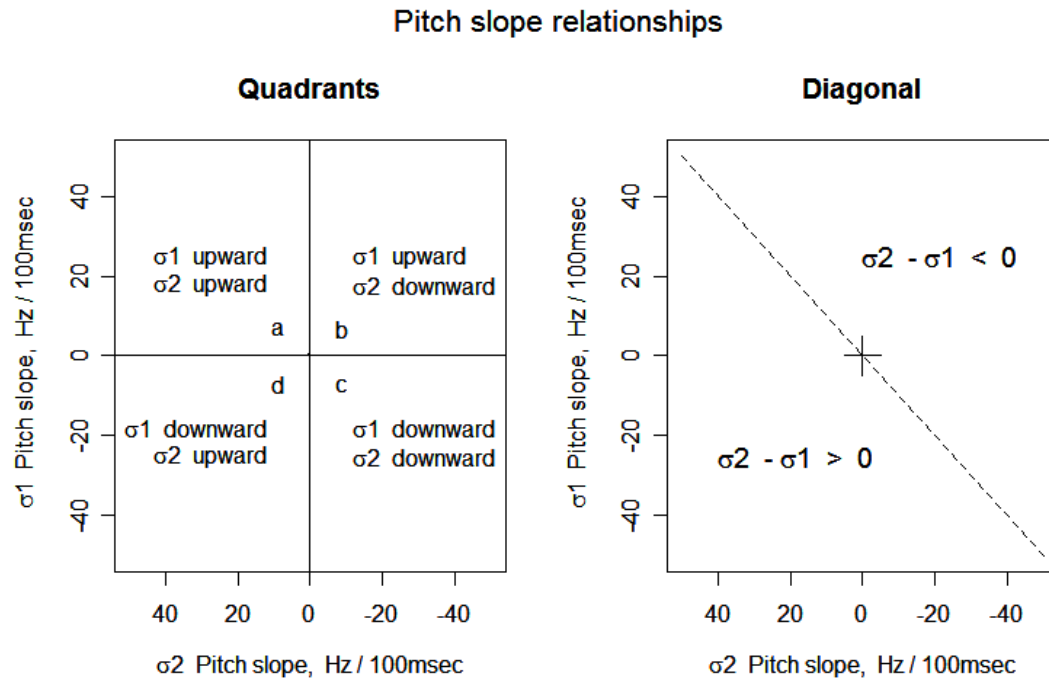
Figure 2.18 Graphing pitch slope measurements



The x- and y-axes divide the graph into four quadrants which represent different relationships between pitch slopes in $\sigma1$ and $\sigma2$. These are shown schematically in the graph on the left in Figure 2.19 below. All points falling above the x-axis have an upward slope in $\sigma1$; all points falling below the x-axis have a downward slope in $\sigma1$. All points falling to the left of the y-axis have an upward slope in $\sigma2$; all points falling to the right of the y-axis have a downward slope in $\sigma2$. Nearly all of the AR_04 isolation forms of nouns plotted in Figure 2.18 above fall in quadrants (b) and (c). That is, in $\sigma1$ pitch sometimes slopes upward and sometimes slopes downward; in $\sigma2$, pitch always slopes downward (with one exception). These

patterns are reflected in the box-and-whisker plot on the right in Figure 2.18. The AR_04 noun [h^əʒ̥ɿ.ma] *glang.ma* ‘alpine willow tree’ – shown in Figure 2.12 and Figure 2.13 – plots in quadrant (b), at (σ_2 , σ_1) coordinates (-40, 4.73).

Figure 2.19 Pitch slope relationships

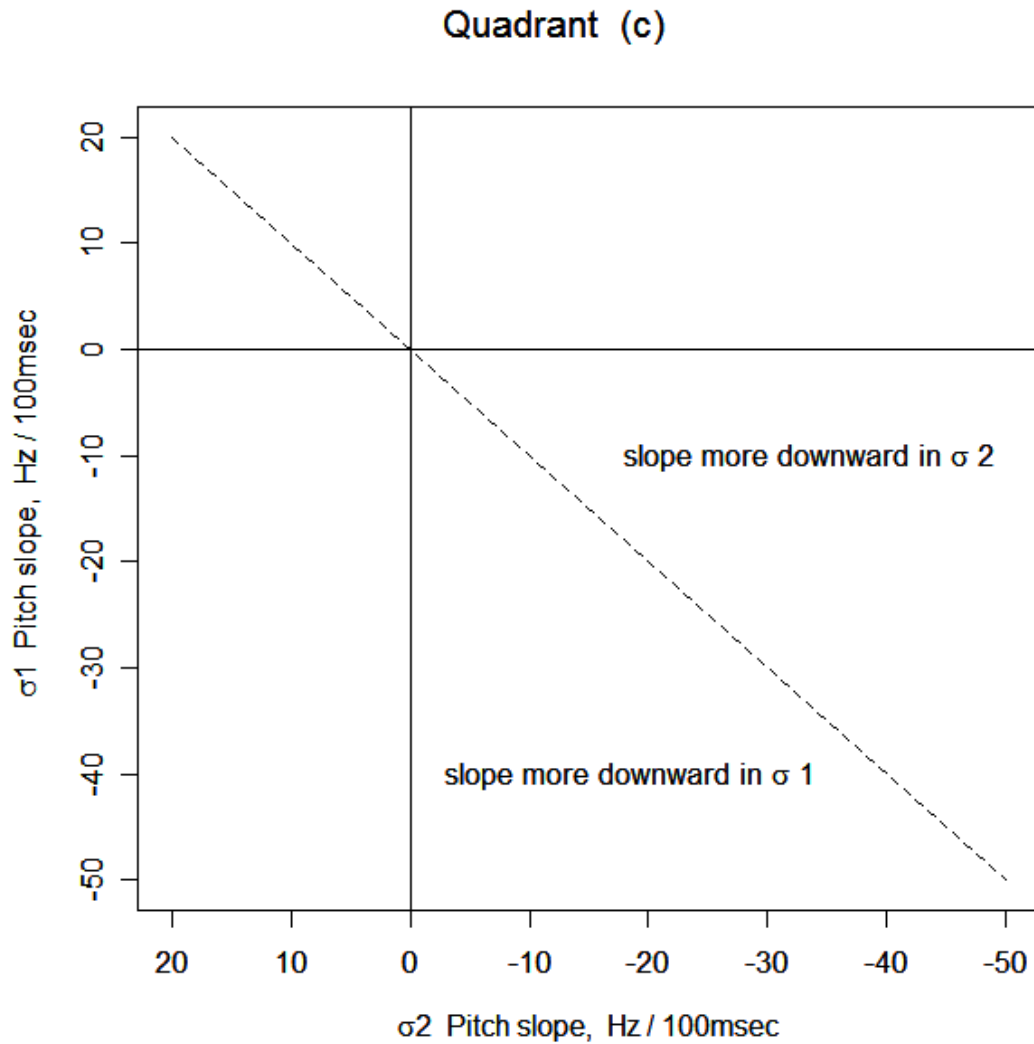


In the graph on the right in Figure 2.19 above, the diagonal line represents all theoretical cases in which the slopes in σ_1 and σ_2 are equal. This includes cases in which pitch slopes downward in both syllables – i.e., words that would fall in quadrant (c) – and cases in which pitch slopes upward in both syllables – i.e., words that would fall in quadrant (a). In words that plot to the right of the diagonal line, the difference in pitch slope across syllables is negative, meaning that pitch slope is “more downward” in σ_2 – i.e., the vector representing σ_2 lies below the vector

representing σ_1 . In words that plot to the left of the diagonal line, the difference in pitch slope across syllables is positive, meaning that pitch slope is “more downward” in σ_1 – i.e., the vector representing σ_2 lies above the vector representing σ_1 .

As noted, in Figure 2.18 nearly all of the AR_04 nouns fall in quadrants (b) and (c). Within quadrant (c), nearly all points fall to the right of the dashed line. A closer look at this quadrant is provided in Figure 2.20 below. Pitch slopes downward on both syllables of words which plot in this quadrant; but points which fall to the right of the dashed line represent words in which pitch slope is more downward in σ_2 – i.e., the vector representing σ_2 lies below that representing σ_1 . This was the case for the noun [xtax.'tɕɿχ] *rta.lcag* ‘horse whip’, shown in Figure 2.14 and Figure 2.15; the (σ_2, σ_1) coordinates of this word are (-56, -14). The noun [xkɿ.'pa] *skud.pa* ‘thread’ – shown in Figure 2.16 and Figure 2.17 – also plots in quadrant (c), but falls to the left of the dashed line, at coordinates (-28, -39). In this case, pitch slopes downward more steeply in σ_1 than in σ_2 – i.e., the vector representing σ_2 lies above that representing σ_1 .

Figure 2.20 Pitch slope relationships in quadrant (c)



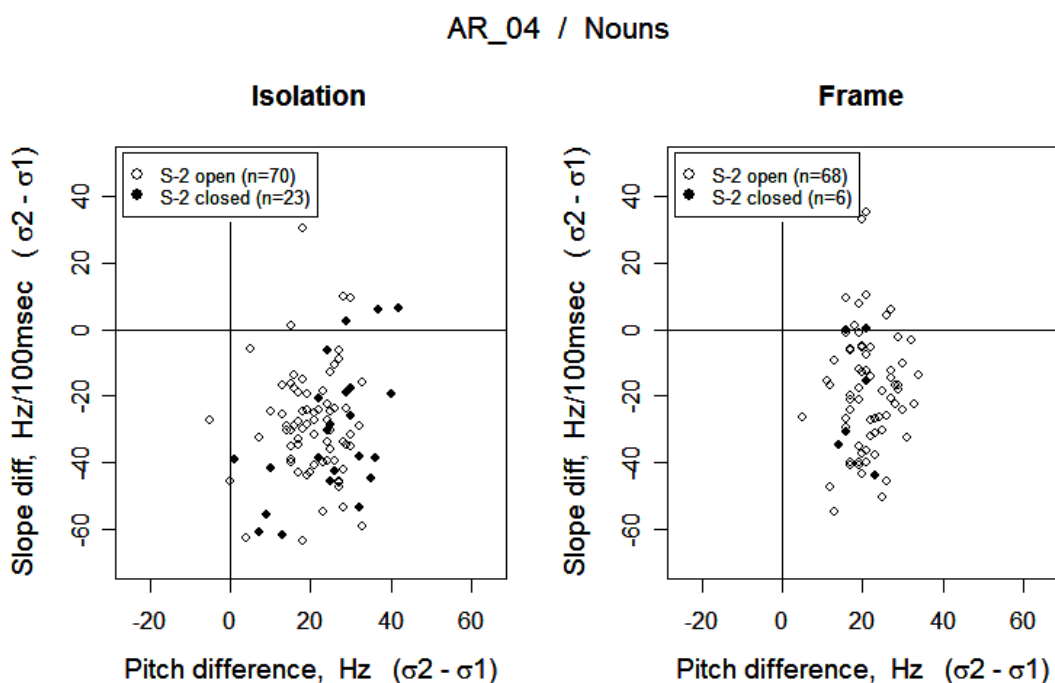
In summary, the graph in Figure 2.18 tells us that, in $\sigma 1$, pitch sometimes slopes upward and sometimes slopes downward; in $\sigma 2$, pitch always slopes downward (with one exception); and pitch slope is almost always “more downward” (i.e., more prominent) in $\sigma 2$ than in $\sigma 1$.

2.4.3 Graphing pitch and pitch slope together

Pitch and pitch slope can be regarded as two reflexes of one acoustic resource – fundamental frequency. In the non-verbs, pitch and pitch slope sometimes reinforce one another as cues for stress, and sometimes complement one another. In verbs, the contrast in pitch across syllables obscures any potential contrast in pitch slope across syllables. (This situation is discussed in section 2.5.3.)

In Figure 2.21 below, pitch and pitch slope are plotted for nouns produced by speaker AR_04. The x-axis shows the difference in pitch across the two syllables of each word, calculated as $\text{pitch}_{\sigma 2} - \text{pitch}_{\sigma 1}$. The one or two nouns which plot to the left of the vertical “Pitch difference = 0” line have a higher pitch on $\sigma 1$; all the rest plot to the right of the line, with a higher pitch on $\sigma 2$, the stressed syllable. (Compare to Figure 2.4 – the vertical line here corresponds to the diagonal line there.)

Figure 2.21 Graphing pitch difference vs. pitch slope difference: AR_04 nouns



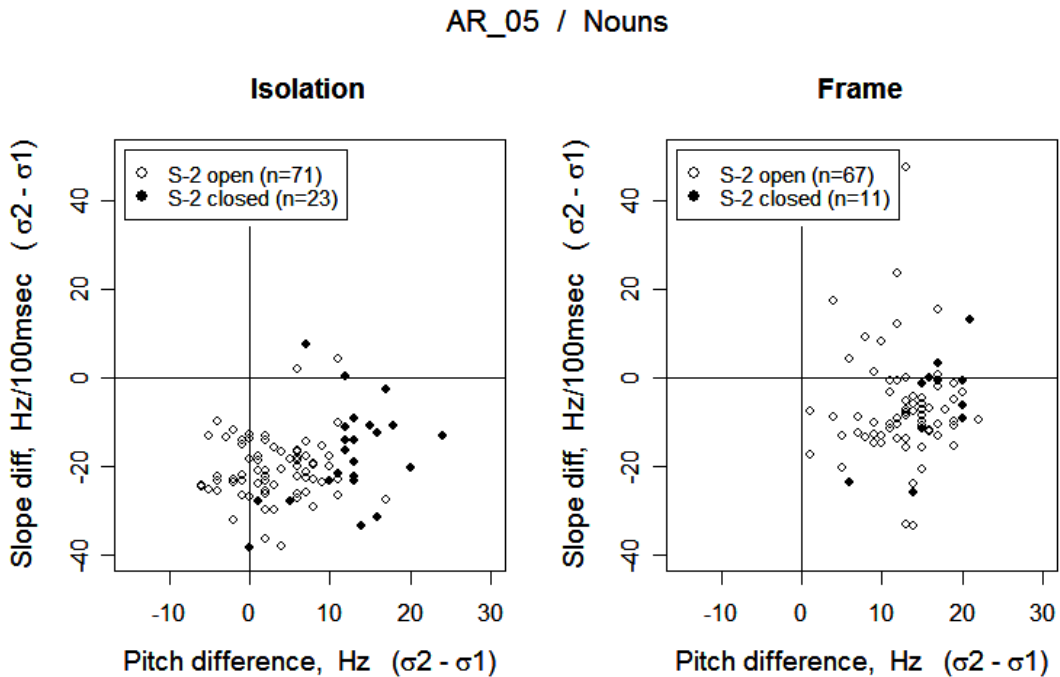
The y-axis in the graphs shows the difference in pitch slope across syllables: $\text{slope}_{\sigma_2} - \text{slope}_{\sigma_1}$. In words which plot below the horizontal “Pitch slope difference = 0” line – those with a negative slope difference – pitch slope is “more downward” in σ_2 than in σ_1 . (These are the points which fall to the right of the dashed line shown in Figure 2.18, Figure 2.19, and Figure 2.20.) In words which plot above the horizontal line, pitch slope is “more downward” in σ_1 than in σ_2 .

Most nouns in Figure 2.21 plot in the lower right quadrant, where both pitch and pitch slope are more prominent in σ_2 , the stressed syllable. For these words, the two reflexes of fundamental frequency reinforce one another as acoustic correlates of stress. A minority of nouns plot above the x-axis; here, it does not matter that pitch slope does not correlate with stress, because pitch itself does. Note that there are no

points in the upper left quadrant of the graph; these would be cases in which neither pitch nor pitch slope was more prominent in σ_2 , the stressed syllable.

A different situation is illustrated by the graph of pitch differences and pitch slope differences for nouns produced by speaker AR_05, in Figure 2.22 below. The isolation forms are of particular interest. Nouns with σ_2 open form a cluster approximately centered on the y-axis, the vertical “Pitch difference = 0” line. This means that pitch is not a consistent and reliable cue for σ_2 stress. (If it were, points would fall well to the right.) However, in these words pitch slope is always more prominent in σ_2 , with points falling well below the horizontal “Pitch slope difference = 0” line. Thus for this group, pitch slope is the reflex of fundamental frequency which is manipulated to convey stress.

Figure 2.22 Graphing pitch difference vs. pitch slope difference: AR_05 nouns



2.5 The notion of significance

As noted above, the objective of this analysis is to determine whether there is a meaningful correlation between acoustic patterns and perceived stress. One aspect of being “meaningful” is that an observed correlation must be attributable to the speaker’s [unconscious] manipulation of the acoustic property in order to convey stress. That is, factors which contribute to an incidental correlation between an acoustic property and stress – factors such as vowel height / vowel quality, syllable closure, position in the utterance, and compensatory lengthening – must be identified and controlled for. This was addressed in sections 2.3 and 2.4 above.

The second aspect of being “meaningful” is that the contrast across syllables must be significant. In sections 2.5.1, 2.5.2, and 2.5.3 below, I discuss statistical, perceptual, and contextual significance, respectively.

2.5.1 Statistical significance

In the course of this study, I used only three simple graphical and mathematical tests to determine whether correlations between acoustic patterns and stress patterns were statistically significant.

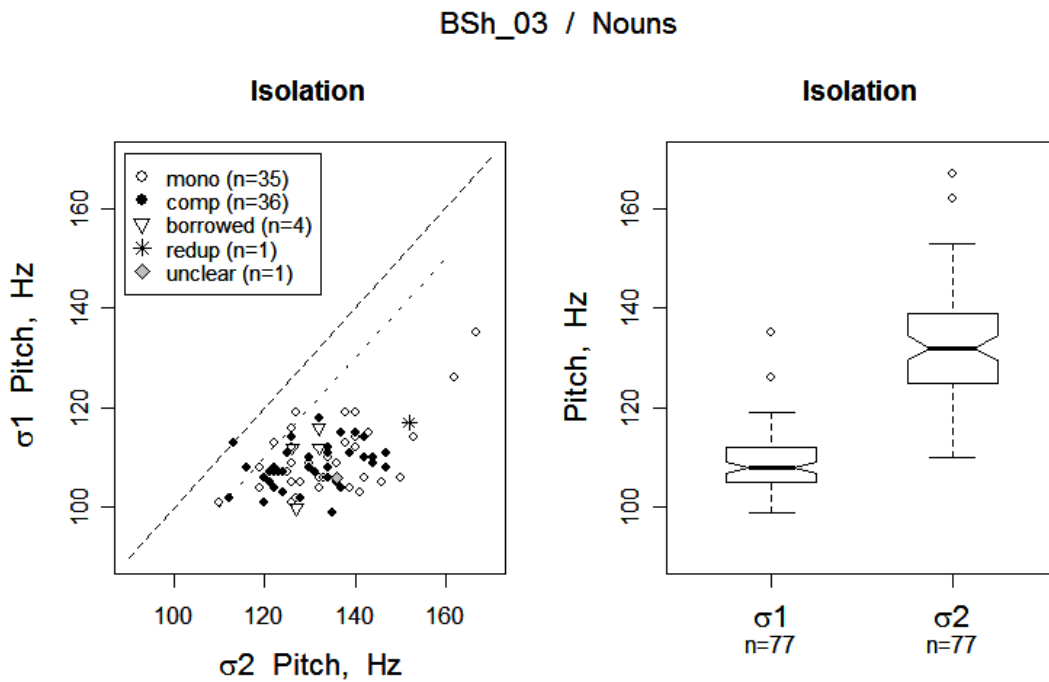
Box-and-whisker plots provide a graphical means of interpreting the distribution of measurements; these are discussed in section 2.5.1.1. Paired-sample t-tests – discussed in section 2.5.1.2 – provide a more accurate assessment of the difference in, say, pitch across syllables for Balti nouns. Finally, in section 2.5.1.3 I

discuss the Welch t-test, which I used only to demonstrate that verbs produced by speaker BM_01 differ significantly by morphological structure.

2.5.1.1 Interpreting box-and-whisker plots

Box-and-whisker plots are helpful in illustrating the distribution of data. As an example, pitch measurements for the isolation forms of BSh_03 nouns are illustrated in Figure 2.23 below. In the distribution plot, all points fall to the right of the dashed line, indicating that pitch is higher in $\sigma 2$. The same data is shown in the box-and-whisker plot at right.

Figure 2.23 Boxplot illustrating distribution of pitch measurements



For each syllable, the thick horizontal line at the waist represents the median value. (The median provides a more useful measure of central tendency here than the

mean, because it is less vulnerable to the effects of one or a few high or low observations.)

The length of the box itself corresponds to the interquartile range (IQR), the range within which 50% of the measured pitch values are clustered. As described in the R documentation for “Box Plot Statistics”¹⁴, the boundaries of the lower and upper “hinges” generally correspond to the first and third quartiles – i.e., 25% of the measurements fall within the lower hinge, and 25% of the measurements fall within the upper hinge. (The hinges of the box need not be symmetrical; one hinge could be quite narrow, if the 25% of the measurements it represents falls within a tight range.) The “whiskers” which sprout from the bottom and top of the box extend, respectively, to the smallest / largest measurements which are within a distance of (1.5 x IQR) from the lower / upper edge of the box. Outliers – if any – are points which lie beyond this range; they are indicated by open circles. If there are no outliers, then the span of the whiskers corresponds to the complete range of measurements.

The notches in the plot correspond roughly to the 95% confidence interval about the median.¹⁵ For the 77 BSh_03 nouns above, the median calculated for σ_1 is 108 Hz. If we were to repeatedly record groups of nouns from this speaker, measure the pitch on the first syllable of each noun, and then repeatedly calculate the median of each group of nouns we record, there is a 95% probability that these median values would fall within the range defined by the notch about 108 Hz we see here. (It

¹⁴ <http://www.r-project.org/>; R Development Core Team (2007)

¹⁵ For detailed calculations see the R documentation for “Box Plot Statistics”, <http://www.r-project.org/>

sometimes happens that the 95% confidence interval exceeds the 25th or 75th percentile, in which case the notch will appear to be “folded”. Examples of this will be seen in Chapters 4 and 5.)

Comparing notches is essentially a graphical analysis of variance, providing a good indication of whether two [or more] medians can be considered statistically different. When there is no overlap of notches, as is the case in Figure 2.23, the null hypothesis (that the two medians are statistically the same, representing different random samples from one common population) is rejected. That is, in this situation, the medians are deemed statistically different. Specifically, here we conclude (with 95% confidence) that there is a significant difference between the median pitch of σ_1 (108 Hz) and the median pitch of σ_2 (132 Hz).

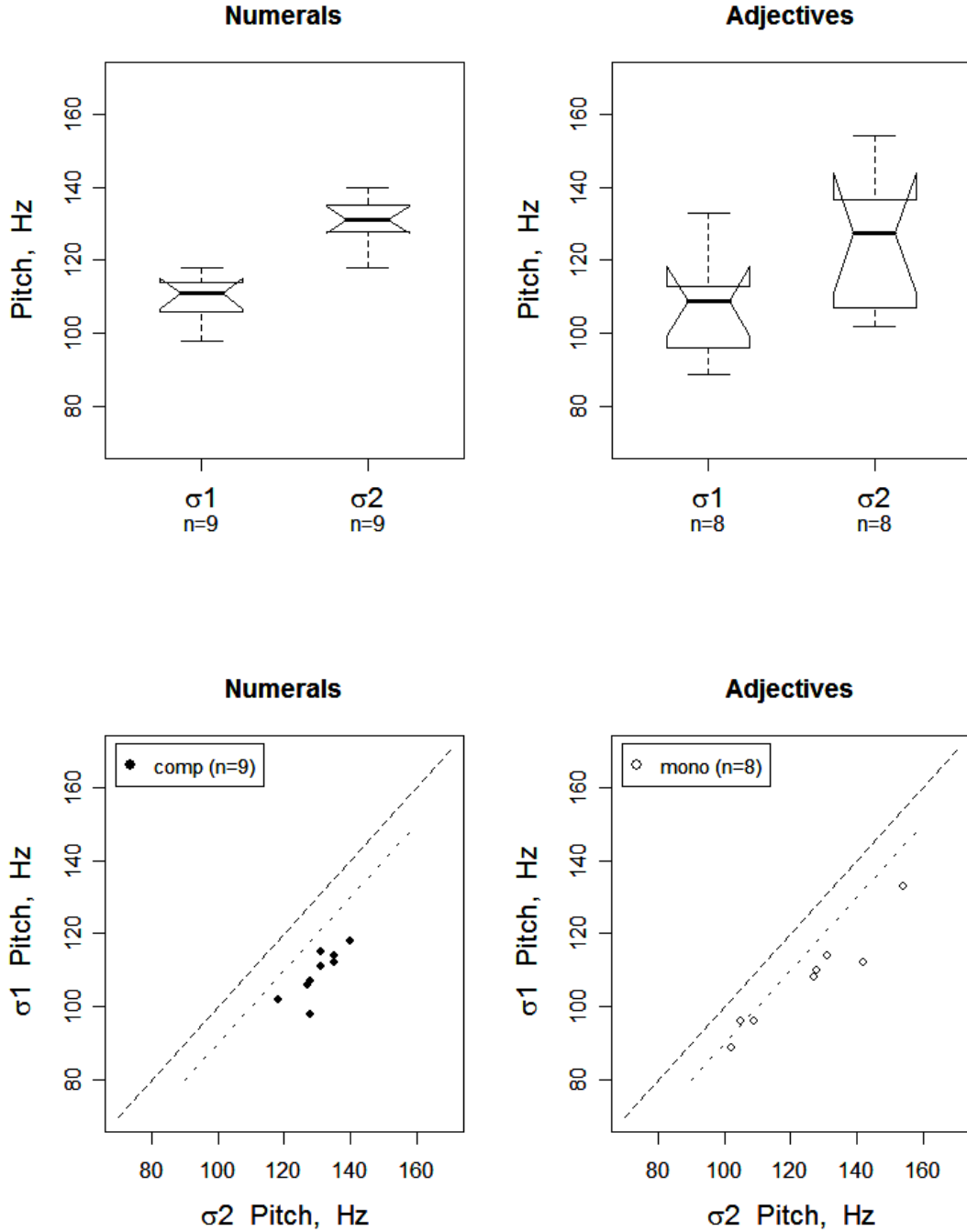
It is important to bear in mind that these box-and-whisker plots represent the distribution of values as if each syllable were independent of the other. Of course, this is not the case, since we are analyzing the properties of disyllabic words. Rather, each σ_1 value corresponds to one particular σ_2 value. Thus the graphical comparison of medians and 95% confidence intervals does not capture the entire story. In fact, for all we can tell from this plot, the highest pitch value measured on σ_1 – the outlier at 135 Hz – might correspond to the lowest pitch value measured on σ_2 – the bottom of the whisker, at 110 Hz. That would suggest the existence of a particular noun whose pitch was higher in σ_1 than in σ_2 , by 25 Hz. In fact, there is no such noun; pitch is always higher in σ_2 , as we already know from the distribution plot at left in Figure 2.23.

Thus in some respects the box-and-whisker plot provides more detailed information than the distribution plot, but in other respects it provides less. We can clearly see here that pitch values measured in σ_1 of BSh_03 nouns fall within a narrower range than do pitch values measured in σ_2 , and we can tell at a glance that the difference between median values for the two syllables is statistically significant. But we have no information about the pattern of contrasts in pitch across syllables in individual words.

In some cases, this does not really matter. For instance, in the box-and-whisker plot at left in Figure 2.24 below, there is no overlap at all in the ranges of pitch measurements in σ_1 and σ_2 of BSh_03 numerals. Since the distribution of measurements in the two syllables (considered independently) is so dramatically different, the paired values across syllables of individual words must also necessarily be different. No matter how measurements are paired up across syllables, pitch will always be higher in σ_2 than in σ_1 . At right, though, the distributions of pitch measurements in σ_1 and σ_2 of BSh_03 adjectives are not so distinct. The boxes – representing the interquartile range – and the notches – representing the 95% confidence interval about the median – both show considerable overlap. The relationship between pitch across syllables in individual words might show a number of patterns: higher-lower, lower-higher, or equal pitch.

Figure 2.24 Boxplots without and with overlap

BSh_03 / Isolation



In fact, though, for adjectives as well as for numerals, pitch is always higher in σ_2 – the stressed syllable – as illustrated in the distribution plots in the bottom part of Figure 2.24. The mean difference in pitch across syllables is 21 Hz for numerals, and 18 Hz for adjectives. The difference between the two lexical categories lies in the distribution of measurements; the greater range in values for both syllables of adjectives is illustrated by the box-and-whisker plot.

Thus when a box-and-whisker plot shows a clear difference across syllables, there really is a difference across syllables. But when a box-and-whisker plot does *not* show a difference across syllables, there still may be a difference. In these cases, a paired-sample t-test offers a more appropriate and more powerful means of examining potential correlations between acoustic patterns and perceived stress patterns.

2.5.1.2 Paired-sample t-tests

A t-test is used to compare two samples, evaluating them in terms of the default assumption, or “null hypothesis”. The null hypothesis states that the two samples are drawn from one common larger population and that there is no significant difference between their mean values; any apparent difference between the means simply reflects the randomness of the tokens selected from that larger pool. The null hypothesis is either accepted – meaning that the mean values are essentially the same – or it is rejected – meaning that the mean values are significantly different, and in fact represent samples from different populations.

Here, of course, the two samples are the acoustic measurements collected from the two syllables of the words in the data set. In this study, the more rigorous paired-sample t-test is appropriate; this test matches corresponding measurements from σ_1 and σ_2 of each individual word, calculates the difference in pitch (or pitch slope, or intensity, or vowel duration) for each word, and then determines whether the mean difference for the entire sample of words defines a statistically significant pattern.

As an example, the results of paired-sample t-tests for pitch measurements from BSh_03 numerals and adjectives are summarized in Table 2.7 below. The crucial test statistic is the p-value. For both lexical categories, $p < 0.05$. This means that the null hypothesis – that there is no difference in pitch across syllables – must be rejected. More precisely, the p-value indicates that there is a less than 5% probability that the mean pitch difference obtained could occur if there were no difference across syllables – if the pitch measurements on σ_1 and σ_2 simply represented the random variation which occurs within a single, normal population. For numerals, the probability of obtaining a mean difference of 21 Hz across syllables in our sample if there really were no difference at all in the larger population of BSh_03 numerals is only 3.287×10^{-7} to 1 – a very unlikely occurrence indeed.

Table 2.7 BSh_03 / Isolation / Pitch data: results of paired-sample t-tests (two-tailed)

Speaker	Lex Cat	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BSh_03	numerals	15.3106	8	3.287e-07	21	18	24
	adjectives	7.7709	7	0.0001097	18	12	23

The 95% confidence limits shown in the table above indicate that, if we repeatedly recorded samples of disyllabic numerals from speaker BSh_03 and calculated the mean difference in pitch across syllables, there is a 95% probability that the mean pitch difference for each sample would fall between 18 and 24 Hz for numerals, and between 12 and 23 Hz for adjectives. (The 95% confidence interval for adjectives is broader – and the p-value larger – because pitch showed greater variability, as discussed above.)

These paired-sample t-tests show that pitch is higher in σ_2 , and that the difference in pitch across syllables is statistically significant. Since both numerals and adjectives are stressed on σ_2 , this indicates a correlation between pitch and stress.

2.5.1.3 Welch t-tests

There is one instance in the analysis in which I use a Welch t-test to compare pitch measurements and intensity measurements in verbs of different morphological compositions. (See section 4.2.2.3.) As explained above, a t-test is a means of comparing two samples to determine if they represent variation within one common larger population, or if they represent two distinct populations. Here, the test is used to determine whether the two verbs types have the same acoustic patterns, or not.

There is no sense in which the measurements are paired, as described above, so the Welch test is the appropriate choice here.

The interpretation of these test results is similar to that described above for paired-sample t-tests.

2.5.2 Perceptual significance

In order to show that pitch, pitch slope, intensity, or vowel duration function as correlates of stress, a demonstration that a contrast across syllables is statistically significant – as discussed above – is a necessary but not sufficient condition. The contrast across syllables must also be perceptually significant.

For example, in nouns produced in isolation by speaker AR_05, the intrinsic correlation between intensity and vowel height can be controlled for by focusing on the subset of nouns with vowels of the same height in both syllables. As shown in Table 2.8 below, a paired-sample t-test confirms that the contrast in intensity across syllables for this [same.height] control group is statistically significant, with a p-value $\ll 0.05$.

Table 2.8 AR_05 / Nouns / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.442	31	0.1593	0.72	-0.30	1.7
same height	4.529	31	8.243e-05	1.8	1.0	2.6
higher.lower	5.144	23	3.266e-05	2.4	1.4	3.4

For this subset of 32 nouns, intensity on σ_2 is, on average, 1.8 dB greater than intensity on σ_1 , indicating a correlation with σ_2 stress. However, the lower 95% confidence limit means that, if one were to repeat the analysis with other samples of nouns recorded from this speaker, an intensity contrast deemed to be statistically significant might be as small as 1.0 dB.

The problem is that Lehiste (1970) has identified 1 dB as the difference limen (or “just-noticeable difference”) for intensity. Thus a mean intensity contrast which is statistically significant may be just at the threshold of perceptibility. If a listener can barely perceive the difference, can it be considered an effective means of conveying stress? I do not believe it can be. Furthermore, since 1.0 dB is a possible *mean* intensity difference, such a value entails that some of the intensity contrasts across syllables in other tokens in this hypothetical sample must be less than 1.0 dB – that is, less than what a listener can detect.

I thus consider perceptual significance to be crucial to the identification of *meaningful* acoustic correlates of stress. But establishing standards for perceptual significance is a tricky matter. Where difference limens have been determined through experimental phonetics, these can be considered a lower boundary of perceptibility. Beyond this, to say that a contrast across syllables is “weak” or “strong” is somewhat arbitrary. And for pitch slope, where experimental results are not available, I can rely only on my own perceptions of what is highly noticeable and what is not.

In sections 2.5.2.1 through 2.5.2.3 below, I discuss levels of perceptual significance for pitch, intensity, vowel duration, and pitch slope, respectively.

2.5.2.1 Perceptual significance for pitch

Lehiste (1970: 62-67) and Laver (1994: 451) cite a number of experimental studies designed to identify the minimal contrast in fundamental frequency / pitch that can be perceived by a listener. This just-noticeable difference, or difference limen, varies as a function of stimulus frequency (i.e., at low frequencies a small difference is noticeable; at high frequencies the difference must be larger to be noticeable), loudness, vowel duration, the listener's experience and expectations, and other factors. For frequencies within the range of a typical male voice (i.e., 80 Hz to 160 Hz), the discriminable difference was found to be on the order of ± 1 Hz. (Any difference smaller than that is sub-liminal.)

These experimental findings are directly relevant to the present study. If, for a set of words, the difference in pitch across syllables is found to be statistically significant but averages only ~ 1 Hz, then we would not consider this difference to be perceptually significant. Thus it would not constitute a meaningful correlate of stress. In fact, such a small difference would be subliminal – i.e., not even perceptible.

So, what if a statistically significant difference in pitch across syllables is, say, 3 Hz? What then? Such a difference is perceptible; but is it “perceptually significant”? Lehiste (1970: 79-80) refers to experimental studies in which listeners were able to discriminate tones in synthesized Thai words when the difference in

fundamental frequency of the tones was on the order of ± 5 Hz. (This was for frequency ranges of 120 to 150 Hz – i.e., within the range of the male voice).

Extrapolating from this study, for better or for worse I make the assumption that what is a meaningful difference for Thai tone is also a meaningful difference for Tibetan stress. That is, if a difference in pitch across syllables is statistically significant and is ~ 5 Hz, then I consider it to be perceptually significant. (So I would not consider the hypothetical 3 Hz difference questioned above to be perceptually significant, even though it is perceptible.) Still, I consider a statistically significant mean difference of 5 Hz to be only a “weak” and “unreliable” cue for stress; a mean pitch difference of 5 Hz entails that the sample includes a number of words with a pitch difference < 5 Hz, which is not perceptually significant.

When a statistically significant mean pitch difference across syllables is ~ 10 Hz, I call it a correlate of stress, usually without further qualification; a difference of ~ 20 Hz is a “strong” or “robust” correlate; a difference of ~ 40 Hz is a “dramatic” correlate.

When I refer to an acoustic parameter as a “reliable cue” for stress, I mean that the correlation between the acoustic pattern and the stress pattern is clear and consistent: e.g., a contrast in pitch across syllables is “robust”, and occurs in all tokens in the sample with very few exceptions. If a listener attended only to this one acoustic parameter, she would be able to readily identify the locus of stress.

While these terms and boundaries are somewhat arbitrary, they are nonetheless of some descriptive and comparative value within the confines of this study.

2.5.2.2 Perceptual significance for intensity

The just-noticeable difference for intensity has been identified as $\pm \sim 1$ dB, based on experiments conducted with synthesized vowels cited by Lehiste (1970: 115-116). In these studies, intensity was found to vary as a function of the intensity of the stimulus sound in the experiment, and also as a function of frequency. The difference limen of ± 1.0 dB corresponds to frequency levels and intensity levels relevant to natural speech.

Intensity also shows considerable variation as a function of vowel quality, as discussed in section 2.3.2.1. Controlling for this factor by comparing like vowels, Lehiste (1970: 121) measured an intensity difference of ~ 2 dB in a word-pair (in actual human speech) where the contrast was known to be perceptually significant.

Based on this, if a mean difference in intensity across syllables is statistically significant and is ~ 2 dB, then I consider it to be perceptually significant. However, I again consider such a mean difference to represent only a “weak” and “unreliable” correlate of stress, since the sample must necessarily include a number of words with an intensity contrast smaller than the mean – i.e., below the threshold for perceptual significance.

In Table 2.8 (page 83), we saw a subset of [same.height] nouns with a mean intensity difference of 1.8 dB and a lower 95% confidence interval of 1.0 dB. Even

though these differences are statistically significant ($p \ll 0.05$), I do not consider them to signal a meaningful correlation with stress, given the threshold values of ~ 1 dB for perceptibility and ~ 2 dB for perceptual significance.

Since intensity values are on a logarithmic scale, the interpretation of measurements is not fully intuitive. Laver (1994: 502) points out that, in terms of speech production, “a doubling of intensity corresponds to a difference of 3 dB”, and refers to other scholars whose work demonstrates that, in terms of speech perception, “a doubling of loudness ... corresponds to a rise in sound-level of approximately 10 dB”.

In this study, I consider a statistically significant mean difference in intensity of 5 dB to be “strong” or “robust”; a mean difference of 10 dB is “dramatic”. These are the kinds of differences observed in Balti verbs.

2.5.2.3 Perceptual significance for vowel duration

For vowel duration, too, Lehiste (1970: 10-13) summarizes the findings of experimental studies designed to determine the minimal contrast that can be perceived by a listener. As was the case with pitch and intensity, the difference limen varies with the stimulus; i.e., if the stimulus is short, a small change will be noticeable; if the stimulus is long, a difference must be longer to be noticeable. As Lehiste summarizes, “...in the range of durations of speech sounds – usually from 30 to about 300 msec – the just-noticeable differences in duration are between 10 and 40 msec” (*ibid*, p.13).

In this study, for both Balti speakers, nearly all vowels were between 50 and 200 msec in duration. For both Amdo speakers, nearly all vowels were between 30 and 200 msec in duration. Following Lehiste, the just-noticeable difference in duration for this range is probably about 10 to 30 msec – smaller for shorter vowels, and larger for longer vowels. In any event, a duration contrast must be larger than 10 msec in order to be perceptible.

But at what point can we consider a perceptible duration contrast to be “perceptually significant”? Lehiste suggests considering the ratio of duration measurements, rather than absolute values (1970: 11), and refers to experimental studies in which listeners were able to discriminate between [stressed] phonemically short and phonemically long vowels when the V/V₁ ratio was close to 50%, though with considerable variation from one language to another (1970: 33-34). Again, for better or for worse, I extrapolate from these studies and make the assumption that what is meaningful in distinguishing phonemic vowel length is also meaningful in distinguishing vowel length in unstressed *vs.* stressed syllables.

As demonstrated in Chapters 4 and 5, it is only in the case of verbs produced by speaker BM_01 that vowel duration seems to show a clear correlation with stress. Here, the duration ratio of unstressed : stressed syllables indeed comes close to 0.50.

2.5.2.4 Perceptual significance for pitch slope

So far as I know, there have been no previous studies quantifying pitch slope as an acoustic correlate of stress. Thus there are neither experimental nor heuristic

guidelines to help anticipate what order of pitch slope or pitch slope contrast might be perceptible or perceptually significant.

What I can say with certainty is that, for all of the Rebkong Amdo speakers I worked with – in Xining, in Rebkong, and in Kathmandu – the dramatic fall in pitch on the second syllable of nouns produced in isolation was an immediately striking and highly perceptible cue for stress. As noted in Chapter 1, for Ndzorge Amdo Tibetan, too, Sun (1986: 58) observed that the perceptual cues for σ_2 stress were “stronger articulatory force and a high-falling tune on the stressed syllable...”.

Based on my work here, it seems that a pitch slope of ~ -15 Hz/100msec – or a pitch slope difference across syllables of ~ 10 Hz/100msec – is perceptually significant.

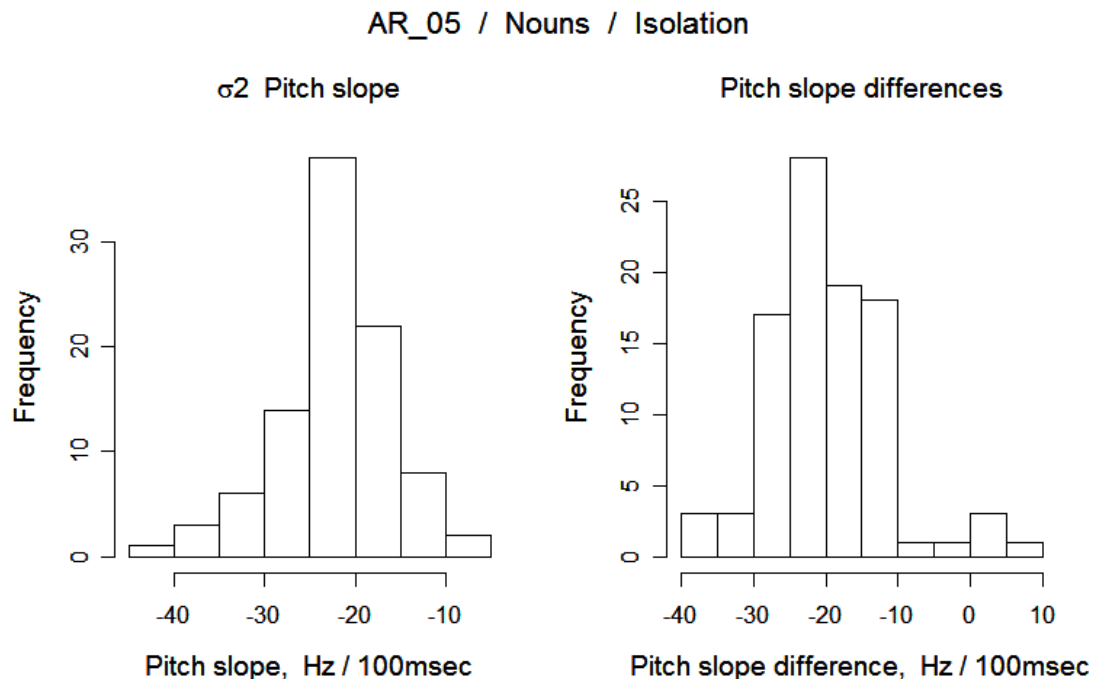
Median pitch slopes and pitch slope differences for nouns produced in isolation by speakers AR_04 and AR_05 are summarized in Table 2.9 below. The slopes produced by speaker AR_05 are the less steep of the two. Since I know from my own listening that such slopes are quite prominent, these values are helpful in defining the limits of perceptual significance. Thus I can be sure that a slope of at least ~ -20 Hz/100msec is perceptually significant.

Table 2.9 Perceptually significant pitch slopes: Rebkong Amdo nouns / isolation

Speaker	n	σ_1 median (Hz/100msec)	σ_2 median (Hz/100msec)	Difference ($\sigma_2 - \sigma_1$)
AR_04	93	-4	-33	-30
AR_05	94	-1	-21	-21

In fact, a consideration of the distribution of pitch slope measurements for AR_05 nouns suggests that the threshold of perceptual significance is a slope even more gentle than -20 Hz/100msec. The histograms in Figure 2.25 below show the frequency of pitch slope measurements and pitch slope differences. In the graph on the left, there is a jump in frequency at -15 Hz/100msec. That is, nouns with σ_2 pitch slopes between -10 and -15 Hz/100msec comprise less than 8% of the total, but nouns with σ_2 pitch slopes between -15 and -20 Hz/100msec suddenly comprise more than 20% of the total. Based on the frequency distribution, I suggest that these slope values between -15 and -20 Hz/100msec are within the range of the speaker's target for perceptually significant pitch slopes. That is, I consider a pitch slope of -15 Hz/100msec to be the threshold for perceptual significance.

Figure 2.25 Frequency of pitch slope measurements and pitch slope differences



Similarly, based on the graph at right above, I consider a pitch slope contrast across syllables of -10 Hz/100msec to be the limit of perceptual significance.

2.5.2.5 Summary of perceptual significance levels

Table 2.10 below summarizes the levels of perceptual significance which are relevant to evaluation of the acoustic parameters considered in this study.

Table 2.10 Perceptual significance levels

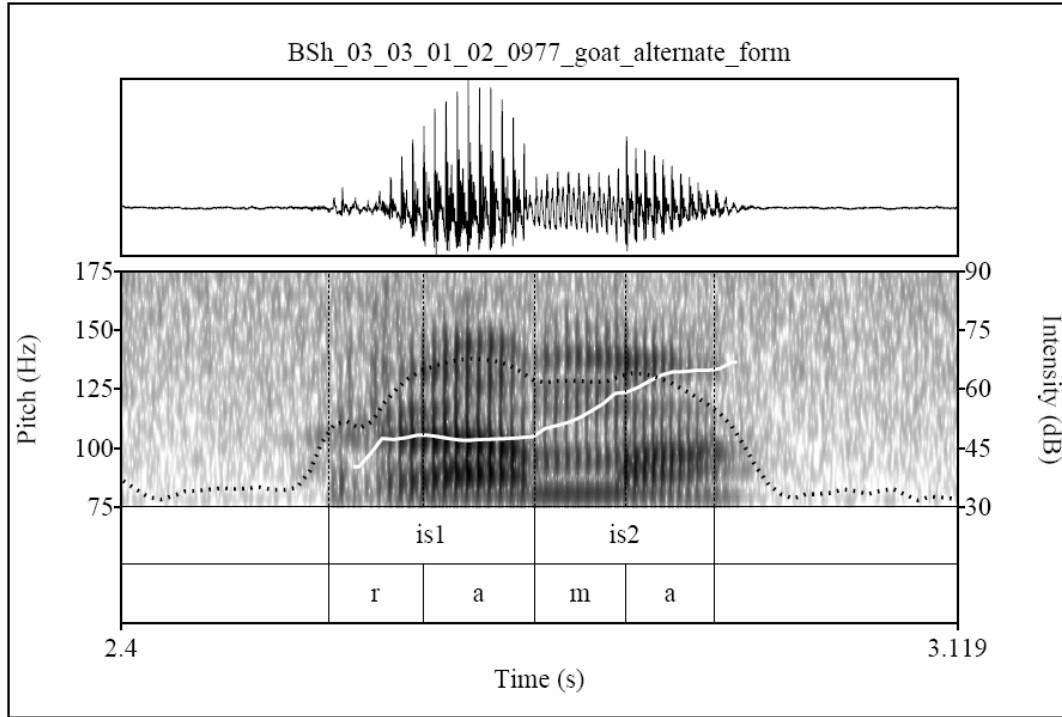
Acoustic parameter	Just-noticeable-difference	Weak difference	Meaningful difference
Pitch	1 Hz	5 Hz	10 Hz
Intensity	1 dB	2 dB	5 dB
V Duration	10-30 msec	–	V/V: 50%
Pitch slope	–	–	10 Hz/100msec

2.5.3 Contextual significance

In order to constitute a meaningful correlate of stress, a contrast in an acoustic parameter must be contextually significant, as well as statistically and perceptually significant. By contextually significant I mean that an acoustic pattern must be distinguishable from background acoustic patterns; it must stand out from its context.

This factor is particularly relevant in consideration of pitch slope data. For Balti speaker BSh_03, for instance, pitch slope patterns arise as a direct outcome of the contrasts in pitch across syllables. The noun [ra.'ma] *ra.ma* ‘goat’, shown in Figure 2.26 below, provides an illustration.

Figure 2.26 Pitch slope patterns arising from pitch patterns: BSh_03 [ra.'ma] ra.ma 'goat'



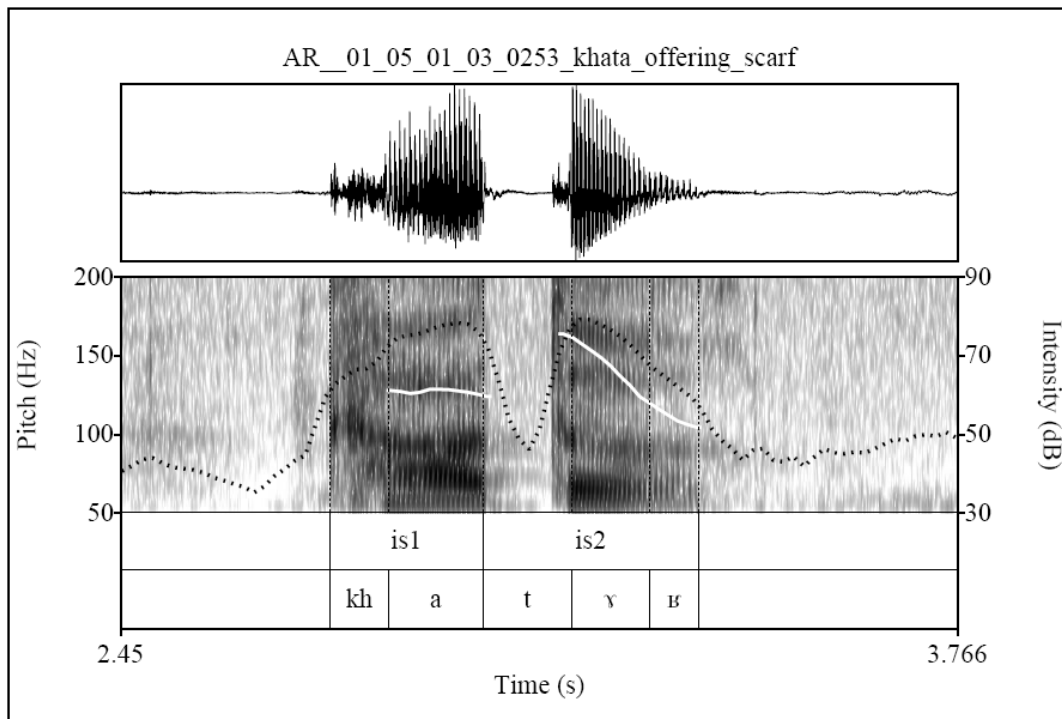
Measured over the medial 50% of the vowel, pitch is 104 Hz in σ_1 and 132 Hz in σ_2 . The 28 Hz pitch increase from σ_1 to σ_2 serves as a strong cue for stress. Pitch slope is 0.52 Hz/100msec in σ_1 , and 12 Hz/100msec in σ_2 , so slope is “more upward” in σ_2 . This upward slope in σ_2 has nothing at all to do with stress; it merely reflects – and is an extension of – the overall increase in pitch from σ_1 to σ_2 . It is not distinct from its acoustic context.

The relationship between pitch and pitch slope observed in this example also occurs in almost all of the other BSh_03 non-verbs, both in isolation and in the sentence frame. Since pitch slope is purely an epiphenomenon of another acoustic

feature, it lacks contextual significance. Thus for this speaker, it cannot be considered an acoustic correlate of stress.

In Rebkong Amdo nouns, where pitch slope is a robust correlate of stress, the opposite pattern pertains. The steep downward pitch slope in σ_2 is in sharp contrast to the overall rise in pitch from σ_1 to σ_2 ; it stands out against its context. This is illustrated by the pitch trace for the isolation form of the AR_05 noun [k^ha.'tʰɿɕ] *kha.btags* ‘khata, offering scarf’ shown in Figure 2.27 below.

Figure 2.27 Contextually significant pitch slope: AR_05 [k^ha.'tʰɿɕ] *kha.btags* ‘khata, offering scarf’



When Sun (1986: 58) refers to “ a high-falling tune ” as a perceptual cue for σ_2 stress in Ndzorge Amdo Tibetan, I believe he may be describing a similar

contextual contrast. My guess is that the “falling” pitch in σ_2 may be particularly salient precisely because it is juxtaposed against a “low” overall pitch level in σ_1 compared to σ_2 .

The methods of data collection and data analysis described in this chapter were used in the evaluation of stress in Balti and Rebkong Amdo Tibetan, as presented in Chapters 4 and 5, respectively.

3. Background on the dialects

In this chapter I provide background information relevant to my investigation of Balti and Rebkong Amdo Tibetan. Previous research is discussed in section 3.1, with a focus on descriptions of stress. Consonant and vowel inventories are presented in section 3.2, and observed syllable types in section 3.3. In sections 3.4 through 3.6 I describe some of the prosodic patterns I have noted in words of different lengths.

3.1 Previous research

My own analysis indicates that stress in both Balti and Rebkong Amdo falls on σ_2 of disyllabic nouns, adjectives, and numerals, and on σ_1 of disyllabic verbs. This characterization differs from previous studies of these dialects in drawing an intrinsic distinction between stress patterns in non-verbs *vs.* verbs. However, this same distinction has been reported for Zhongu Tibetan, a distinctly different dialect, with which I begin.

3.1.1 Zhongu Tibetan

Zhongu Tibetan is spoken in the Aba (Ngaba) Tibetan Autonomous Prefecture in Sichuan Province, China – geographically distal from the areas in which both Balti and Rebkong Amdo are spoken. Sun (2003) provides the first linguistic description of Zhongu, which he characterizes as an “obscure”, “peculiar”, and “idiosyncratic” variety of Tibetan: Zhongu is unusual in exhibiting extreme reduction of consonant

clusters, but without the concomitant development of phonemic tonal contrasts observed in the central Innovative dialects.¹⁶

One element of Zhongu that is relevant here is that Sun distinguishes between a σ_2 stress pattern in disyllabic nouns, and a σ_1 stress pattern in disyllabic verb complexes – the same pattern I observe and document in Rebkong Amdo and Balti. This has not been reported, so far as I know, for any other Tibetan dialects. In fact, it is unusual, cross-linguistically, for different lexical categories to exhibit different stress patterns.¹⁷

Sun explicitly marks “stress accent” in his transcriptions only in the section where he describes it (2003: 778). There, the example he provides of a noun in isolation is the compound [χse-ná] *gser.nag* (gold+black) ‘gold’. His glossary contains many of the same disyllabic words which I recorded in my study – including what I identify as both monomorphemic and compound forms – and he nowhere suggests that they are stressed anywhere but on σ_2 . His examples of disyllabic verbs include two Noun + Verbalizer constructions – [pá-lɛ] *bag.len* (bride+take) ‘to marry a wife’, and [ná-ɲɛ] *rna.nyan* (ear+listen) ‘to listen’ – but none which could be interpreted as an infinitive / verbal noun / citation form. In his glossary, for example,

¹⁶ In fact, Sun (2003: 797) considers Zhongu to be so distinct and divergent from other varieties of Tibetan that it should be regarded as “language-like”, and thus should not be categorized as a member of any other dialect group (such as Khams).

¹⁷ English is one of these rare cases, with a limited number of noun-verb minimal pairs such as 'convert - con'vert, 'import - im'port, 'rebel - re'bel, 'confines - con'fines, 'contract - con'tract, 'address - ad'dress, 'defect - de'fect, 'contrast - con'trast, and 'insight - in'cite.

the verb ‘carry’ is given simply as [k^hə] *khur*, rather than something like [k^hur.ba] *khur.ba* ‘to carry’ which I recorded in Balti.

Finally, Sun notes that “...stress placement is not always predictable and must sometimes be lexically marked.” He provides as illustration the minimal pair [mé-rə] ‘ideophone mimicking moving currents’ vs. [me-rə] *dmər.?* ‘to be red’.

3.1.2 Balti Tibetan

Previous studies of Balti include works by Grierson (1909), Read (1934), Sprigg (1966, 2002), and Bielmeier (1985a, 1988a)¹⁸.

Grierson (1909: 35) observed that “Balti does not appear to possess a marked system of tones. In this respect it agrees with Purik and Ladakhi.” He does not offer any remarks about stress.

Read (1934: 3) writes:

- (2.5) The Tibetan language, and consequently all its dialects, is really a collection of independent short syllables. However many syllables the word may contain, each one must be given equal emphasis and never be cut short. The word *polo* (a ball) is not “poll-o”, but “po-lo”; likewise in the word *gor-gyal-chan* (disobedient) equal emphasis must be placed on each syllable.

¹⁸ Read’s grammar is based on the variety of Balti spoken in the village of Khapalu. Sprigg, at the time he conducted research on Balti, was not able to travel to Baltistan, and so worked in Rawalpindi, Pakistan with a young man also from Khapalu; this young man, Zakir Hussein Baltistani, was also familiar with the Skardu variety of Balti. The material in Bielmeier’s book (1985a) is based on a story told by a Balti story teller from a village near Khapalu. His narration was recorded on tape by Klaus Sagaster. Bielmeier analyzed this tape with the assistance of M. Iqbal, a native speaker of the Skardu variety, resident in Islamabad at that time. Bielmeier later worked (mainly in Skardu) with several Balti speakers from different areas of Baltistan. (Bielmeier p.c., 2008)

It is not unreasonable to describe Tibetan as a collection of short syllables. However, Read is alone in asserting that all syllables receive “equal emphasis”, and I do not see how this characterization can be reconciled with those of other scholars.¹⁹

Indeed, Sprigg (1966, 2002) and Bielmeier (1985a, 1988a:47) are quite clear about their observations of stress: both report that stress falls on the final syllable of a word, regardless of lexical category (and regardless of word length). The σ_2 stress I observe in disyllabic non-verbs is explicitly accounted for by this description. The σ_1 stress I observe in disyllabic verbs is accounted for – only partially, as I explain below – by their analysis of certain morphemes as non-stress-bearing, causing stress to shift to the preceding syllable. Sprigg’s characterization (2002: 4-5) is as follows:

- (2.6) The stress-bearing syllable of a word should be taken to be the final syllable, unless (*i*) some other syllable has been marked as stress-bearing, by a grave accent, or (*ii*) the final syllable is one of the following (non-stress-bearing) suffixes: -can (some words only), ci (-s)-e, -en, -i, -ing, -la, -mo (except bruk-mo), -mo (pron. ngo), na, -pa (II), -pa (VI), -pho, -phu, -phun, -po, -re, -sang, -shik, -tsa (II ‘some’), or -tu, in which case it is the syllable immediately before this suffix that is the stress-bearing syllable (unless that syllable is also one of these non-stress-bearing syllables; but that is rarely so).

The type (*i*) exceptions above are lexical. Aside from words borrowed from Urdu, there are probably not more than a few dozen of these in the entire dictionary. Most easily found are words beginning with **a**²⁰ (mostly interjections or borrowings) and deictics beginning with **de-** or **e-**. Other autochthonous disyllabic words marked

¹⁹ Perhaps Read was attending to vowel quality, rather than stress? But that is mere conjecture.

²⁰ Sprigg presents his dictionary entries in terms of morphological composition (transcribed at the phonemic level). I show his forms here in bold type to distinguish them from Written Tibetan forms (which I show in italics) and phonetic forms (which I show in plain type, using the IPA).

explicitly with σ_1 stress include **bràa-shing** 'brad.shing? 'rake', **dàaman** *lda.man*²¹ 'small drum', **kùru** *ku.rug* 'colt of an ass', **phàDing** 'dried apricots', and **tsùt-mo** *?mo* 'doll'. I did not happen to elicit any of these words in my own work.²²

The type (ii) exceptions listed above are morphological and/or phonological. These “non-stress-bearing suffixes” from (2.2) are listed in Table 3.1 below. The “Meaning / function” in the third column here is Sprigg’s definition (except where enclosed in square brackets), as provided on the page indicated in the fourth column. The lower case letters in the second column correspond to the order in which these “suffixes” are listed by Sprigg in (2.6) above. The Roman numerals in the first column reflect my own groupings, as discussed below.

I consider many of Sprigg’s type (ii) exceptions to be clitics (which I, too, have found to be unstressed, as discussed in section 3.5.3): oblique case markers / clause connectors in I, quantifiers in II, participial endings in III, and others in IV (including (o), which seems to be lexically unique).

²¹ According to Bielmeier (p.c., 2008), this word is a loan, and the spelling provided in Jaeschke ([1881] 1958: 289a) is secondary.

²² Sprigg’s dictionary also includes a number of trisyllabic lexical exceptions which are stressed on σ_1 : e.g., **gà-ba-r-met** 'nowhere' and **kàaDo-shing** *kha.to.shing* 'cross used in ancient torture'. A handful of tri- and quadri-syllabic words – compounds and reduplicated forms – are marked as stressed on a medial syllable – e.g., **daltùmur** *bdar.rdo.?* 'soft smooth stone used for honing' – but most are not, and so we must assume Sprigg found them to be stressed on the final syllable. Longer expressions bear no indication of stress placement, either: e.g. **khi-skor mi-skor bya** 'to convince through using several people, enforce through using several people to plead', **tsoks-na tsoks** 'alike, exact, same', and **gzhu-bu khur-pa** 'man who carries a bow; best man at a wedding'.

Table 3.1 Sprigg’s list of non-stress-bearing suffixes (2002: 5)

	“Suffix”	Meaning / function	Page
	e. -i	<i>gen</i> of, -’s, -s’	77
	f. -ing	in	78
I.	g. -la	at, on, to, for	95
	j. -na	from, with, and; when, and then	119
	u. -tu	<i>loc</i> , equivalent to ‘-wards’	172
	b. ci	any, a certain, some	46
II.	q. -re	equivalent to -ful, -ish	136
	t. -tsa (II, ‘some’)	[some, about ?]	168
III.	c. (-s)-e	<i>participial</i> , equivalent to -ing	55
	d. -en	<i>pres part</i> , equivalent to -ing	55
	a. -can	‘having’, ‘-able’	39
IV.	r. -sang	<i>particle</i> also, too, even	148
	s. -shik	<i>imp</i> please do it	152
	o. -phun	[only in <i>bila-phun</i> ‘butterfly’]	131, 30
	h. -mo (except <i>bruk-mo</i>)	female, equivalent to <i>-ess</i>	117, 33
	i. -mo (pron. [-ngo])	[see -mo (I), -mo (II), -mo (III)]	117
V.	m. -pho	adjectival suffix (pron. [-fo])	129
	p. -po	suffix designating concrete nouns	132
	n. -phu	young one, of animals	130
VI.	k. -pa (II)	<i>inf</i> , equivalent to ‘to’	126
	l. -pa (VI)	past	126

Of particular significance are the nominal / adjectival suffixes in group V, and the verbal suffix (k) in group VI. All of these are very common in the language, and I consider them in some detail here in order to make two points: First, where Sprigg sees non-stress-bearing suffixes on nouns and adjectives (group V), I see a contrast in vowel height across syllables, whose intrinsic acoustic properties account for the perceived stress pattern. Second, where Sprigg sees a non-stress-bearing suffix on verbs (VI.k), I see a different stress pattern altogether, by taking into consideration verbs of a different morphological composition.

Exception (V.h) here is evidently Sprigg’s suffix **-mo** (I). He identifies **bruk-mò** (which seems to occur only in the quadrisyllabic compound **lha-mò bruk-mò**

‘fabulous golden-haired woman, wife of Gesar’; 2002:101) as an anomaly stressed on σ_2 . This word is presented in clear contrast to a list of other examples in which he considers **-mo** to indeed be unstressed. This list includes the disyllabic words **baks-mo** *bag.ma* ‘bride’, **dak-mo** *dag.mo* ‘housekeeper, person in charge of all food, etc.’ (related to the adjective *dag.po* ‘clean’), and **lyaks-mo** *legs.po* ‘good, clean, well, nice’. For these words – which I elicited in my own study – I found that the acoustic correlates of stress provide mixed cues to the listener: intensity and vowel duration are more prominent on σ_1 , while pitch is more prominent on σ_2 . This can be attributed to the contrast in vowel height across syllables; as discussed in section 2.3, low vowels are intrinsically longer and have a higher intensity than high vowels, while high vowels have a higher pitch than low vowels. It seems likely that the intrinsic effects lending prominence to σ_1 are what Sprigg noted as stress; this, in turn, may have led him to identify σ_2 as a non-stress-bearing suffix.

In fact, the intrinsic effects of vowel quality contrast may underlie all of the other nominal / adjectival “non-stress-bearing suffixes” in Table 3.1. Nearly all of them end in [-o], and since [a] is the most common vowel in Tibetan, the result is a high frequency of words with an [a.o] / [lower.higher] vowel height contrast.

Sprigg’s account of exception (V.i) is phonologically based; it seems to refer to **-mo** (I), **-mo** (II), and **-mo** (III) whenever they are pronounced as [-ngo], which occurs when the preceding sound is a vowel. (As indicated above, **-mo** (I) marks ‘female, equivalent to *-ess*’; **-mo** (II) is an adjectival suffix, while **-mo** (III) is a ‘substantive’ (i.e., nominal) suffix.) In the few words I recorded with a [-ngo] suffix,

however, stress in almost all cases fell on σ_2 , rather than on σ_1 . These included some of Sprigg's examples with **-mo** (I): **bu-mo** *bu.mo* 'girl, daughter' and **nò-mo** *no.mu* 'younger sister' were pronounced, in my recordings, as [bu.'ŋo] and [no.'ŋo]²³. For **bya-mo** *bya.mo* 'hen', stress correlates were mixed; again, contrasts in vowel height across syllables likely play a role.

Exceptions (V.m) and (V.p) all have a male or masculine sense; for these, picking out specific entries in Sprigg's dictionary is a little confusing. **-pho** and **pho** (both pronounced [-fo]) are identified as an adjectival suffix and a substantive, respectively (2002: 129). For examples, the reader is directed to entries for **mar-pho** *dmar.po* 'red', **ser-pho** *ser.po* 'yellow', and **graks-pho** *grang.mo* ? 'cold'?²⁴.

Contrasts in vowel height may play a role here, as described above. This may also be the case in other illustrations of **-pho** and **pho** offered by Sprigg, including **baks-pho** *bag.po* 'bridegroom', **bya-pho** *bya.po* 'cockerel, cock', and **rgyal-pho** *rgyal.po* 'king', and also in his examples with **-po**: **daks-po** *bdag.po* 'husband', **gong-po** ? (no gloss provided), **mak-po** *mag.pa* 'bridegroom', **rgyal-po** *rgyal.po* 'king', **rtsan-po** *rtsan.po* 'prayer answered by god'. An exception with a higher vowel in σ_1 is **rtsis-po** *rtsis.po* 'number, account'.

²³ As noted in the preceding paragraph, Sprigg identified **lha-mò** *lha.mo* 'goddess' as a lexical exception to his proposed phonological alternation, pronounced [l̥a.'ŋo] rather than [l̥'a.ŋo]; I, too, found this word to be stressed on σ_2 .

²⁴ Bielmaner (p.c. 2008) points out that the adjective 'cold' in Balti is [graks.'mo]. It is not clear what Sprigg is referring to with his form **graks-pho**. Also, for 'king', Bielmaner has observed only [gyal.'pho], and not [gyal.'po] in Balti.

Finally, exception (V.n) **-phu** is a diminutive suffix. As examples, Sprigg offers **bilà-phu** *byi.la.phru* ‘kitten’, **bya-phu** *bya.phru* ‘small chicken’, **dren-phu** *dren.phru* ‘bear cub’, **gong-phu** *gong.phru* ‘small snow pheasant’, **gri-phu** *gri.phru* ‘small knife’, **lu-phù-u** *lug.phru* ‘lamb’, and **ITing-phu** *lteng.phru* ‘small shallow well’. I recorded several words with the related diminutive suffix *-phrug* from speakers AR_04 and AR_05: [çə.ʋriɣ] and [çə.riɣ] *bya.phrug* ‘baby bird, chick’; [lə.riɣ] *lu.phrug* ‘kitten’; and [ʔlɔ̃.ɰiɣ] *glang.phrug* ‘baby male ox’. None of these examples could be segmented reliably, and so were not useful for quantitative acoustic analysis, but the acoustic information can be looked at all the same. In all four tokens, intensity is higher on σ_1 and pitch is higher on σ_2 . These contradictory cues are exactly what one might predict, since in all four tokens there is a [lower.higher] vowel height contrast. (The stem which is the historical source of this diminutive is unstressed when it occurs in the Balti noun [p^hru.^ʋɸra] *phru.phra* ‘children’; in this case, both pitch and intensity are higher on σ_2 .)

In all of the above, then, Sprigg accounts for his perception of prominence on σ_1 by suggesting that σ_2 is a non-stress-bearing suffix. For a large majority of these cases, I suggest that his perception of prominence on σ_1 may be attributable instead to the intrinsic effects of a contrast in vowel height across syllables.

This hypothesis can be tested by considering disyllabic nouns and adjectives with these type (V) suffixes, but in which there is *not* a contrast in vowel height. My

own data includes only a few such cases, which are listed in Table 3.2 below. As shown, when there is no contrast in vowel height, there are no acoustic cues suggestive of a shift in stress to σ_1 . Instead, in these control cases, pitch and intensity are higher on σ_2 , indicating σ_2 stress.²⁵

Table 3.2 Balti [o.o] nouns and adjectives

Speaker	Gloss	WT	IPA	Δ Pitch (Hz)	Δ Int (dB)	Δ Pitch slope (Hz/100msec)
BM_01	girl	<i>bu.mo</i>	bo.'ŋo	25	0	-43
	thick, fat	<i>sbom.po</i>	bom.'bo	2	1	-39
BSh_03	thick, fat	<i>sbom.po</i>	bom.'bo	19	4	9.8
	rare	<i>dkon.po ?</i>	ʂkøn. 'mo	17	1	21
	high	<i>mthon.mo</i>	thon.'mo	21	2	22

An entirely different situation is presented by exception (VI.k) in Table 3.1 – the verbal suffix “-**pa** (II) *inf*, equivalent to ‘to’”.

Once Sprigg interpreted the other suffixes in Table 3.1 as non-stress-bearing, he might analogously have interpreted this **-pa** the same way, which is one means of accounting for the σ_1 stress which is so strongly and consistently perceived and acoustically manifested in verbs. In this case, though, I think the σ_1 stress on verbs is

²⁵ I do not include a comparison of vowel duration across syllables in these examples because, as discussed in section 2.3.2.2, a number of interfering factors come into play: (a) vowels are intrinsically longer in open syllables than in closed syllables; (b) vowels are intrinsically longer in the final syllable of an utterance. Since these effects cannot be factored out, it is not possible to isolate a potential correlation between vowel duration and stress, and so such a comparison would be meaningless.

The pitch slope contrasts presented here are also not relevant to stress, since they are not contextually significant. This is discussed in sections 2.5.3 and 4.3.

intrinsic and distinctive. I am led to this conclusion by considering – in addition to infinitive / citation verb forms like [tʰaq.pa] *'thag.pa* ‘to grind’ and [χnot.pa] *gnod.pa* ‘to harm’ – verbs with other morphological structures. These include N+Vblzr forms like [lut.taŋ] *lud.btang* (manure+send) ‘to spread manure’ and [sna.bya] *rna.bya* (ear+do) ‘to listen’, as well as N + V compounds like [zgo.tɕukʰ]²⁶ *sgo.bcug* (door+close) ‘to close a door’ and [za:n.za] *zan.za* (food+eat) ‘to eat food’.

In these cases the morphemes in σ_2 are richer in phonetic and semantic content than Sprigg’s set of non-stress-bearing suffixes, and are not counted among them. Thus σ_1 stress cannot be explained as a shift in placement from σ_2 . Vowel quality does not play a role either; in all of these words, pitch and intensity are higher in σ_1 than in σ_2 regardless of contrasts in vowel quality across syllables. Thus I believe that the stress pattern observed in disyllabic verbs of all structural types is most efficiently accounted for by simply saying that Balti verbs are stressed on σ_1 , in contrast to the non-verbs.

Both Sprigg and Bielmeier noted a handful of minimal pairs in Balti in which verbs stressed on σ_1 contrast with nouns stressed on σ_2 – e.g., Sprigg’s verbal noun / infinitive / citation form [rgɔs.pa] *dgos.pa* ‘needing / to need’ vs. the noun [rgɔs.'pa]

²⁶ The final aspiration here is quite unusual, but this is indeed how the speaker produced this token.

dgos.pa ‘need’, and Bielmeier’s [‘xlat.pa] **glad.pa*²⁷ ‘to be tired’ vs. [xlat.’pa]
klad.pa / *glad.pa* ‘brain’. Based on such minimal pairs, Bielmeier describes stress as
“marginally phonemic”. (Though of course, as noted by Pike (1948), even when
stress is not phonemic, it is still part of what a speaker knows about her language.)

Finally, for disyllabic nouns, both authors also reported perceiving a higher
pitch on the second syllable than on the first. This is exactly what I confirm through
acoustic measurements, as discussed in Chapter 4.

For other Western Archaic Dialects, such as Sham and Ladakhi, Bielmeier
again reports σ_2 stress (1988a: 48).

Regarding Purik, Zemp (2006) notes that previous studies of the language did
not address stress, and that he himself did not take much note of it during his
fieldwork, as it never seemed prominent or important. In his preliminary analysis,
Zemp seems to have analyzed stress patterns in disyllabic words as they occurred in
recorded stories – i.e., not in isolation or in a controlled setting – and he measured
only intensity as a potential acoustic correlate. He seems to say that stress on
disyllabic words is variable, and that intensity measurements were not always
consistent with perceived stress. He does note, though, that most disyllabic nouns
have a higher pitch on σ_2 .

²⁷ This reconstructed Proto-Tibetan form was provided by Roland Bielmeier (p.c. 2008).

3.1.3 Amdo Tibetan

In Rebkong Amdo, just as in Balti, my research shows that disyllabic nouns, adjectives, and numerals are stressed on σ_2 , while disyllabic verbs are stressed on σ_1 .

de Roerich (1958) is a grammatical sketch of the very same variety of Amdo that I consider here, with a number of transcribed and translated narratives. de Roerich's discussion of phonology includes no mention of stress at all; he says only that the dialect is non-tonal (1958: 29).

For Ndzorge Amdo, Sun (1986: 58) reports that stress is fixed and non-phonemic, falling on the last syllable of polysyllabic words regardless of their length. He observes that the perceptual cues for stress are “stronger articulatory force and a high-falling tune on the stressed syllable, regardless of whether the onset of this syllable is voiced or voiceless.”

In Themchen Amdo, as described by Haller (2004: 28), disyllabic words are usually stressed on the second syllable. Haller does not mention whether or not stress is phonemically contrastive. There is no discussion of stress patterns on longer words.

Neither Sun (1986) nor Haller mentions a contrast between stress patterns in verbs *vs.* non-verbs for Amdo. It may be that verbs and non-verbs exhibit the same stress patterns. Alternatively, it may be that they did not particularly attend to the potential for a distinction here. Since it is not very common for languages to exhibit different stress patterns in different lexical categories, one might not think to check this. Furthermore, while disyllabic nouns and adjectives are ubiquitous and easily

isolable, eliciting disyllabic forms of verbs requires a deliberate effort; if one were not specifically looking for a stress pattern in verbs, one might not notice it.

Makley et al. (1999) make no mention of stress in their initial phonological overview of the Labrang variety of Amdo.

3.2 Consonant and vowel inventories

Consonant and vowel inventories for Balti and Rebkong Amdo are presented in sections 3.2.1 and 3.2.2 below, respectively. These inventories should be considered preliminary – they are limited to the segments and clusters encountered in the words I elicited during this focused investigation. A much more careful evaluation of Balti and Rebkong Amdo phonetics and phonology would be required to develop authoritative inventories of consonant and vowel phonemes.²⁸

3.2.1 Balti inventories

Table 3.3 below provides an inventory of Balti simplex consonants – i.e., those that cannot be analyzed as clusters.

²⁸ Or – perhaps more useful – an inventory of syllable onsets and rhymes.

Table 3.3 Balti: inventory of simplex consonants²⁹

bilab	ap-dent	retro	pal	post-pal	velar	uvular
p	t	ʈ			k	q
p ^h	t ^h	ʈ ^h			k ^h	
b	d	ɖ			g	
	ts	tʂ	tɕ			
	dz	ɖʂ	dʒ			
	s	ʂ	ɕ	ç	x	χ
	z	ʐ	ʒ		ɣ	ʁ
m	n		ny		ŋ	
	l	ɭ				
	ɭ	ɬ				
	r	ɽ				
w			y		w	

Consonant clusters are presented in Table 3.4. This inventory is almost certainly not complete, as it is limited to onset clusters encountered in the words in my study. (See also Sprigg 2002: 5-13.)

²⁹ As noted in section 2.1.3, in my transcriptions, I use [ny] rather than IPA [ɲ], and [y] rather than IPA [j] in order to avoid potential confusion with the Wylie (1957) transliteration system.

Table 3.4 Balti: Inventory of consonant clusters (preliminary)

f-	ɸtɕ-							
v-	vd-							
d-	dzɾ-							
s-	sp-	st-	sk-	sm-	sn-	sŋ-		
z-	zb-	zg-						
ʂ-	ʂt-	ʂk-						
ç-	çt-	çl-						
x-	xt-	xs-	xl-					
χ-	χm-	χn-	χs-	χ ^l -	χp-			
m-	mb-							
n-	nd-	nz-						
ŋ-	ŋg-							
r-	rd-	rg-	rɣ-					
ʁ-	ʁz-							
ʈ-	ʈs-							
-r	pr-	br-	t ^h ɾ-	kr-	tʂɾ-	ndr-	dzɾ-	str-
-l	bl-	çl-	xl-	χl-				
-y	py-	spy-	by-	ʂky-	gy-	rgy-		

Vowels are shown in Table 3.5.

Table 3.5 Balti: vowel inventory

i	u
e	o
a	

3.2.2 Rebkong Amdo inventories

Table 3.6 provides an inventory of simplex consonants observed in Rebkong Amdo. The voiced bilabial [b] sometimes occurs word-initially as [mb-] or [ʙ].

Aspiration in the velar [k^h] is often so strong as to sound like [k^x].

Table 3.6 Rebkong Amdo: inventory of simplex consonants

bilab	lab-dent	ap-dent	retro	pal	post-pal	velar	uvular	glottal
p		t	ʈ	c		k		
p ^h		t ^h				k ^h		
b		d				g		
		ts	tʂ	tɕ				
		ts ^h	tʂ ^h	tɕ ^h				
		dz	dʐ	dʑ				
	f	s	ʂ	ɕ	ç	x	χ	h
		s ^h	ʂ ^h	ɕ ^h	ç ^h			
	v	z	ʐ	ʑ		ɣ	ʁ	
m		n		ny		ŋ		
m̥		n̥				ŋ̥		
		l	l̥					
		ʎ	ʎ̥					
		r	r̥					
w	ʋ			y		w		

Table 3.7 below provides an inventory of consonant clusters noted in my data.

Clusters beginning in [h^o-] are discussed in section 3.3.2

Table 3.7 Rebkong Amdo: Inventory of consonant clusters

k-	ks-					
d-	dzy-					
ç-	çt-	çs-	çts-	çtç-		
x-	xt-	xts-	xs-	xtç-	xcy-	xk-
χ-	χk-					
h-	^{hə} d-	^{hə} g-	^{hə} z-	^{hə} dʒ-	^{hə} l-	^{hə} ʃ-
n-	nt-	nd-	nts-	ndz-	ntç-	
ŋ-	ŋk-	ŋg-				

The vowel inventory is shown in Table 3.8 below.

Table 3.8 Rebkong Amdo: vowel inventory

i	u
e	ɤ, o
a	

Reduced vowel allophones – [ɪ], [ɨ], [ɤ], and [ʌ] – are very common in Rebkong Amdo, and lip-rounding is not very pronounced even in [o], and [u]. The speakers I worked with in Xining, Rebkong, and Kathmandu all spoke with their jaws somewhat closed and still.

The vowel inventory also includes the diphthongs [ai] and [oi]. These are only found in σ1, and only in words in which σ1 has an alveolar coda in the Written Tibetan form. Examples include [^hmãĩ.'k^xɤŋ] *smān.khang* ‘hospital’, [^{hə}lai.'pa]

klad.pa / *glad.pa* ‘brain’, [tsɔiŋ.'ma] *btson.pa* ‘prisoner’, [χɔim.'bo] *dpon.po* ‘official, chief’, and [ntɕai.'mo] *rgyal.mo* ‘queen’.

In natural speech, vowels may be lengthened and nasalized to compensate for the deletion of a final nasal coda.

3.3 The syllable template

The syllable templates which I observed in Balti and Rebkong Amdo are described in sections 3.3.1 and 3.3.2, respectively. The inventories include both heavy and light syllables. (Of course, syllable weight is not relevant to the placement of stress; stress is fixed, according to lexical category.) In both dialects, complex consonant clusters are more varied and more frequent in onset position than in coda position.

3.3.1 Balti syllable types

The syllable templates which I have observed in my Balti data are listed in Table 3.9 below. It is possible that other types exist, but I have not encountered them thus far. And there are certainly other types of open syllables with long vowels – e.g., CVV or CCCVV – but all of these are cases where a coda has been deleted and compensatory lengthening – and sometimes also nasalization – has occurred.

Table 3.9 Balti / Possible syllable types

	σ_1	σ_2
	V-	*
open	CV-	-CV
	CCV-	-CCV
closed	CVC-	-CVC
	CCVC-	-CCVC
	CCCVC-	-CCCVC
	CVCC-	-CVCC
	CVVC-	*

* Not attested in my data.

Examples of the different types of open syllables are provided in Table 3.10.

Table 3.10 Balti / Examples of open syllables

Template	Examples				
	Lex cat	Syllable	Word	WT	Gloss
V	N	ʔo-	ʔo.'spis	'o.spri	cream
CV	N	tɕa-	tɕa.'p ^h e	ja.phye	tea flour
		-ga	ɕtʌz.'ga	rta.sga	horse saddle
	V	*	*	*	*
CCV	N	-za	'za:n.za	zan.za	to eat food
		sta-	sta.'re	sta.re	axe
	V	*	*	*	*
		zgo -	'zgo.tɕuk	sgo.bcug	to close a door
sna-	'sna.bya	snya.bya	to listen		

* Not attested in my data.

Examples of the different types of closed syllables are provided in Table 3.11.

Table 3.11 Balti / Examples of closed syllables

Template	Examples				
	Lex cat	Syllable	Word	WT	Gloss
CVC	N	mak-	mak.'pa	<i>mag.pa</i>	bridegroom
		-tek	ʃa.'tek	?	frog
	V	ʒik- *	'ʒik.p ^h a *	<i>'jigs.pa</i> *	to be afraid *
CCVC	N	zbuk-	zbuk.'pa	<i>sbud.pa</i>	bellows
		-rlat	ŋgo.'rlat	<i>mgo.glad</i>	brain, mind
	V	stun- - ɸtɕøʃ	'stun.ma 'zan.ɸtɕøʃ	<i>srung.ma</i> <i>zan.bcos</i>	to keep, protect to cook food
CCCVC	N	stren-	stren.'ma	<i>sran.ma</i>	pea
		-striŋ	bu.'striŋ	<i>bu.sring</i>	woman
	V	strãŋ- *	'strãŋ.ma *	<i>srang.ma</i> *	to straighten *
CVCC	N	-t ^h aɣʃ	gøn.'t ^h aɣʃ	<i>dgong.?</i>	evening
	A	tsokʃ-	tsokʃ.'tsokʃ	<i>tsogs.tsogs</i>	same, alike
	V	* *	* *	* *	* *
CVVC	N	t ^h aũŋ-	t ^h aũŋ.'boŋ	<i>thang.bong</i>	donkey
	A	lear-	lear.'mo	<i>legs.po</i>	good
	V	za:n-	'za:n.za	<i>zan.za</i>	to eat food

* Not attested in my data.

3.3.2 Rebkong Amdo syllable types

The syllable types I have observed in Rebkong Amdo are presented in Table 3.12 below.

Table 3.12 Rebkong Amdo / Possible syllable types

	$\sigma 1$	$\sigma 2$
open	V-	*
	CV-	-CV
	CCV-	-CCV
	CCVV-	*
closed	CVC-	-CVC
	CCVC-	-CCVC
	CCCVC-	-CCCVC
	CVCC-	-CVCC
	CVVC-	*

* Not attested in my data.

Examples of the different types of open syllables are provided in Table 3.13 below. (As discussed in chapter 2, I recorded, segmented, and analyzed only a few Rebkong Amdo verbs. All of these had a Noun + Verbalizer morphological structure; none were citation forms, in which it would have been more likely to encounter an open $\sigma 2$. Their absence here reflects the limitations of my sample.)

Table 3.13 Rebkong Amdo / Examples of open syllables

Template	Examples				
	Lex cat	Syllable	Word	WT	Gloss
V	N	o-	o.'ma	'o.ma	milk
CV	N	t ^h o-	t ^h o.'pa	thod.pa	forehead, skull
		-go	^ə vuŋ.'go	pus.mgo	knee
	V	vu - *	'vu.dzəp *	wu.brgyab *	to shoot a gun *
CCV	N	xkɣ- *	xkɣ.'pa *	skud.pa *	thread *
		V	nta- *	'nta.hən *	mda'. 'phen *
CVV	N	ʂai- *	ʂai.'ma *	sran.ma *	soybean, legume *
		V	* *	* *	* *
CCVV	N	çtai- *	çtai.'mo *	ltad.mo *	show, spectacle *
		V	* *	* *	* *

Examples of closed syllables are provided in Table 3.14 below.

Table 3.14 Rebkong Amdo / Examples of closed syllables

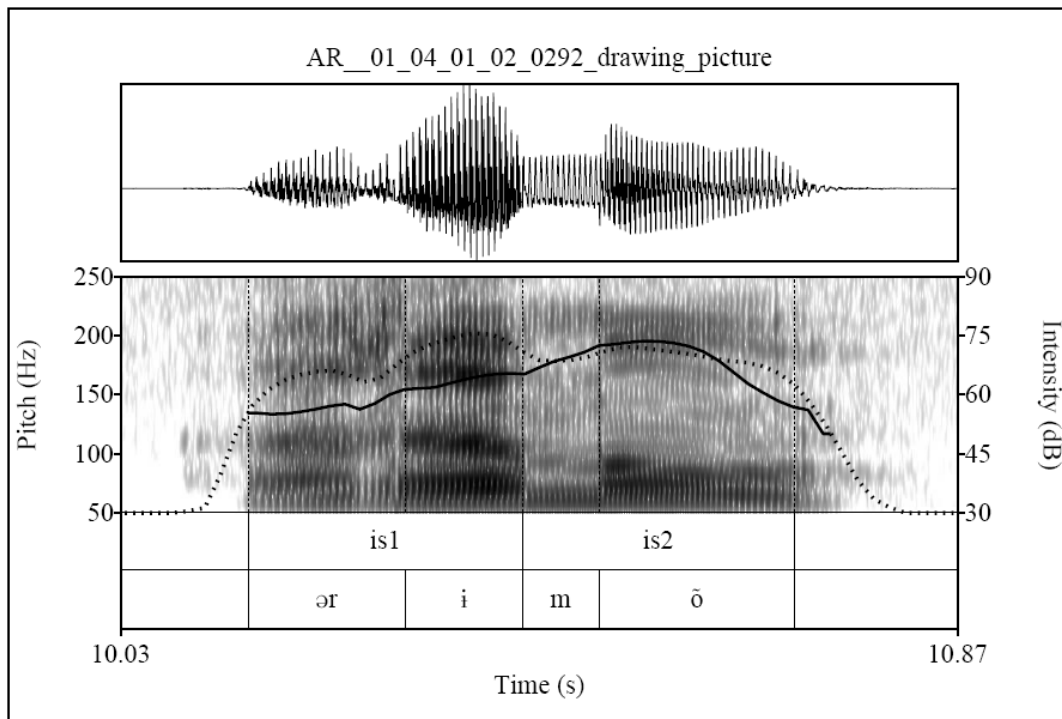
Template	Examples				
	Lex cat	Syllable	Word	WT	Gloss
CVC	N	lex-	lex.'ka	<i>las.ka</i>	work
		-kɪl	p ^h ar.'kɪl	<i>pha.skad</i>	father tongue
	V	k ^h ʌŋ-	'k ^h ʌŋ.bʌp	<i>gangs.babs / kha.ba.babs ?</i>	to snow
		-hɛn	'nta.hɛn	<i>mda'. 'phen</i>	to shoot an arrow
CCVC	N	xtsam-	xtsam.'ba	<i>rtsam.pa</i>	tsampa
		- hmin	xtsa.'hmin	<i>rtswa.sman</i>	medicinal plant
	V	^{hə} nʌm-	^{hə} nʌm.bʌv	<i>gnam.babs</i>	to rain
		- tɛyɣv	ʒɛ̃.tɛɣv	<i>lan.brgyab</i>	to answer
CCVCC	N	xkarn-	xkarn.'da	<i>skar.zla</i>	shooting star
	V	*	*	*	*
CVVC	N	χɔim-	χɔim.'bo	<i>dpon.po</i>	official, chief
	A	ts ^h ɔim-	ts ^h ɔim.'bu	<i>tshon.po</i>	fat
	V	*	*	*	*
CCVVC	N	hmain-	hmain.'tsa	<i>sman.rtswa</i>	medicinal plant
	A	χkɔim-	χkɔim.'bu	<i>dkon.po</i>	rare
	V	*	*	*	*

* Not attested in my data.

In a number of words I transcribe an [ʔ-] or [hʔ-] at the beginning of the first syllable. These sounds are visible in the waveform and spectrogram, but I do not consider them to constitute an independent syllable.³⁰ In some cases these fragments precede a voiced syllable onset; here, I think they represent nothing more than early voicing which is part of careful enunciation. An example of this is the noun [ʔri.'mõ] *ri.mo* ‘drawing, picture’, shown in Figure 3.1 below.

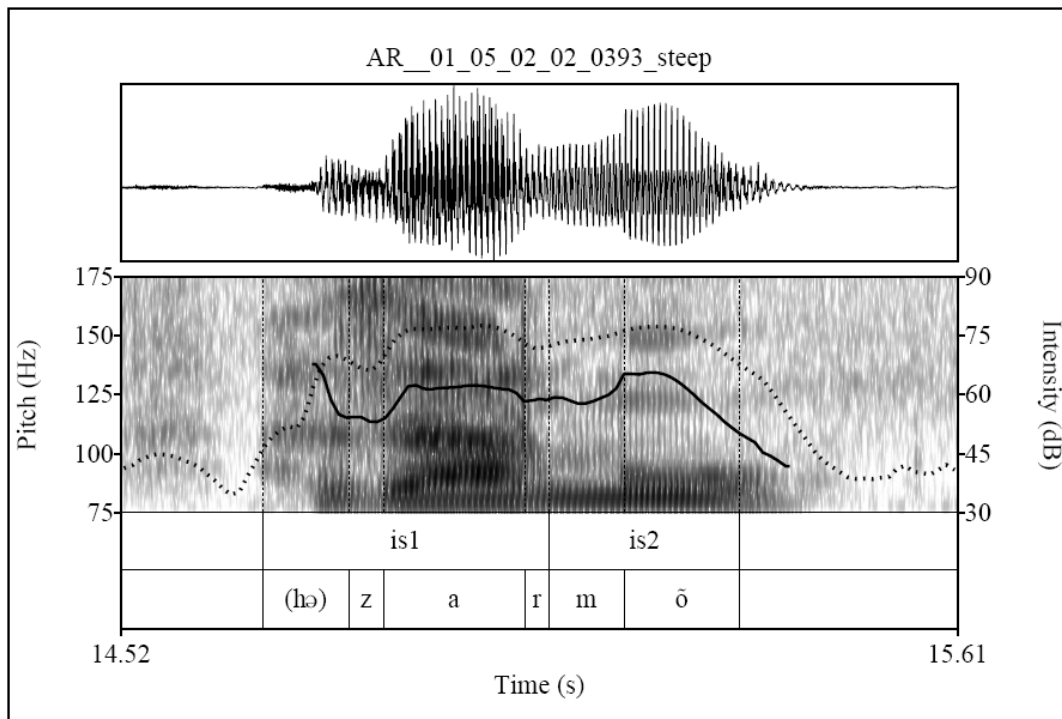
³⁰ These are reminiscent of the sesquyllables described by Matisoff (1973).

Figure 3.1 [-] preceding a voiced $\sigma 1$ onset: AR_04 [ʔri.'mõ] *ri.mo* ‘drawing, picture’



In other cases these extrametrical bits correspond to – and seem to be relicts of – a consonant cluster in Written Tibetan. An example is the adjective [ʔzɑɾ.'mõ] *gzar.mo* ‘steep’, shown in Figure 3.2 below. The onset sequence [ʔz-] corresponds to the WT cluster *gz-*.

Figure 3.2 [h^o-] corresponding to a WT cluster: AR_05 [h^ozar.'mō] *gzar.mo* 'steep'



When such words are produced within the sentence frame, these initial vocalic fragments often disappear.

3.4 Monosyllabic words

In the Innovative dialects of Tibetan, there is a seemingly endless supply of monosyllabic words which are minimal pairs contrasting only in tone. Monosyllables thus provide ubiquitous and robust evidence that these dialects are tonal.

When we turn to monosyllabic words in Balti and Rebkong Amdo, such evidence is virtually non-existent. In fact, I have encountered only one anomalous

minimal pair in Balti³¹, and none at all in Rebkong Amdo. The tonal distinctions that are a defining characteristic of the Innovative dialects simply do not occur here.

3.4.1 Comparison to an Innovative dialect

To illustrate, we can compare data from Balti and Rebkong Amdo to examples from Tokpe Gola Tibetan, an Innovative dialect spoken in northeastern Nepal which is prosodically quite similar to Lhasa Tibetan and which I have investigated in some detail. As shown in Table 3.15 below, the Tokpe Gola words for ‘arrow’ and ‘horse’ constitute a minimal pair, contrasting only in terms of tone. The Balti and Rebkong Amdo cognates of these same words differ instead in their segmental content. (This table is arranged iconically in terms of geographic location, with Balti to the “west” and Rebkong Amdo to the “east”.)

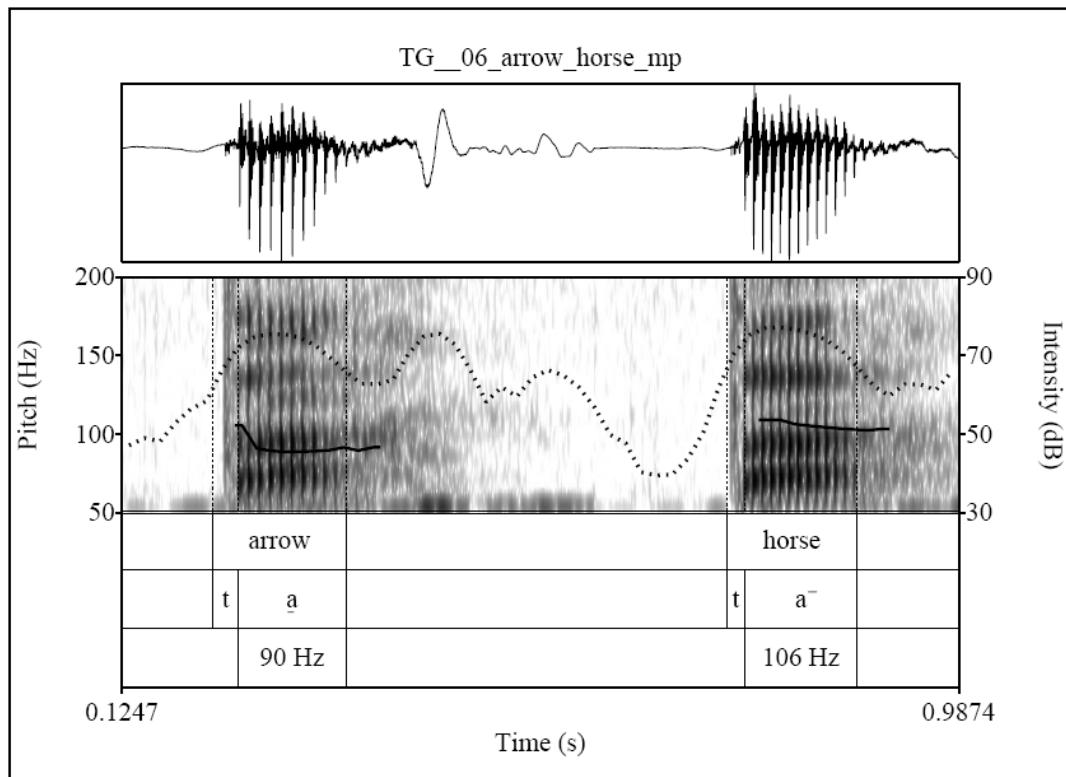
Table 3.15 ‘arrow’, ‘horse’: Non-tonal vs. tonal dialects

Gloss	Written Tibetan	Balti	Tokpe Gola	Rebkong Amdo
arrow	<i>mda'</i>	nda	t̪a	nda
horse	<i>rta</i>	ʎta	tā	xta

³¹ One of my Balti language consultants provided me with a single surprising minimal pair: the words [ra□] *WT* ‘goat’ and [ra□] *WT* ‘pen, enclosure’. These words appear to contrast in pitch, with average fundamental frequency measured over the vowels as 127 Hz and 110 Hz, respectively. My skepticism about this example is shared by Nicolas Tournadre (p.c., 2008), and such contrasts have not been reported elsewhere in the literature. My consultant pointed out to me that [ra□] ‘goat’ is a short form of [ra.'ma]. As he was enthused about having recently begun to study Written Tibetan (and thus tonal Central Tibetan), I think it is possible he may have had in mind the fact that such contracted forms often take on a high tone in the tonal dialects. The high pitch of [ra□] is not consistent with the same speaker’s low pitch in [ra.'ma] in Figure 4.25, in Chapter 4.

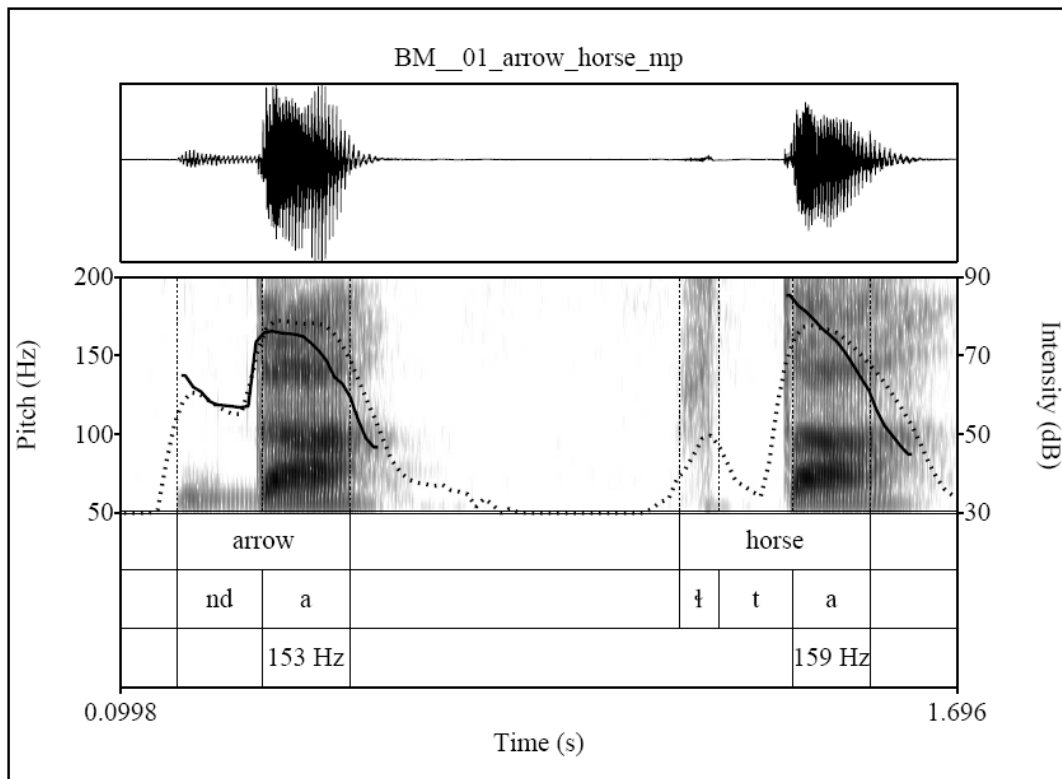
Pitch traces for the two Tokpe Gola words are illustrated in Figure 3.3 below. The speaker maintains a level pitch across the vowels in both words; the average pitch over the vowel in [tā] *mda* 'arrow' is 16 Hz lower than the average pitch over the vowel in [tā] *rta* 'horse'. This pitch contrast is perceptually significant, as discussed in section 2.5.2. (16 Hz may not appear to be a dramatic difference in Figure 3.3, but I chose this scale of 50 Hz - 200 Hz to be consistent across all the illustrations below.)

Figure 3.3 Tokpe Gola (TG_06) / Pitch traces for [t̩] *mda* ‘arrow’ and [t̩] *rta* ‘horse’



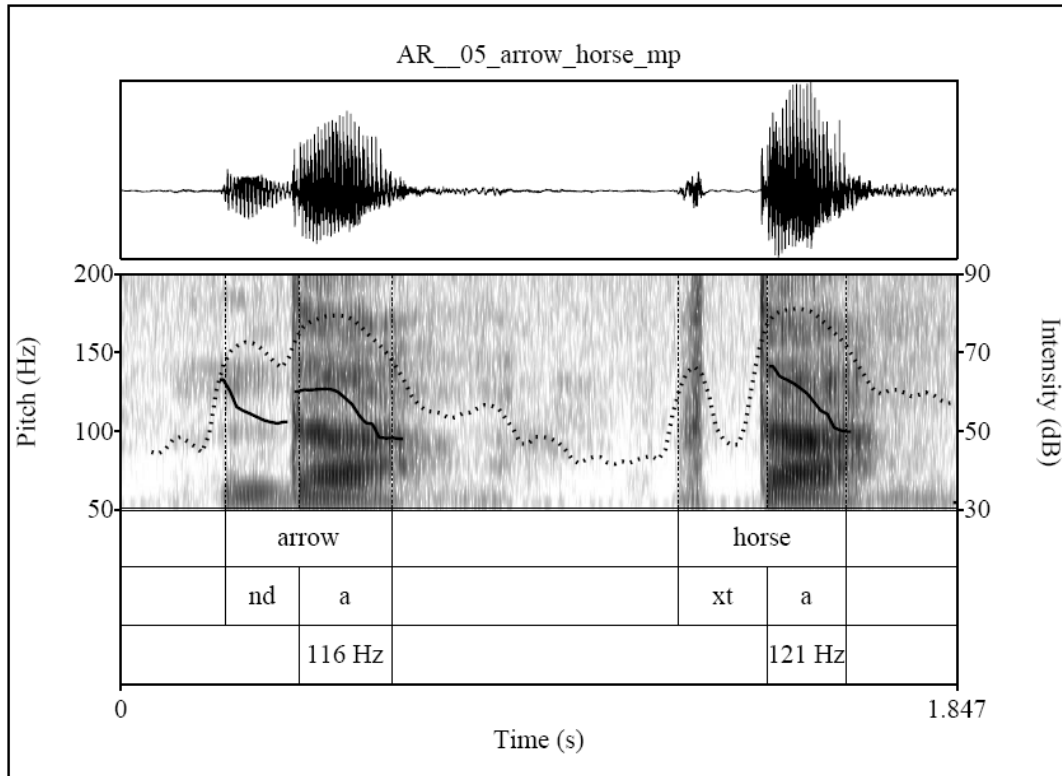
For vowels in the same two words in Balti, however, the difference in average pitch over the vowels in [nda] *mda* ‘arrow’ and [ṭa] *rta* ‘horse’ is only 6 Hz, as illustrated in Figure 3.4 below. Both words show the same falling pitch contour over the vowel; the slightly higher pitch at the beginning of ‘horse’ is likely a result of the transition from the tense vocal folds associated with the high frequency of the fricative and with the voiceless stop onset.

Figure 3.4 Balti (BM_01) / Pitch traces for [nda] *mda* ‘arrow’ and [ɬta] *rta* ‘horse’



Almost exactly the same patterns are observed in the pitch traces of the same words in Rebkong Amdo, as shown in Figure 3.5 below. Here, the difference in average pitch is only 5 Hz.

Figure 3.5 Rebkong Amdo (AR_05) / Pitch traces for [nda] *mda* ‘arrow’ and [xta] *rta* ‘horse’



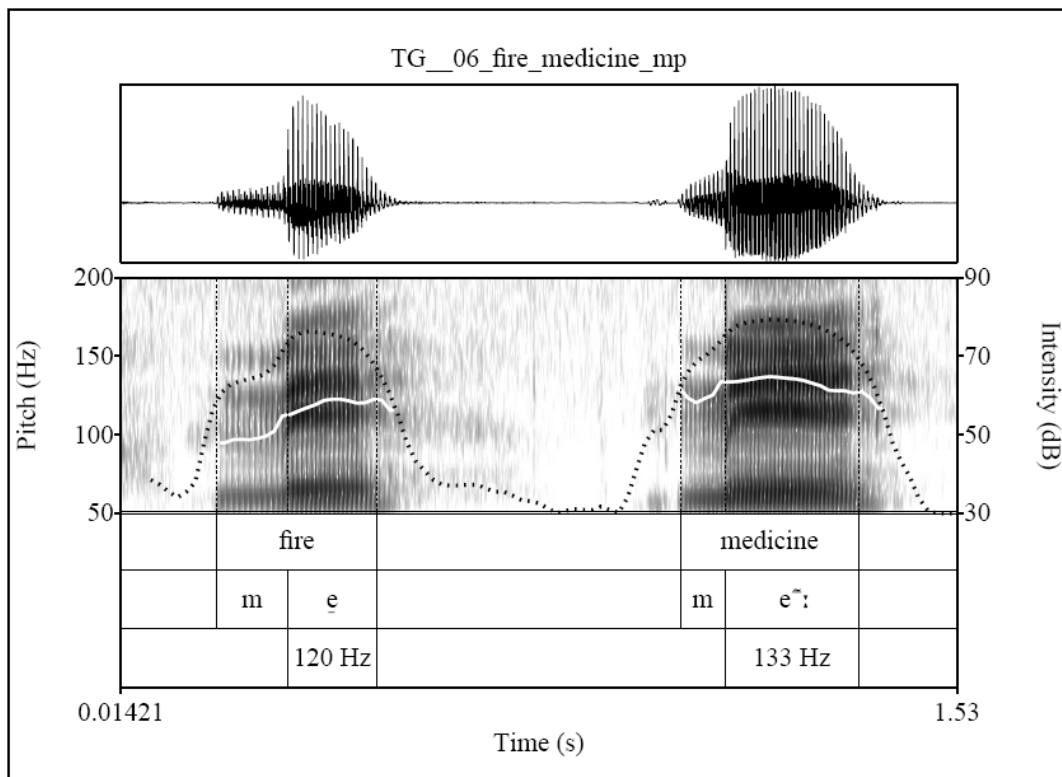
A similar example is the case of *me* ‘fire’ and *sman* ‘medicine’. As shown in Table 3.16 below, the Balti and Rebkong Amdo forms preserve traces of the segmental content and contrasts of 7th century Written Tibetan, while in Tokpe Gola the distinction is made by the suprasegmental properties of tone and vowel duration.

Table 3.16 ‘fire’, ‘medicine’: Non-tonal vs. tonal dialects

Gloss	Written Tibetan	Balti	Tokpe Gola	Rebkong Amdo
fire	<i>me</i>	me	m̄e	nyi
medicine	<i>sman</i>	sman	m̄ẽ:	h _a men

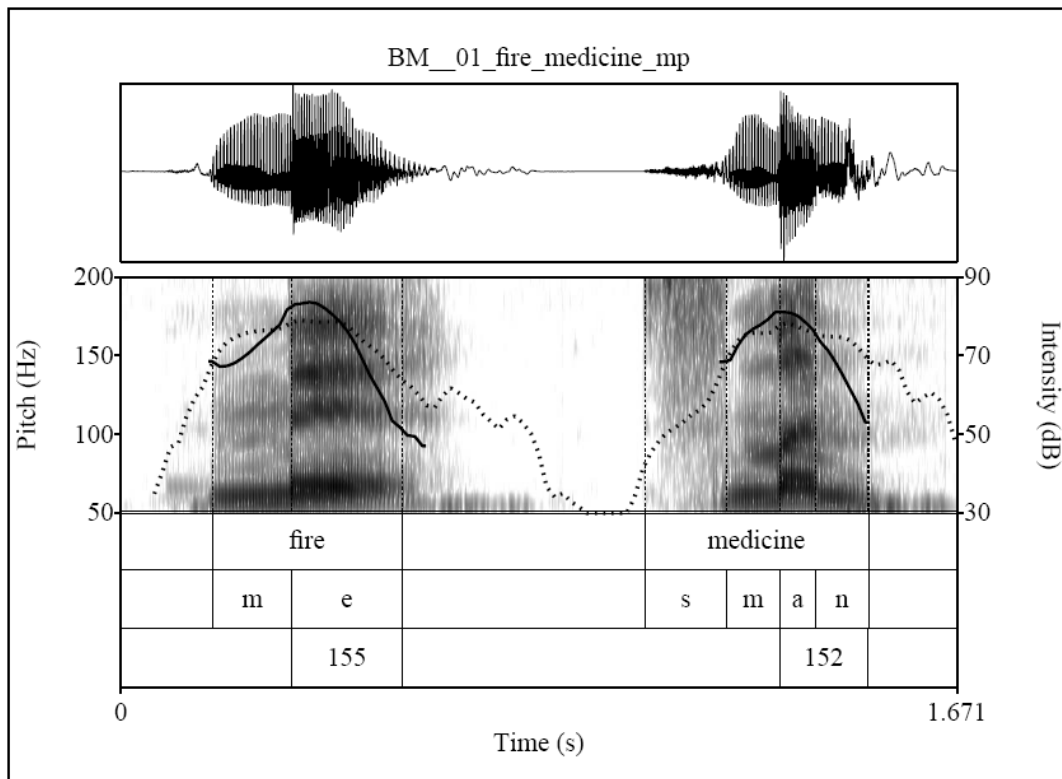
Pitch traces for the Tokpe Gola pair are shown in Figure 3.6 below. The difference in average pitch across the vowels in the two words is 13 Hz, a perceptually significant contrast.

Figure 3.6 Tokpe Gola (TG_06) / Pitch traces for [mɛ] *me* ‘fire’ and [mẽ:] *sman* ‘medicine’



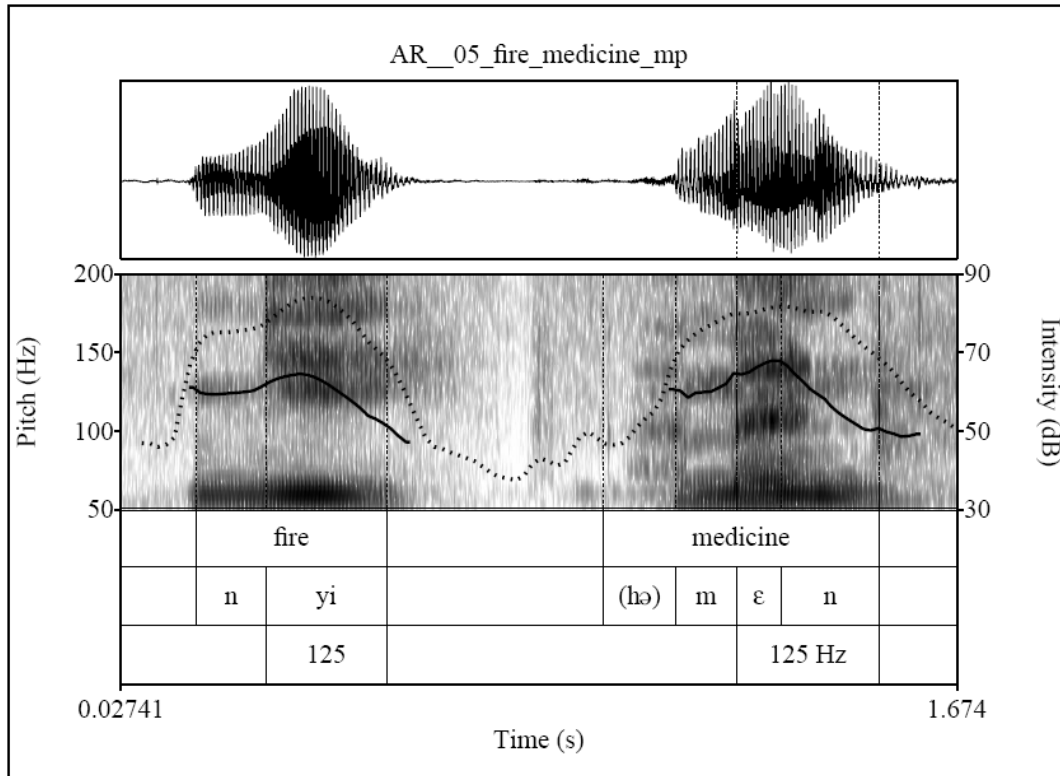
Pitch traces for the two Balti cognates are compared in Figure 3.7 below. The rhymes show similar curves, and the average pitch values differ by only 3 Hz. To my ears, there is no readily detectable difference in pitch in the two words.

Figure 3.7 Balti (BM_01) / Pitch traces for [me] *me* ‘fire’ and [sman] *sman* ‘medicine’



This is again the case in Rebkong Amdo, as shown in Figure 3.8 below; here, too, the average pitch values over the rhyme are exactly the same.

Figure 3.8 Rebkong Amdo (AR_05) / Pitch traces for [nyi] *me* ‘fire’ and [h^omen] *sman* ‘medicine’

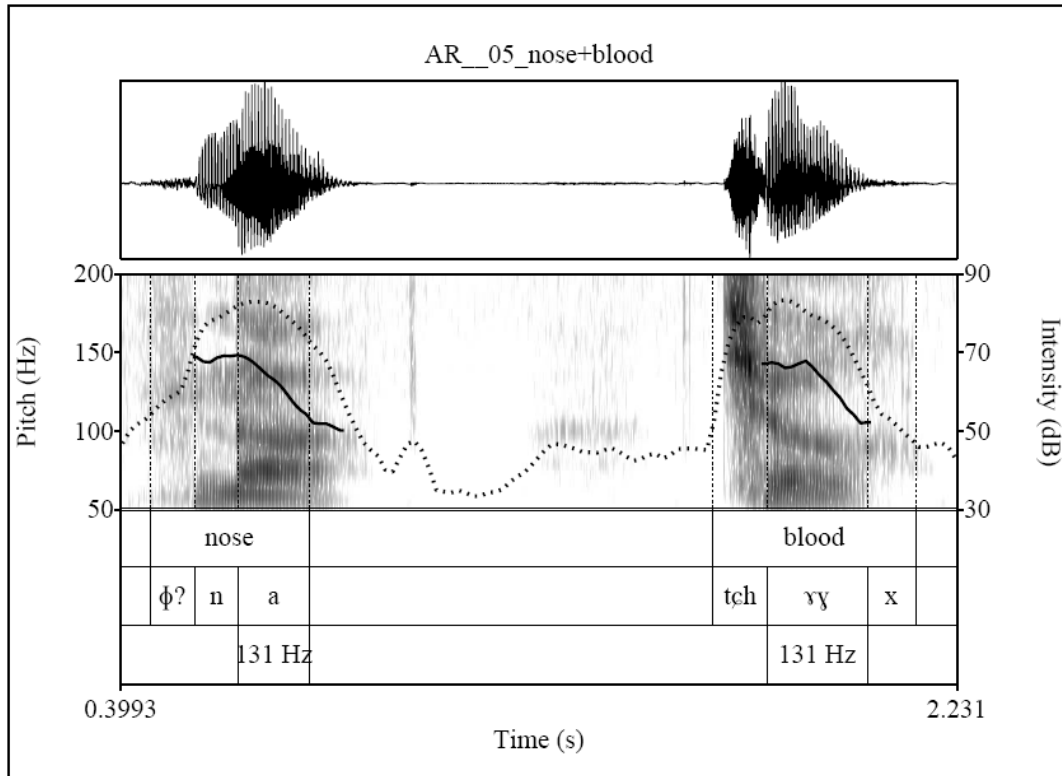


3.4.2 Pitch in compound formation

In her consideration of monosyllabic words in the Aba (Nga.ba) variety of Amdo, Huang (1995:45) says that “[t]one of any kind is completely lacking ...syllables of all types invariably carry a high falling pitch 53.”³² This high falling pitch is just what we have observed for Rebkong Amdo (and Balti) above, and it is illustrated again by the two words [ɸna] *sna* ‘nose’ and [tɕ^hɣ] *khrag* ‘blood’ (as produced by speaker AR_05) spliced together in Figure 3.9 below. For both words, the pitch pattern is more or less the same as in the preceding examples.

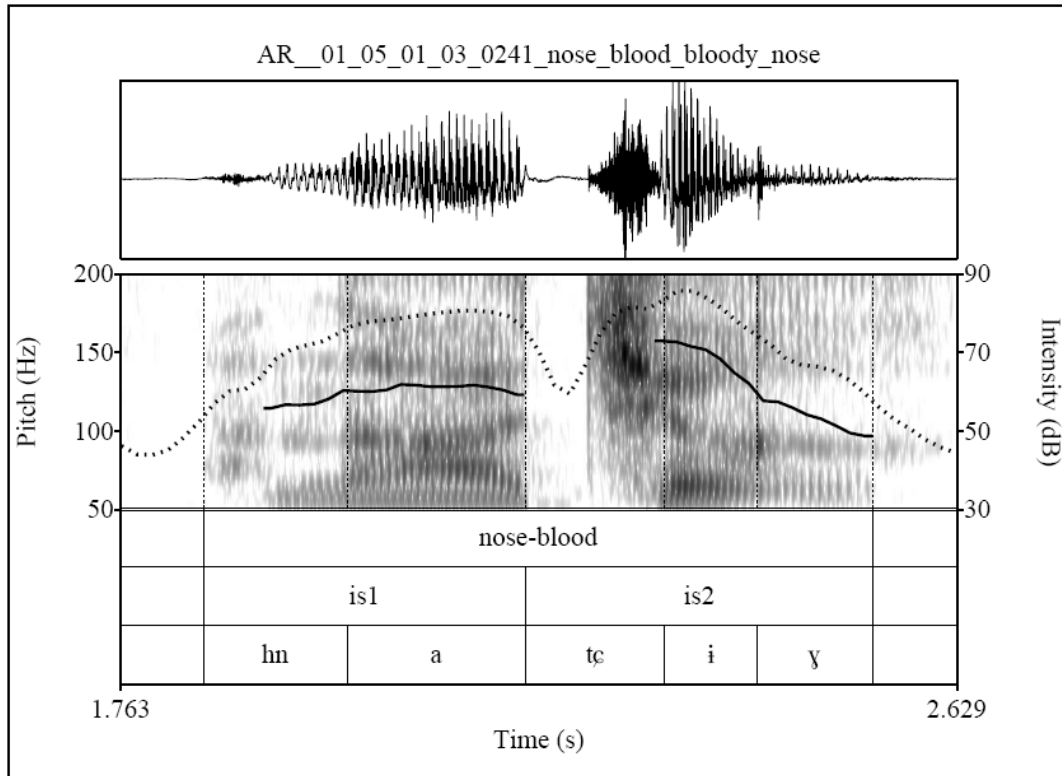
³² Huang’s reference to “syllables of all types” refers to various possible onsets and rhymes.

Figure 3.9 Rebkong Amdo (AR_05) / Pitch traces for [ɸna] *sna* ‘nose’ and [tɕ^hiɣ] *khrag* ‘blood’



What is interesting here is that when these monosyllabic words are combined to form the compound [ʰna.ʰtɕiɣ] *sna.khrag* ‘nose blood, bloody nose’, an entirely different pitch pattern emerges, as illustrated in Figure 3.10 below. In σ_1 , pitch is fairly flat across the vowel; in σ_2 , pitch slopes downward at a steep slope over the rhyme. (In effect, the same pitch pattern that conveys stress in monosyllabic words is found in the stressed syllable of disyllabic words.) In Chapter 4, I show that this pitch slope is one of the primary means of conveying σ_2 stress in Rebkong Amdo nouns.

Figure 3.10 Rebkong Amdo (AR_05) / Pitch trace [ʰna.ˈtɕiɣ] sna.khrag ‘nose blood, bloody nose’



The pitch trace in Figure 3.10 shows that two things have *not* happened. First, the two monosyllabic words have not each preserved their individual pitch patterns. If they had, we would see a high falling pitch on both σ_1 and σ_2 . Second, the high-falling pitch pattern which occurs in the monosyllables cannot be interpreted as a tone pattern common to all nouns, distributed over the domain of the word. If that were the case, we would expect to see a high tone on σ_1 , and a falling or low tone on σ_2 , matching the pitch contour of the monosyllabic words. These are two phenomena that we might expect to see if Rebkong Amdo were a tone language.

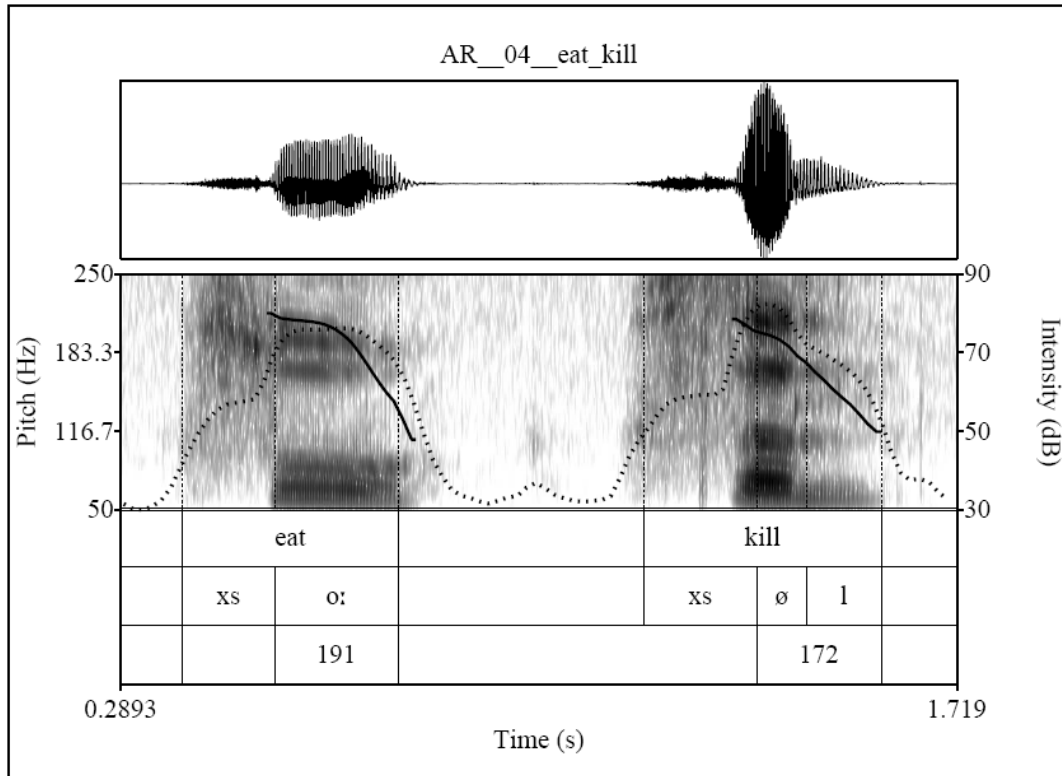
Instead, the pitch trace here can be regarded as a template for disyllabic non-verbs in Rebkong Amdo. The same pattern can be seen in the many figures in Chapter 4.

3.4.3 Pitch in monosyllabic verbs

As a final observation, monosyllabic verbs can exhibit the same pitch patterns as the nouns described above. The pitch traces of the imperative forms of the verbs [xso:] *gso* / *gsos* ‘eat’³³ and [xsøɭ] *gsod*, *bsad* ‘kill’ are spliced together in Figure 3.11 below. These exhibit the same falling contour that was illustrated for nouns in Figure 3.5 and Figure 3.8.

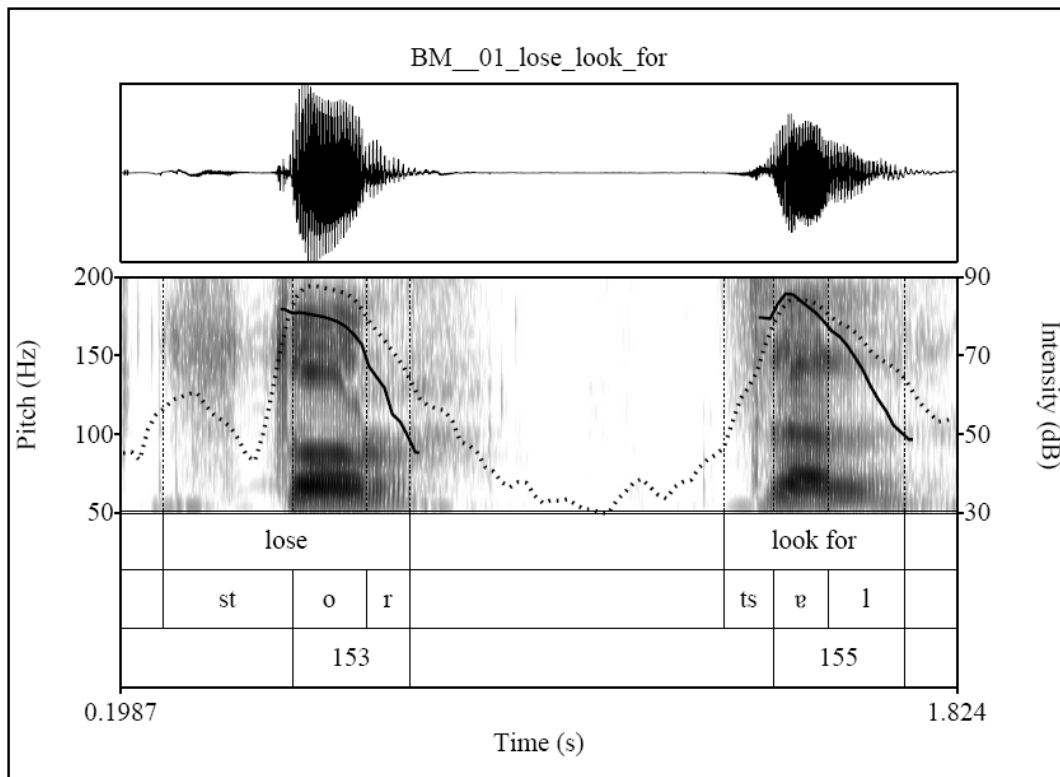
³³ Roland Bielmeier (p.c., 2008) suggests that this Balti form may be more appropriately translated as ‘to feed, to nourish, to raise’.

Figure 3.11 Rebkong Amdo (AR_04) / Pitch traces for [xso:] *gso* / *gsos* ‘eat’ and [xsø] *gsod*, *bsad* ‘kill’



In Balti, too, monosyllabic verbs exhibit the same pitch curves as monosyllabic nouns. The bare stem forms of the verbs [stor] *stor* ‘lose’ and [tsəl] *btsal* ‘look for’ are shown in Figure 3.12 below. These show much the same pattern as nouns produced by the same speaker, illustrated in Figure 3.4 and Figure 3.7.

Figure 3.12 Balti (BM_01) / Pitch traces for [stor] *stor* ‘lose’ and [tsəl] *btsal* ‘look for’



As I go on to demonstrate in the following chapters, for both Balti and Rebkong Amdo the pitch patterns for disyllabic nouns are distinctly different than those for disyllabic verbs. Thus it is interesting to observe that pitch is not distinctive in monosyllabic nouns *vs.* monosyllabic verbs.

3.5 Disyllabic words

Disyllabic words are at the heart of this study. In both Balti and Rebkong Amdo, disyllabic nouns, adjectives, and numerals are stressed on σ_2 , while disyllabic verbs are stressed on σ_1 . The acoustic correlates of these stress patterns are considered in Chapters 4 and 5. Here I discuss other types of evidence of stress,

including compounds, borrowed words, clitics, and noun-verb minimal pairs. I also comment on some lexical exceptions and ambiguous cases.

3.5.1 Compound words

Robust evidence of the correlation between lexical category, syllable position, and stress is provided by cases in which the same morphological and phonetic material is stressed differently depending on its position in a word. Compound numerals illustrate this beautifully. For instance, in Balti the monosyllables [χsum] *gsum* ‘three’ and [ft̪ɕu] *bcu* ‘ten’ are combined to form [t̪ɕuχ.'sum] *bcu.gsum* (ten+three) ‘thirteen’ and [sum.'t̪ɕu] *gsum.bcu* (three+ten) ‘thirty’ (all as produced by speaker BSh_03). In ‘thirteen’ it is the component meaning ‘three’ that is stressed, while in ‘thirty’ it is the component meaning ‘ten’ that is stressed. Regardless of content or meaning, stress always falls on σ₂; it is linked to position in the word, not to segmental or semantic content. We see the same pattern in the Rebkong Amdo numerals, as illustrated by [t̪ɕɤb.'dɤn] *bcu.bdun* (ten+seven) ‘seventeen’ and [h̥dɤn.'t̪ɕɤ] *bdun.cu* (seven+ten) ‘seventy’.

Such pairs can be found among nouns, too. For instance, in Rebkong Amdo, the words *chu* ‘water’ and *thig.pa* ‘drop’ can be compounded to form [t̪ɕh̥ɤ.'t̪iɤ] *chu.thig* ‘drop of water’ as well as [t̪h̥iɤ.'t̪ɕɤ] *thig.chu* ‘leak’ (i.e., a drop of water that

has come through the roof). (These forms were produced by speaker AR_05.) Stress is consistently on σ_2 , no matter which way the components are ordered.

There are other types of word-pairs which show that an element will be stressed or unstressed according to its position in the word. Balti speaker BM_01 produced the monomorphemic noun [$\chi\lambda t.'pa$] *klad.pa* / *glad.pa* ‘brain’ and the compound noun [$\eta go.'rlat$] *mgo.glad* (head+brain) ‘brain, mind’. In the former, the element corresponding to *klad* / *glad* is unstressed, because it occurs in σ_1 ; in the latter, the element corresponding to *glad* is stressed, because it occurs in σ_2 . Likewise in the nominalized form [$\zeta es.'k^{\times}an$] *shes.mkhan* ‘expert’, the element corresponding to *shes* is unstressed, occurring in σ_1 of a noun, but in the infinitive [$\zeta es.pa$] *shes.pa* ‘to know’ it is stressed, occurring in σ_1 of a verb.

Again, similar examples can be found in Rebkong Amdo. In the monomorphemic noun [$tho\chi.'wa$]³⁴ *thag.pa* ‘rope’, stress falls on σ_2 ; the semantic head is unstressed. But in the compound noun [$xt\ \square\square\square.'t\ \square\square\square$] *lcags.thag* (iron+rope) ‘chain’, the component corresponding to *thag* is now stressed because it is the second syllable. Again, we see that stress is associated with syllable position, not with segmental or semantic content.

³⁴ This is among the words that could not be reliably segmented, and so could not be included in my acoustic analysis. The stress pattern is nonetheless perceptually quite clear.

3.5.2 Borrowed words

Both Balti speakers provided borrowed nouns which are consistent with the σ_2 stress pattern, as illustrated by the examples in Table 3.17 below. In [ho.'tʃil] ‘hotel’, the σ_2 stress in the Balti form is the same as in the English source. However, when ‘teacher’, ‘bottle’ and ‘thermos’ are borrowed, English σ_1 stress shifts to Balti σ_2 stress. In the words ‘school’ and ‘film’, an epenthesis vowel serves to create a preferred two-syllable word from a dis-preferred one-syllable word. In [su.'kul], stress remains on the original vowel. In the case of ‘film’, however, a native speaker of English might expect the resultant disyllabic word to be stressed on σ_1 , since this is the original vowel. But when the word is produced in Balti, the epenthesis vowel in σ_2 is stressed instead; the vowel here is not a phonetically reduced one. Finally, the trisyllabic word [has.pi.'tal] suggests a tendency for final stress in monomorphemic words longer than two syllables. (Though this is difficult to confirm in native Balti trisyllabic words, as discussed in section 3.6 below).

Table 3.17 Balti: Stress in borrowed nouns

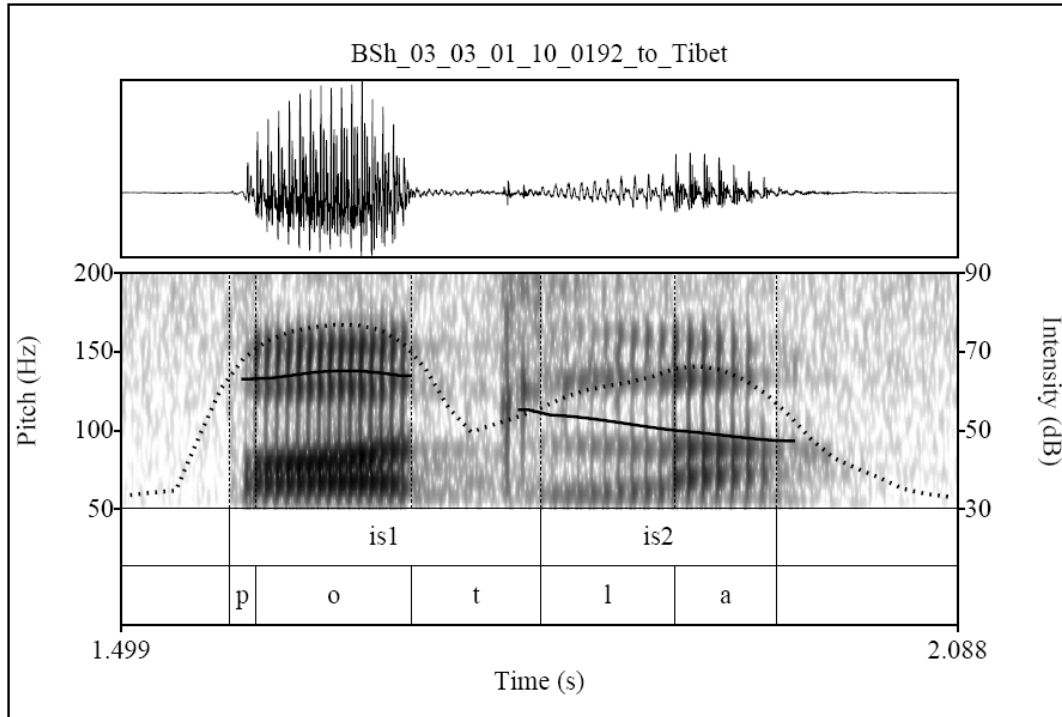
	English	Balti
a.	hotel	ho.'tʃɪl
b.	teacher	tʃi.'tʃəʃ
c.	bottle	mbʌ.'tʃɪl
d.	thermos	t ^θ ar.'mʊʃ
e.	school	su.'kul
f.	film	ʃi.'lim
g.	hospital	hʌs.pi.'tʃal

The examples above should be considered lightly; it is possible that these words were borrowed into Balti indirectly, via Urdu or some other language, and may preserve the stress pattern of some such intermediate language. On the other hand, most of them were provided by speaker BM_01 who had worked as a porter for many years with Western mountain climbers and was thus exposed to English directly.

3.5.3 Clitics

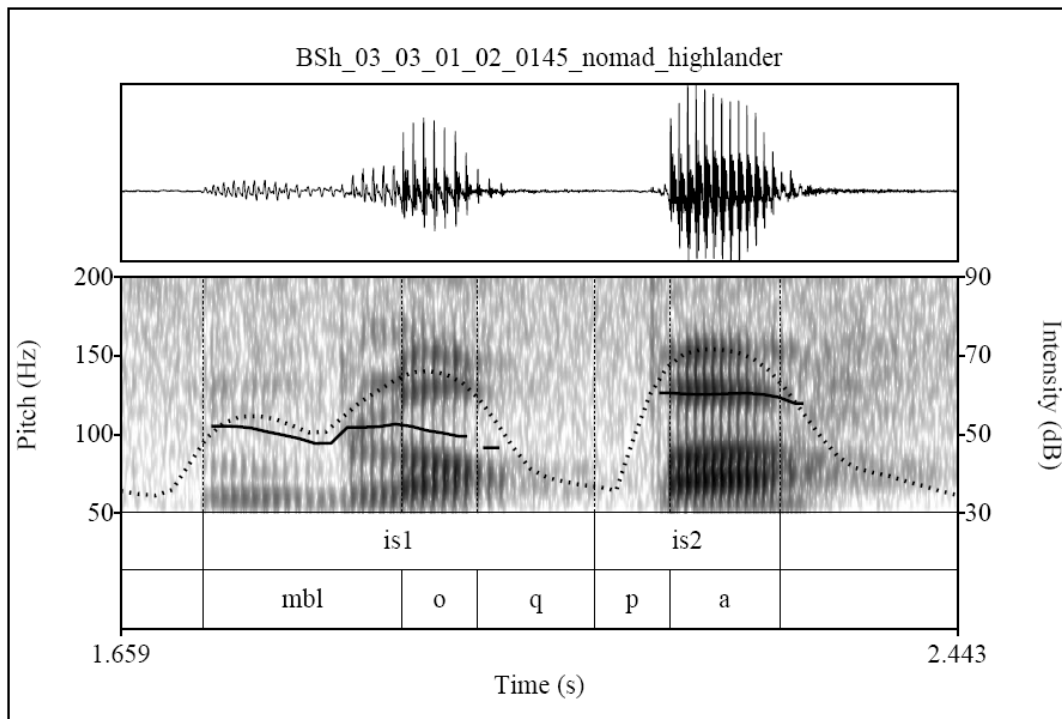
As noted in section 3.1.2, clitics do not bear stress. Thus disyllabic words composed of a noun + clitic – such as Balti [ˈndu.nu] *mdun.nu* ‘before, in front of’ and [ˈbod.la] *bod.la* ‘to Tibet’ – are stressed on σ_1 . For such structures, both pitch and intensity are higher in σ_1 than in σ_2 , as illustrated in Figure 3.13 below (Δ pitch = -39 Hz, Δ intensity = -11 dB).

Figure 3.13 Balti cliticized noun (BSh_03) / Pitch and intensity for ['bod.la] *bod.la* ‘to Tibet’



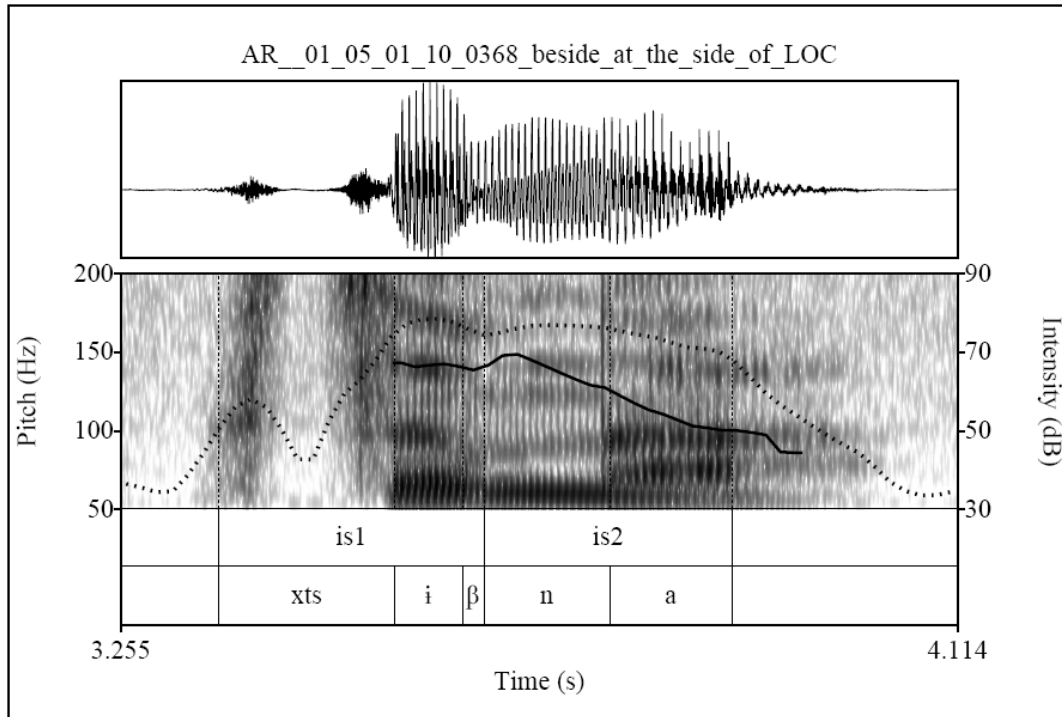
This is the inverse of the pattern observed in a typical monomorphemic noun, such as [mbloq.'pa] *'brog.pa* ‘nomad, pastoralist’, shown in Figure 3.14 (Δ pitch = +23 Hz, Δ intensity = +5.3 dB).

Figure 3.14 BSh_03 / Pitch and intensity for [mbloq.'pa] 'brog.pa 'nomad, pastoralist'



Cliticized forms in Rebkong Amdo show the same patterns. Since the locative clitic [=na] is non-stress-bearing, in the word [ˈxtsɪβ.na] *rtsib.na* 'beside, at the side of', stress falls squarely and prominently on σ_1 . This is illustrated in Figure 3.15 below. (Δ pitch = -31 Hz, Δ intensity = -4 dB).

Figure 3.15 Rebkong Amdo cliticized noun / Pitch and intensity for [ˈxtsiβ.na] *rtsib.na* ‘beside, at the side of’



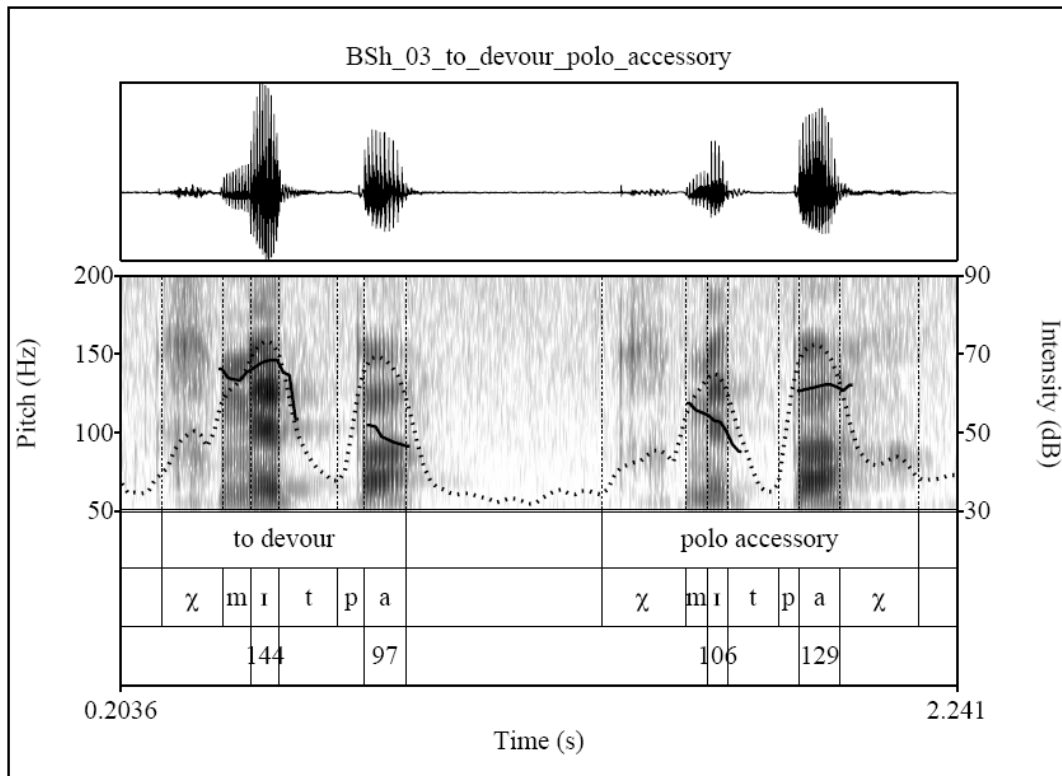
3.5.4 Noun-verb minimal pairs

As mentioned in section 3.1.2 above, in their descriptions of Balti both Sprigg (1966) and Bielmeier (1988a) included a handful of noun-verb minimal pairs which differ only in terms of stress placement. I elicited some of these same contrasts during my own work with speaker BSh_03. This prompted him, over the course of a few days, to provide me with a larger set of minimal pairs and near-minimal pairs. In these, without exception, verbs are stressed on σ_1 , and non-verbs are stressed on σ_2 .

One such near-minimal pair is illustrated in Figure 3.16 below. In the verb [ɣmit.pa] *mid.pa* ‘to devour, to swallow’, both pitch and intensity are higher in σ_1

than in σ_2 ; the converse is true in the noun [χmit.'paχ] ?? ‘polo accessory for a horse’.

Figure 3.16 BSh_03 / Pitch and intensity curves for ['χmit.pa] *mid.pa* ‘to devour, to swallow’ and [χmit.'paχ] ? ‘polo accessory for a horse’



A complete list of the noun-verb minimal pairs and near-minimal pairs provided by speaker BSh_03 is provided in Table 3.18 below.

Table 3.18 Balti noun-verb minimal pairs and near-minimal pairs

	Gloss	Lex cat	Transcription	WT
a.	to stretch	V	'çkyãŋ.ma	<i>brkyang.ma</i>
	to fill	V	'skãŋ.ma	<i>bkang.ma</i> ?
	foot	N	kãŋ.'ma	<i>rkang.pa</i>
b.	to clothe (trans.)	V	'skøn.ma	<i>skon.ma</i>
	rare	A	çkøn.'mo	<i>dkon.po</i>
c.	to wear (clothes)	V	'gøn.ma	<i>gon.ma</i>
	clothing	N	gøn.'tçes	<i>gon.chas</i>
d.	to carry	V	'k ^h ur.ba	<i>'khur.ba</i>
	bread	N	k ^h ur.'ba	<i>'khur.ba</i>
e.	to be tired	V	'xlat.pa	<i>*glad.pa</i>
	brain	N	xlat.'pa	<i>klad.pa / glad.pa</i>
f.	to arrive	V	't ^h øn.ma	<i>'thon.ma</i>
	high	A	t ^h øn.'mo	<i>mthon.mo</i>
g.	to grind	V	't ^h aq.pa	<i>'thag.pa</i>
	rope	N	t ^h aq.'pa	<i>thag.pa</i>
h.	to assign a name	V	'taq.pa	<i>btag.pa</i>
	birch tree	N	staq.'pa	<i>stag.pa</i>
i.	to give (by hand?)	V	'min.ma	<i>sbyin.ma</i> ?
	to be ripe	V	'smin.ma	<i>smin.ma</i>
	eyebrow	N	smin.'ma	<i>smin.ma</i>
j.	to light, ignite	V	'spar.ba	<i>spar.ba</i>
	forehead	N	spal.'ba	<i>dpral.ba</i>
k.	to walk	V	'drul.ba	<i>'grul.ba</i>
	vegetable garden	N	drum.'ba	<i>ldum.ra</i>
l.	to harm	V	'χnot.pa	<i>gnod.pa</i>
	sick person, patient	N	nɔt.'pa	<i>nad.pa</i>

* Reconstructed Proto-Tibetan form, provided by Roland Bielmeier (p.c. 2008). No Written Tibetan form could be identified.

Table 3.18 (cont.) Balti noun-verb minimal pairs

	Gloss	Lex cat	Transcription	WT
m.	to be clean, pure	V	'daq.pa	<i>dag.pa</i>
	boss, head man	N	ndaχ.'po	<i>bdag.po</i>
n.	to beat	V	^o rduŋ.ma	<i>rdung.ma</i>
	beam	N	^ə ɛduŋ.'ma	<i>gdung.ma</i>
o.	to straighten	V	'strā:ma	<i>srang.ma</i>
	pea, legume	N	stren.'ma	<i>sran.ma</i>
p.	to get up, to rise	V	'laŋ.ma	<i>langs.ma</i>
	lama	N	la.'ma	<i>bla.ma</i>
q.	to ride (variation)	V	'dzrøn.ma	??
	guest	N	ɖrøn.'pa	<i>mgron.pa</i>
r.	to laugh, get wild	V	'rgot.pa	<i>rgod.pa</i>
	stomach	N	ʈrot.'pa	<i>grod.pa</i>
s.	to survive, to be nourished	V	'χsøn.ma	<i>gson.ma</i>
	digit (finger, toe)	N	sen.'mo	<i>sen.mo</i>
t.	to drip	V	^o ɛzar.ba	<i>'dzar.ba / gzar.ba / bzar.ba</i>
	new	N	sar.'p ^h a	<i>gsar.pa</i>
u.	to devour, swallow	V	'χmit.pa	<i>mid.pa</i>
	polo accessory for horse	N	χmit.'paχ	??
v.	to entangle	V	ʈ ^h r̥is.pa	<i>'khris.pa</i>
	difficult	A	ʈ ^h r̥is.'pa	<i>mkhregs.pa</i>
w.	to fight, to wrestle	V	ʈ ^h r̥il.ba	<i>'khril.pa</i>
	shame	N	ʈ ^h r̥el.'ba	<i>khrel.ba</i>
x.	to shave	V	'blaq.pa	<i>'breg.pa</i>
	nomad, pastoralist	N	bloq.'pa	<i>'brog.pa</i>
y.	to dry, evaporate	V	^o ras.pa	<i>ras.pa ?</i>
	bone	N	^ə rus.'pa	<i>rus.pa</i>

Interestingly, with the exception of [nʌt.'pa] *nad.pa* ‘sick person, patient’ and [ʰχnot.pa] *gnod.pa* ‘to harm, to injure’, none of these pairs have an obvious semantic relationship. My consultant did not offer me minimal pairs like Sprigg’s verbal noun / infinitive form [ʰrgɔs.pa] ‘needing / to need’ vs. the noun [rgɔs.'pa] ‘need’. Yet in a dictionary of Central Tibetan like Goldstein (2001) one can easily find Written Tibetan forms like *khur.pa* ‘porter, coolie’ or *grul.pa* ‘traveler, passenger, visitor, guest’, similar to (d) and (k) in the table above. And, according to Bielmeier (p.c., 2008), the word [kʰur.'pa] ‘porter’ does, indeed, occur in Balti. I am not sure whether or not semantic cognates for other words in the list *do* exist in Balti, but this question certainly merits further investigation.³⁵

3.5.5 Lexical exceptions and anomalies

Sprigg (1966) and Bielmeier (1988a) also noted that there were occasional lexical exceptions to the general stress pattern in Balti. I found this to be true as well; a few very clear lexical anomalies are briefly mentioned in Chapter 4. There were a number of other cases in which the stress pattern was perceptually (and acoustically) ambiguous. For instance, in the adjective [sar.pʰa] *gsar.pa* ‘new’, the σ_1 vowel is longer and has a higher intensity than the σ_2 vowel, but the σ_2 vowel has a higher

³⁵ It is possible that my consultant was most eager to provide me with examples similar to the few I elicited, like (e) in Table 3.18, where the forms are not semantically related, and simply did not think the “more predictable” pairs would be of interest to me, so did not offer them. Sprigg (1966) and Bielmeier (1988a) include minimal pairs of both types – those in which the members are semantically related, and those in which they are not, the latter apparently being more common.

pitch. For a native speaker of English – accustomed to attending to all of these factors as cues for stress – this represents conflicting information, and it is impossible to state with certainty that one syllable is perceived as more prominent than the other.

Nonetheless, such cases were in the minority. The patterns I report are based on exposure to a considerably larger sample than that evaluated here: less than one-third of the words I recorded from speakers BM_01 and BSh_03 could be reliably segmented and analyzed, as discussed in Chapter 2. Thus my judgments are based on having elicited, listened to, and discussed a much larger sample. In addition, as noted in Chapter 2, I also elicited and closely transcribed my 500-word list with a third speaker of Balti; there, too, the patterns of σ_2 stress on non-verbs and σ_1 stress on verbs is overwhelming.

3.6 Words of three or more syllables

In sections 3.6.1 and 3.6.2 below I offer some brief and tentative remarks about the stress patterns which may exist in words longer than two syllables, in Balti and Rebkong Amdo. This is an area which certainly merits further investigation.

3.6.1 Longer words in Balti

Table 3.19 below offers a few examples of nouns and verbs which are composed of more than two syllables. The stress marks here should be considered tentative; acoustic cues conflict with one another, so relative syllable prominence is not unambiguous. At this point, I can be sure only that it would *not* be accurate to say that nouns are always stressed on the last syllable, or that verbs are always stressed on

the first syllable, although that would fit tidily with what is observed on disyllabic words. A more careful analysis is clearly needed; this should take into consideration morphological composition and structure.

Table 3.19 Balti / Words of more than two syllables

Gloss	IPA	WT	Composition	Structure
water drop	tʃu.tʰik.pa	<i>chu.thig.pa</i>	N + N	[σ1] + [σ2. σ3] ?
kitten	bi.'la.pʰru	<i>byi.la.'phru</i>	N + Dimin	[σ1.σ2] + [σ3]
limbs	lak.'pa.kã.ma	<i>lag.pa.rkang.ma</i>	N + N	[σ1.σ2] + [σ3.σ4]
to ride a horse	'ʂta.ɹøŋ.ma	<i>rta.zhon.ma</i>	N + V	[σ1] + [σ2.σ3]
to light a fire	'me.spar.ba	<i>me.spar.ba</i>	N + V	[σ1] + [σ2.σ3]
talk! converse!	ɹxpe.'ra.ton	<i>dpe.sgra.thong</i>	N + V	[σ1.σ2] + [σ3]
to rain	tʃʰar.'pa.tã.ma	<i>char.pa.btang.ma</i>	N + V	[σ1.σ2] + [σ3.σ4]

3.6.2 Longer words in Rebkong Amdo

In listening to and examining the acoustic signals of Rebkong Amdo nouns, adjectives, and verbs of three or four (or in a few cases, five) syllables, I have found it very difficult to identify a stress pattern. In all cases, I have an impression that there is a primary stress on the final syllable, regardless of the length of the word – but I am not completely confident even of this. Furthermore, in many cases there seems to be a secondary stress on an earlier syllable. But I cannot be certain of this, either, as acoustic signals are mixed and contradictory: one syllable may have a higher intensity than another, but may have a shorter vowel; one syllable may have a higher pitch or a more dramatic pitch slope than another, but a lower intensity. I am not sure it is possible for a native speaker of English to draw a distinction between more prominent

and less prominent syllables here. And I am not sure a native speaker of Rebkong Amdo could do this either – as described in section 2.1.4, when I attempted to conduct “clapping” tests on words of only two syllables, the Rebkong Amdo speakers I worked with were able to count the number of syllables through clapping, but did not seem to understand my interest in relative syllable prominence.

Variations in morphological composition add another level of complexity.

Some examples are provided in Table 3.20 below.

Table 3.20 Rebkong Amdo / Words of more than two syllables

Gloss	IPA	WT	Composition	Structure
elephant	°lõ.wo.tɕe	<i>glang.po.che</i>	N + Adj	[σ1] + [σ2.σ3] ?
white hair	xtɕa.ka.ru	<i>skra.dkar.po</i>	N + Adj	[σ1] + [σ2.σ3]
sheep’s head	li.ɣŋ.go	<i>lug.gyi.mgo</i>	N + Gen + N	[σ1 + σ2] + [σ3]
neither farmer nor nomad	°rõ.man.dɔy	<i>rong.ma.'brog</i>	N + Neg + N	[σ1] + [σ2] + [σ3]
white prayer flag	dar.tɕok.ka.ru	<i>dar.lcog.dkar.po</i>	N + Adj	[σ1.σ2] + [σ3.σ4]
crooked	xtɕa.ɣa.xtɕo.yi	<i>kyag.ge.kyog.ge</i>	Reduplicated	[σ1.σ2] + [σ3.σ4]
to dream	nyə.ləm.nyi	<i>gnyid.lam.rmi</i>	N + V	[σ1.σ2] + [σ3]
to work	le.ka.le.'go.ki	<i>las.ka.las.??</i>	N + V +	[σ1.σ2] + [σ3] + ?

Thus I have little to report on the stress patterns of words longer than two syllables in Balti and Rebkong Amdo. All that we can be certain of, at present, is that the σ1 stress observed in disyllabic verbs does not reflect a pattern of initial stress in verbs of all lengths; likewise the σ2 stress observed in disyllabic non-verbs does not reflect a pattern of final stress in non-verbs of all lengths.

4. Acoustic correlates of stress in Balti Tibetan

In this chapter I present my analysis of the acoustic correlates of stress in Balti, based on recordings from speakers BSh_03 and BM_01. As noted in section 2.1.1.1, speaker BSh_03 was well-educated and fluent in English, and elicitation sessions were focused and controlled. Consequently, patterns are clearly manifested and distinctions are sharp. Speaker BM_01, on the other hand, was older, uneducated, and sometimes mumbled or else called out words with tremendous enthusiasm. As a consequence, for BM_01 the four acoustic correlates exhibit a broader range in values, and patterns are somewhat more diffuse. In addition, I did not conduct the elicitation session myself, but attended and recorded a session conducted by a colleague for his own research purposes. Thus I was only able to record words in isolation; there is no data for frame forms. Despite these differences, Balti's patterns of σ_2 stress for nouns, adjectives, and numerals, and σ_1 stress for verbs are clearly manifested.

In section 4.1 I summarize my findings by lexical category: non-verbs in section 4.1.1 and verbs in section 4.1.2.

The details of the analysis are presented in sections 4.2 through 4.5, organized by acoustic parameter. For instance, all pitch data is considered in section 4.2 – first for speaker BSh_03 (first nouns, then adjectives, then numerals, then verbs), and then for speaker BM_01 (nouns, then adjectives, then verbs). Pitch slope is likewise considered in section 4.3, intensity data in section 4.4, and vowel duration data in section 4.5.

Finally, in section 3.6 I provide a recapitulation. Table 4.36 on page 323 provides a grand summary of the role played by each acoustic parameter in conveying stress.

4.1 Summary of acoustic correlates of stress in Balti

The acoustic correlates of stress are summarized for non-verbs in section 4.1.1, and for verbs in section 4.1.2.

4.1.1 Non-verbs

Nouns, adjectives, and numerals are all clearly stressed on σ_2 , in sharp contrast to verbs, which are stressed on σ_1 . For both Balti speakers, the analysis of nouns is based on a fairly large sample. The set of adjectives and numerals which were recorded and which could be analyzed is quite limited. Nonetheless, it is sufficient to demonstrate that words in these lexical categories behave acoustically like nouns, and not like verbs.

4.1.1.1 Nouns

For both Balti speakers, I recorded more tokens of nouns than of any other lexical category.

For speaker BSh_03, I was able to segment and analyze 77 isolation forms and 64 frame forms. These groups included more than thirty monomorphemic forms – such as [tut.'pa] *dud.pa* ‘smoke’, [ɓax'mo] *bag.mo* ‘bride’, and [ra.'ma] *ra.ma* ‘goat’ – and more than thirty compound forms – such as [le.'mik] *lde.mig* ‘key’ (?+eye),

[tʃa.'pʰe] *ja.phye* ‘tsampa’ (tea+flour), and [pʰaq.'ʃa] *phag.sha* ‘pork’ (pig+meat).

There was also one reduplicated form ([ʧoq.'ʧoq] *shog.shog* ‘paper’), and a few borrowed words (such as [ti.'tʃəŋ] ‘teacher’ – from English – and [mu.'tik] *mu.tig* ‘pearl’ – a Middle Indian loan from Sanskrit *mauktika*-³⁶).

From speaker BM_01, I was able to segment and analyze 72 disyllabic nouns. This set included 39 monomorphemic nouns – such as [ʔspal.'ba] *dpral.ba* ‘forehead’, [bʊŋ.'bu] *bong.bu* ‘donkey’, and [bo.'ŋo] *pu.mo* ‘girl’ – and 19 compound nouns – such as [nas.'pʰe] *nas.phye* (barley+flour) ‘barley flour, tsampa’, [ŋgo.'rel] *mgo.ral* (head+hair knot) ‘hair on the head’, [ʔrgyam.'so] *rgya.mtsho* (vast+lake) ‘big river’, and [byaβ.'zøŋ] *bya.bzhon* (chicken+egg) ‘egg’. There were also two nominalized forms – [ʧes.'kʰan] *shes.mkhan* ‘wise person; one who knows’ and [ʧtsaŋ.'kʰan] *bslang.mkhan* † / *ltsang.mkhan?* ‘beggar, mendicant; one who begs’ – and a handful of borrowed nouns, including [mɛr.'man] ‘foreigner, guest’ from Urdu; [mbʌ.'tʃl] ‘bottle’ from English; and the reduplicated noun [tʃu.'tʃu] ‘breast’ from Burushaski (p.c., Bielmeier 2008).

For speaker BSh_03, isolation and frame forms show the same patterns. The only acoustic cue for σ_2 stress is pitch, which shows a robust and reliable correlation. Pitch slope and intensity are definitively *not* correlates of stress. For vowel duration,

³⁶ Thanks to Roland Bielmeier for this etymological information (p.c. 2008).

the subset which remains once vowel quality and syllable closure type are controlled for contains only a few tokens; these tokens show no evidence of a correlation with stress.

For speaker BM_01, pitch is higher on σ_2 for most nouns, and is again the most consistent and reliable correlate of stress. Intensity is higher on σ_2 only under limited, controlled conditions (i.e., only in those words with vowels of the same height in both syllables). Pitch slopes downward fairly steeply in σ_2 in most nouns, but I believe this is an artifact of the final falling intonation contour rather than a correlate of stress. Nothing at all can be concluded about the potential role of vowel duration: since words were recorded only in isolation, there was no way to tease out the effects of utterance-final lengthening. Syllable closure also influences vowel duration.

4.1.1.2 Adjectives

The sample of adjectives recorded from speaker BSh_03 was quite small: only eight isolation forms and seven frame forms could be segmented and analyzed. All of these are monomorphemic, such as [t^hon.'mo] *mthon.mo* 'high' and [mar.'p^ho] *dmar.po* 'red'. A larger set of adjectives (n=26) was recorded from speaker BM_01, including monomorphemic: [lear.'mo] *legs.mo* 'good'; compound: [sɲo.'raŋ]

sngo.rang ‘blue color’ (with *rang* a loan from Urdu / Persian³⁷); reduplicated: [tʰäk.'tʰäk] ?? ‘fat’; and borrowed: [sas.'ta] ‘cheap, inexpensive’ (from Hindi).

Like nouns, adjectives in Balti are stressed on the second syllable. Pitch is a robust acoustic correlate of this stress pattern for both speakers. It can be demonstrated that intensity and vowel duration are *not* correlates of stress for speaker BM_01, but for BSh_03, the data set is too small for any determination to be made. Pitch slope does not convey stress for either speaker; for BSh_03, pitch slope reflects the contrast in pitch across syllables, and for BM_01 pitch slope reflects an utterance-final falling intonation.

4.1.1.3 Numerals

Numerals in Balti behave like nouns and adjectives, with a clear stress on σ_2 conveyed primarily by pitch. Intensity is definitively not a correlate of stress. The sample is too small and too constrained to determine the potential roles of either pitch slope or vowel duration.

Numerals were recorded only from speaker BSh_03, and only in isolation. The disyllabic forms which could be analyzed are all compounds: the teens, such as [tɕu.'ruk] *bcu.drug* ‘sixteen’, and the tens, such as [ʔɤaf.'tɕu] *lnɡa.bcu* ‘fifty’. In the pair [tɕuɣ.'sum] *bcu.gsum* ‘thirteen’ and [xsum.'tɕu] *gsum.bcu* ‘thirty’, the same elements are combined in reverse order, confirming that stress is governed by syllable position rather than by segmental or semantic content.

³⁷ Etymology from Bielmeier (p.c., 2009)

4.1.2 Verbs

Verbs in Balti Tibetan are stressed on σ_1 , in stark contrast to the nouns, adjectives, and numerals discussed above.

For speaker BSh_03, all but one of the more than thirty verbs recorded were citation forms, composed of a verb stem followed most often by [-pa], [-ma], or [-ba], and occasionally by [-p^ha] or [-va]. Examples include [ʼmɪn.ma] *sbyin.ma?* ‘to give’; [k^hur.ba] *ʼk^hur.ba* ‘to carry’; and [ʼtʃe.va] ?? ‘to raze, to destroy’. The latter is the only one of the citation forms in which σ_1 is open. The one verb with a Noun + Verbalizer morphological structure is the form [ʼza:n.za] *zan.za* ‘to eat food’, composed of ‘food’ + ‘eat’. This is also the only verb with a long vowel in either syllable.

Many of the verbs elicited from speaker BM_01 were given as monosyllabic forms (e.g., [ŋus] *ngus* ‘cry’; [t^hɔŋ] *thon?* ‘arrive’) or as trisyllabic N+Vblzr forms (e.g., [baχ.ʼston byas] *bag.ston.byas* ‘to marry’; [nyi.ʼlam t^hɔŋ] *nyi.lam.thong* ‘to dream’). Of the disyllabic forms, many included glides (e.g., [ʼrda.wa] *rda.pa?* / *rda.ba?* ‘to kill’; [ʼlaz.byā] *las.byā* ‘to work’), and so could not be segmented for analysis. In the end, only fourteen verbs could be analyzed, including both citation forms (e.g., [ʼspar.ba] *spar.ba* ‘to light [a fire], to ignite’) and N+Vblzr forms (e.g., [ʼskat.zer] *skad.zer* ‘to call, shout’).

For the non-verbs, pitch was the only robust correlate of stress. But this is not the case for verbs. Intensity is a very strong cue for stress in verbs produced by both speakers. Vowel duration is also generally longer in σ_1 – and this despite factors such as syllable closure and vowel height which are predicted to favor a longer vowel in σ_2 . Pitch slope is not a correlate of stress for either speaker; the slope patterns observed reflect the interaction of lexical pitch with utterance-level intonation.

4.2 Pitch

Pitch is a strong correlate of stress for all lexical categories, for both Balti speakers considered. For speaker BSh_03, analysis of nouns, adjectives, numerals, and verbs produced in isolation, and nouns, adjectives, and verbs produced within the sentence frame, are presented in section 4.2.1.

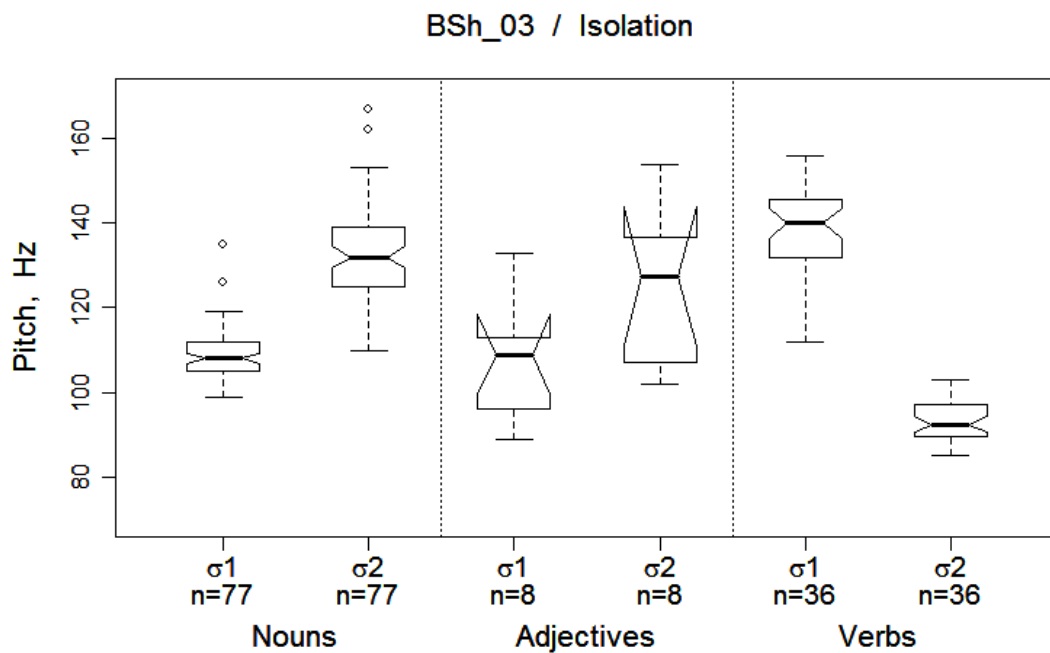
For speaker BM_01, analysis of nouns, adjectives, and verbs produced in isolation are presented in section 4.2.2. For this speaker, the magnitude of the pitch contrast across syllables is influenced by whether σ_2 is open or closed. The intrinsic variation of pitch as a function of vowel height does not appear to play a significant role.

4.2.1 Pitch for speaker BSh_03

For speaker BSh_03, pitch is the primary cue for stress for both isolation and frame forms, for all lexical categories. With only a few exceptions, pitch alone would be sufficient to convey stress.

For isolation forms, the general pattern of pitch contrasts across syllables is illustrated in Figure 4.1 below. (Numerals are not included here for lack of graphical space, but are addressed separately in section 4.2.1.3.) Pitch clearly distinguishes the σ_2 stress perceived on non-verbs from the σ_1 stress perceived on verbs.

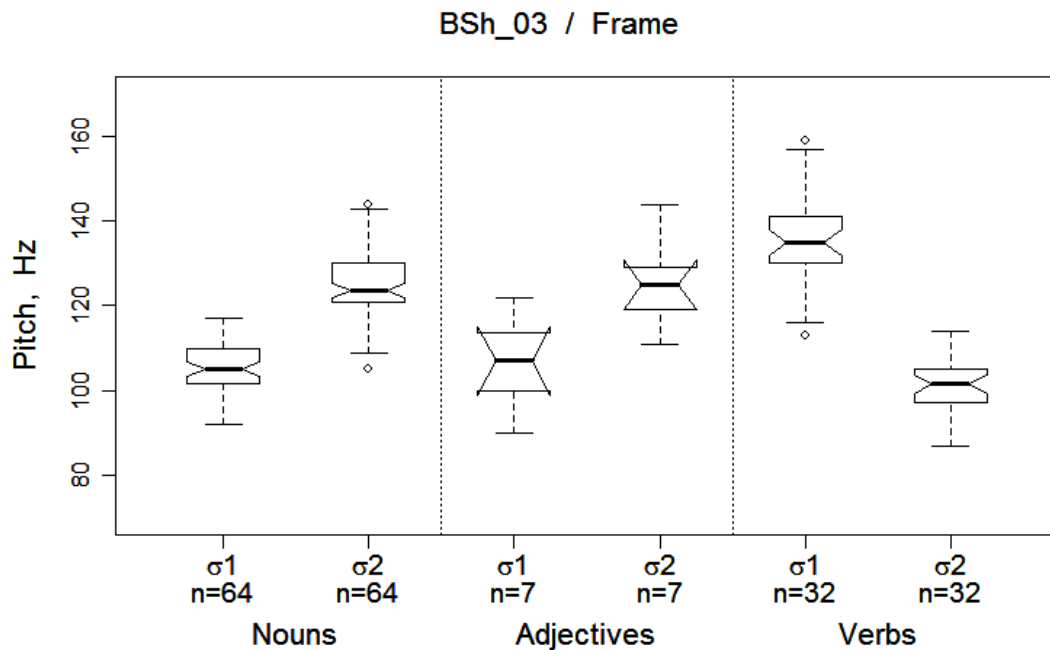
Figure 4.1 BSh_03 / Isolation: Pitch contrasts



Pitch contrasts for BSh_03 frame forms are illustrated in Figure 4.2 below.

The pitch range on the target word is more tightly constrained within the fixed context of the carrier sentence; for the non-verbs, the boxes are more compact and the range of values is narrowed.

Figure 4.2 BSh_03 / Frame: Pitch contrasts



As reported by Lehiste (1970: 68) and discussed in section 2.3.2.1, pitch (fundamental frequency) varies inherently as a function of vowel quality: F0 tends to be higher on high vowels, and lower on low vowels. For words of all lexical categories produced by speaker BSh_03, though, any such variation is moot: it is never sufficient to yield a higher pitch on the unstressed syllable. The correlation between pitch and stress is robust.

Pitch data for BSh_03 nouns, adjectives, numerals, and verbs are discussed in greater detail in sections 4.2.1.1 through 4.2.1.4 below.

4.2.1.1 BSh_03 Nouns

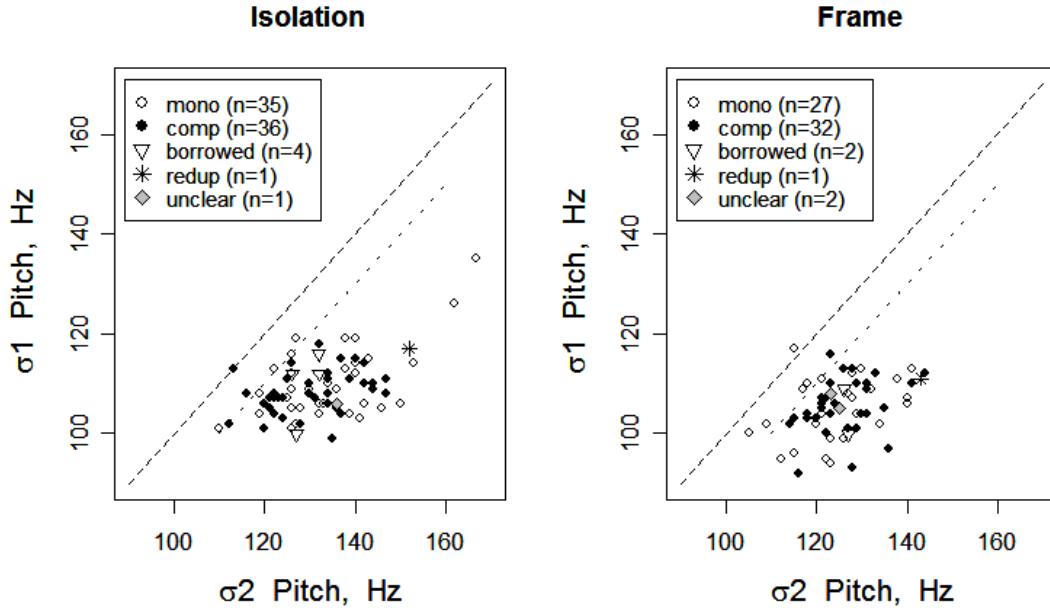
For speaker BSh_03, pitch is higher on σ_2 for nearly all nouns, regardless of origin or morphological structure, and regardless of whether they were produced in

isolation or in the sentence frame. As shown in the top part of Figure 4.3 below, the points define a fairly tight cluster for both settings, reflecting this speaker's consistent production; pitch ranges from ~100 to ~120 Hz for σ_1 , and from ~120 to ~140 Hz for σ_2 . The only exception among the isolation forms is [ɲgøŋ.'tʰɛs] *gon.chas* 'clothing', which lies squarely on the dashed line. The only exception among the frame forms is [xpe.'ra] *dpe.sgra* 'conversation, talk'; for this word, the pitch is a trivial 2 Hz higher on the vowel in σ_1 . This consistent correspondence between higher pitch on σ_2 and perceived stress on σ_2 means that pitch is a strong cue for stress. Most points even fall to the right of the dotted reference line at "Pitch difference = +10 Hz".

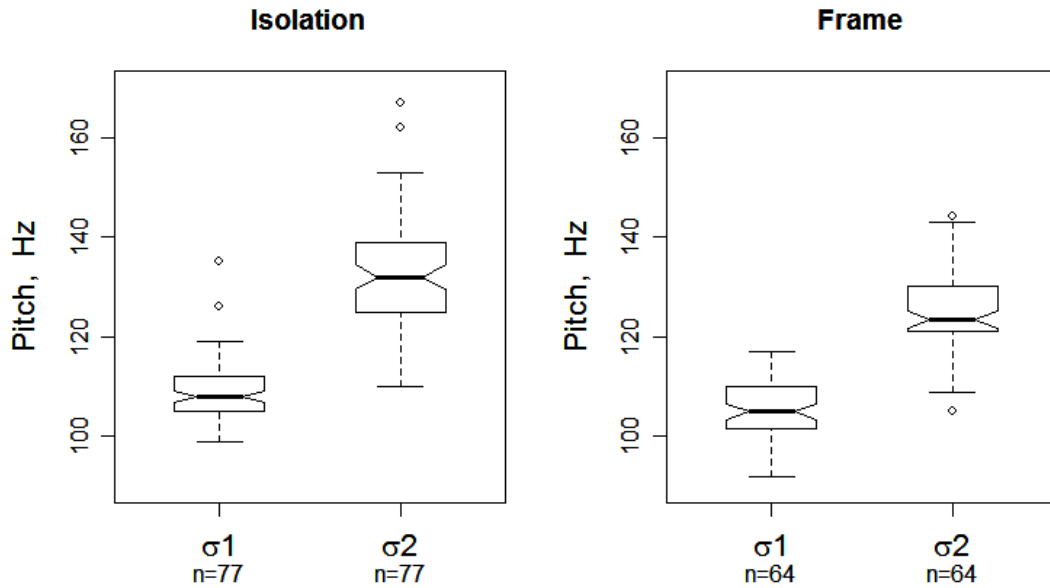
The box-and-whisker plots in the bottom part of Figure 4.3 show that there is no overlap of the notches – representing the 95% confidence interval about the median – or of the boxes – representing the interquartile range.

Figure 4.3 BSh_03 / Nouns: Pitch

BSh_03 / Nouns



BSh_03 / Nouns



The results of paired-sample t-tests for isolation and frame forms are summarized in Table 4.1 below. Since the p-values are $\ll 0.05$, we must reject the null hypothesis – which asserts that there is no difference between the two syllables – and conclude that the contrast in pitch is statistically significant. It is also perceptually significant: the mean increase in pitch from σ_1 to σ_2 is 23 Hz for isolation forms and 20 Hz for frame forms, and the smallest probable mean difference across syllables is 18 Hz.

Table 4.1 BSh_03 / Nouns / Pitch: results of paired-sample t-tests (two-tailed)

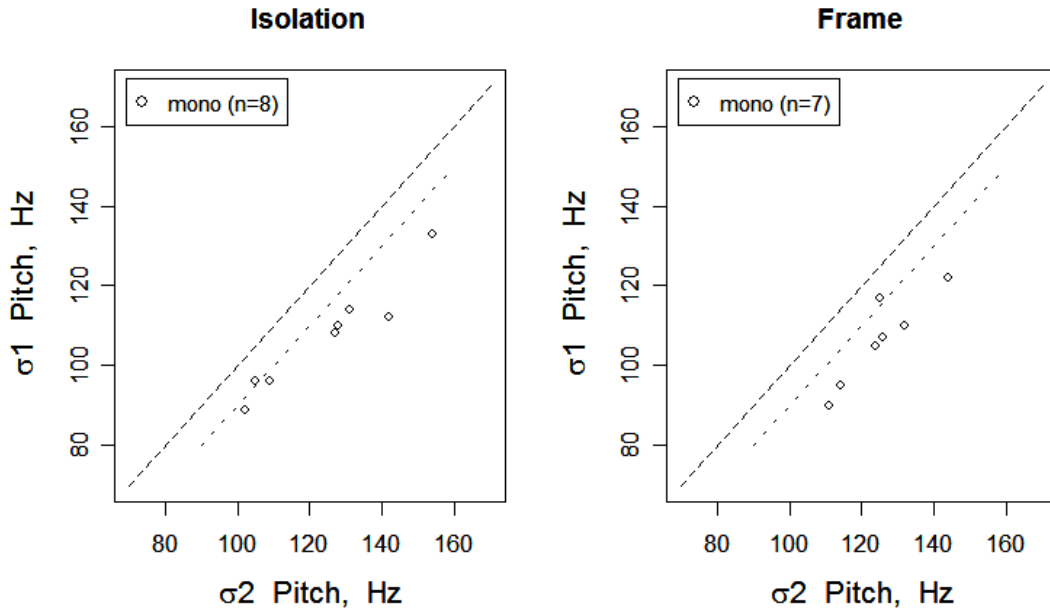
Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BSh_03	isolation	22.0007	76	$< 2.2e-16$	23	21	25
	frame	19.738	63	$< 2.2e-16$	20	18	22

4.2.1.2 BSh_03 Adjectives

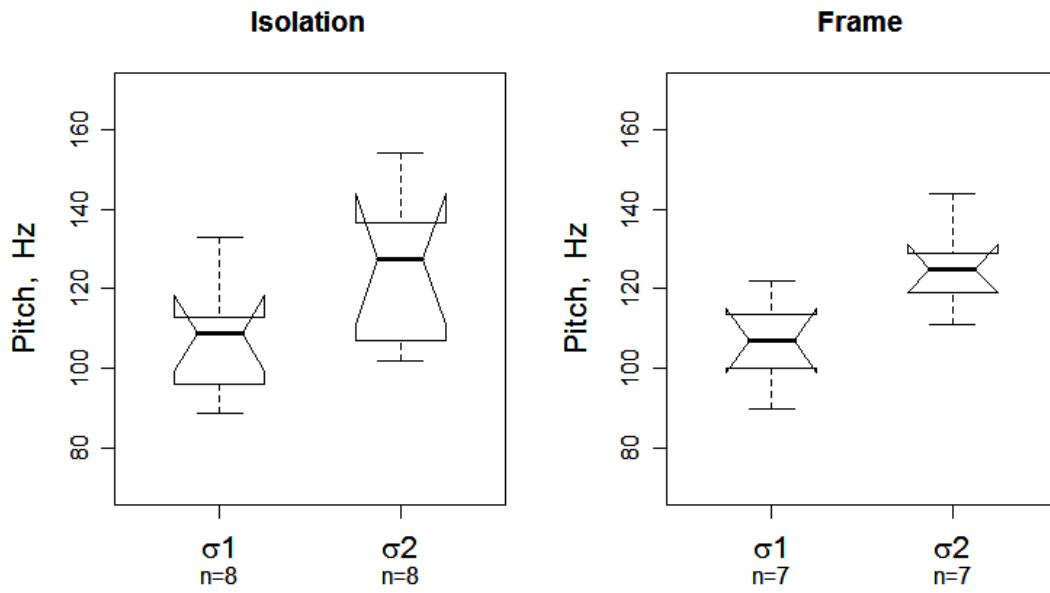
Pitch differences across syllables for BSh_03 adjectives are plotted in Figure 4.4 below. Without exception, pitch is higher in σ_2 , the stressed syllable, and in most cases the difference is greater than +10 Hz. For the isolation forms, the extent of the boxes and the whiskers in the bottom part of the figure reflect the wide range in values for both syllables. The distinction is more crisp for the frame forms, with no overlap of either the notches or the boxes.

Figure 4.4 BSh_03 / Adjectives: Pitch

BSh_03 / Adjectives



BSh_03 / Adjectives



Paired-sample t-tests, summarized in Table 2.7 above, confirm that the contrast in pitch across syllables is statistically significant. Even though the sample size is quite small, $p \ll 0.05$ for both isolation and frame forms. The contrast in pitch is also perceptually significant, averaging 18 Hz. Thus it is quite clear that pitch is a robust correlate of the σ_2 stress perceived on adjectives produced by speaker BSh_03.

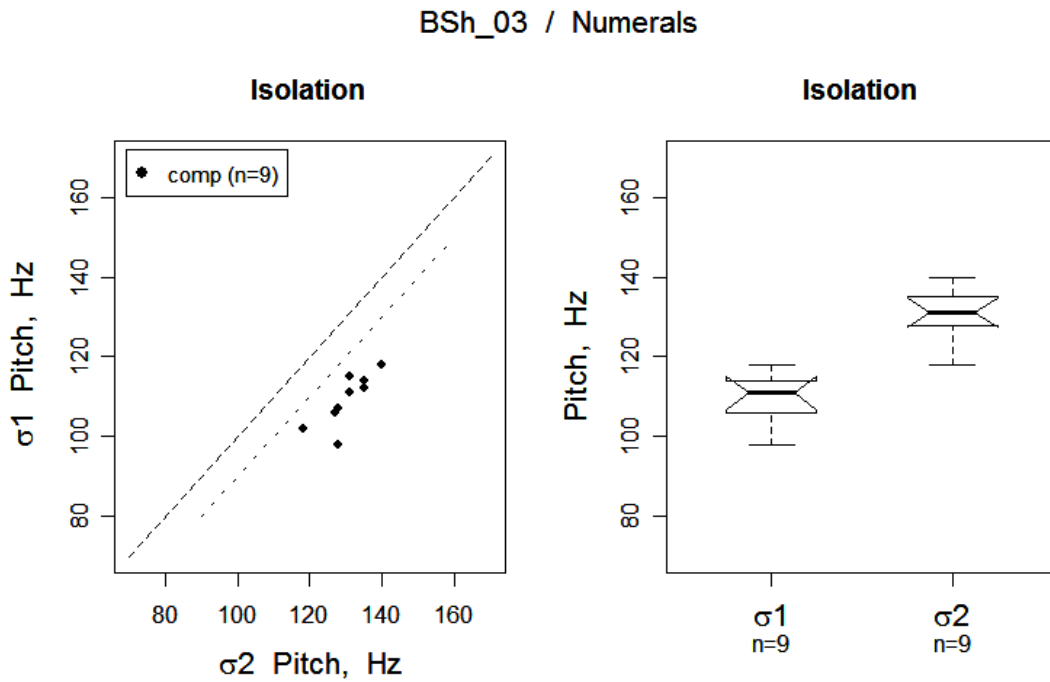
Table 4.2 BSh_03 / Adjectives / Pitch: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BSh_03	isolation	7.7709	7	0.0001097	18	12	23
	frame	10.1103	6	5.441e-05	18	14	23

4.2.1.3 BSh_03 Numerals

For numerals, too, pitch is uniformly higher on σ_2 – the stressed syllable – as illustrated in Figure 4.5. In the distribution plot on the left below, all points fall to the right of the dashed “Pitch difference = 0” line, and even to the right of the dotted “Pitch difference = +10 Hz” reference line. In the box-and-whisker plot on the right, there is no overlap of either the notches or the boxes.

Figure 4.5 BSh_03 / Numerals: Pitch



A paired-sample t-test, summarized in Table 4.3 below, confirms the significance of the contrast in pitch across syllables: $p \ll 0.05$; the mean difference in pitch is 21 Hz.

Table 4.3 BSh_03 / Numerals / Pitch: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BSh_03	isolation	15.3106	8	3.287e-07	21	18	24

Thus for numerals, as for the other non-verbs, pitch is a robust correlate of the perceived σ_2 stress pattern.

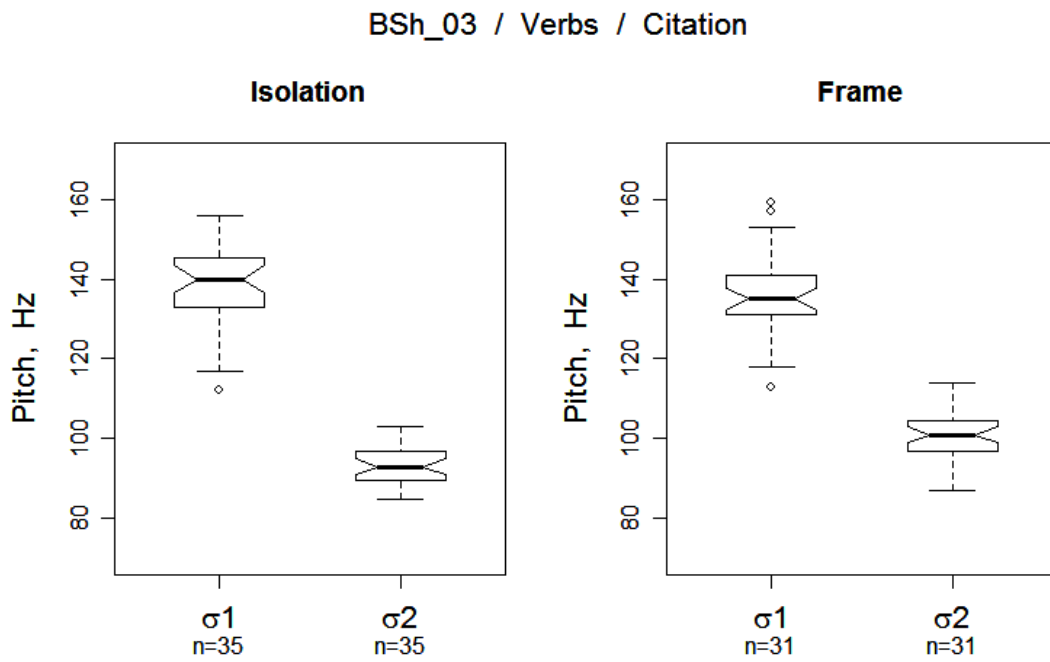
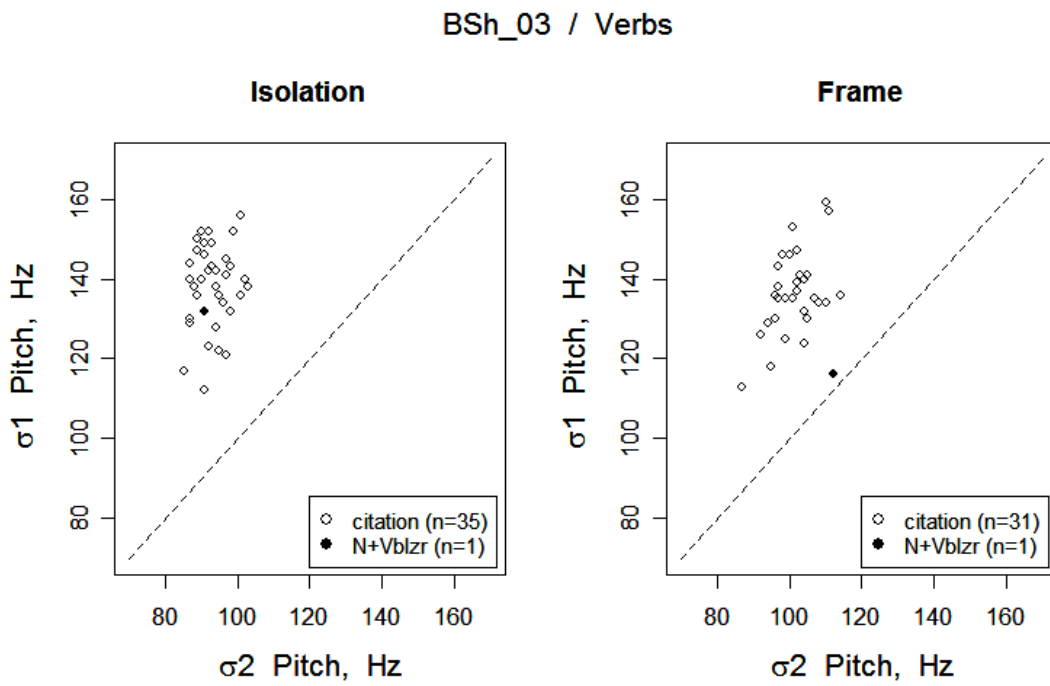
4.2.1.4 BSh_03 Verbs

Pitch contrasts across syllables for BSh_03 verbs are plotted in Figure 4.6 below. In isolation and in the sentence frame, all points fall to the left of the dashed line. That is, for all verbs, pitch is higher on σ_1 , the stressed syllable.

In the distribution plots, the isolation forms define a clear vertical band: pitch in σ_1 is highly variable, ranging from ~110 Hz to ~160 Hz, but in σ_2 pitch is tightly constrained between 85 Hz and 103 Hz. This suggests a baseline pitch level for this speaker of ~85 Hz. The frame forms define a slightly more diffuse vertical band, with σ_2 pitch ranging from 87 Hz to 114 Hz.

In isolation, the lone N+Vblzr form – ['za:n.za] *zan.za* ‘to eat food’ – plots among the citation forms. In the sentence frame, though, it is distinct from the others, falling much closer to the “zero” line with a pitch difference of only -4 Hz. I have excluded this form from the box-and-whisker plots and the statistical tests below. It is unique not only in its morphology, but also as the only verb token with a long vowel in either syllable (see section 4.5.1.4). It cannot be assumed to manifest the same acoustic properties as the citation forms, and so cannot be grouped together with them. Thus nothing at all can be said about the BSh_03 N+Vblzr verbs.

Figure 4.6 BSh_03 / Verbs: Pitch



The box-and-whisker plots above show that the distinction in pitch across syllables is dramatic, with no overlap of any points in the full range of measured values. Paired-sample t-tests are summarized in Table 4.4 below. For both isolation and frame forms, the difference in pitch is statistically highly significant, with $p \ll 0.05$. It is also perceptually highly salient, averaging -45 Hz for the isolation forms and -35 Hz for the frame forms. Thus pitch is a robust and dramatic cue for σ_1 stress in verbs produced by speaker BSh_03.

Table 4.4 BSh_03 / Verbs / Citation / Pitch: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BSh_03 *	isolation	-24.3064	34	< 2.2e-16	-45	-49	-41
	frame	-22.3284	30	< 2.2e-16	-35	-38	-32

* The lone BSh_03 N+Verbalizer form was excluded.

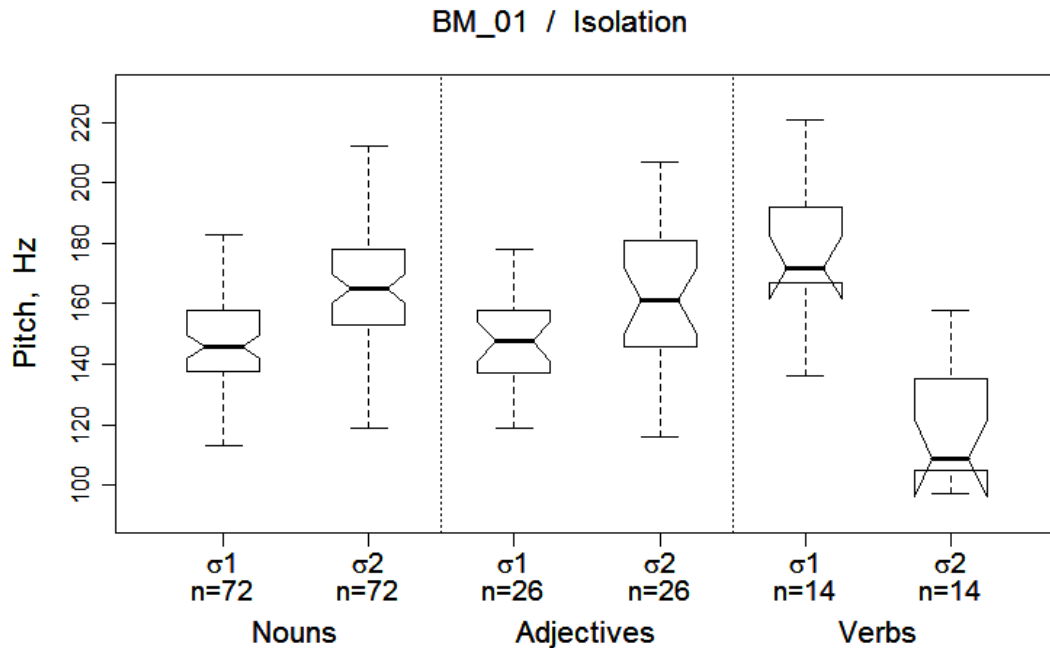
4.2.2 Pitch for speaker BM_01

As noted previously, acoustic parameters show less consistency for speaker BM_01, since target words were sometimes mumbled, drawled, exaggerated, or produced in triumphant contrast to initial “erroneous” responses to a misunderstanding of the word requested during elicitation. Nonetheless, here, too, pitch shows a clear correlation with stress for all lexical categories.

The overall patterns of pitch contrast across syllables are illustrated in Figure 4.7 below. For the non-verbs – stressed on σ_2 – pitch is generally higher on σ_2 . For

the verbs – stressed on $\sigma 1$ – pitch is generally higher on $\sigma 1$. These patterns are discussed in greater detail in sections 4.2.2.1 through 4.2.2.3 below.

Figure 4.7 BM_01 / Pitch contrasts

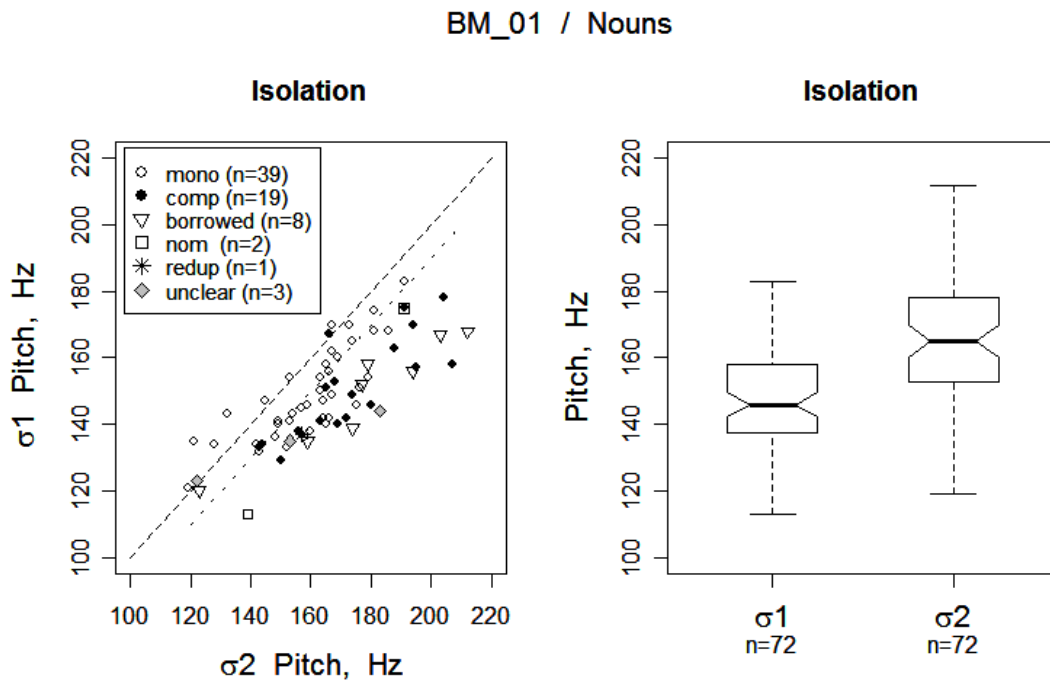


4.2.2.1 BM_01 Nouns

Pitch differences across syllables for BM_01 nouns are illustrated in Figure 4.8 below. Almost all of the nouns have a higher pitch on $\sigma 2$ – the stressed syllable – and thus fall to the right of the dashed “Pitch difference = 0” line. In some of them, though, the pitch difference is not even 10 Hz, as indicated by the dotted reference line. In the box-and-whisker plot to the right, there is a slight overlap of the boxes representing the interquartile ranges on the two syllables. But the notches do not

overlap at all, indicating that there is a significant difference between the median pitch of $\sigma 1$ and the median pitch of $\sigma 2$.

Figure 4.8 BM_01 / Nouns: Pitch



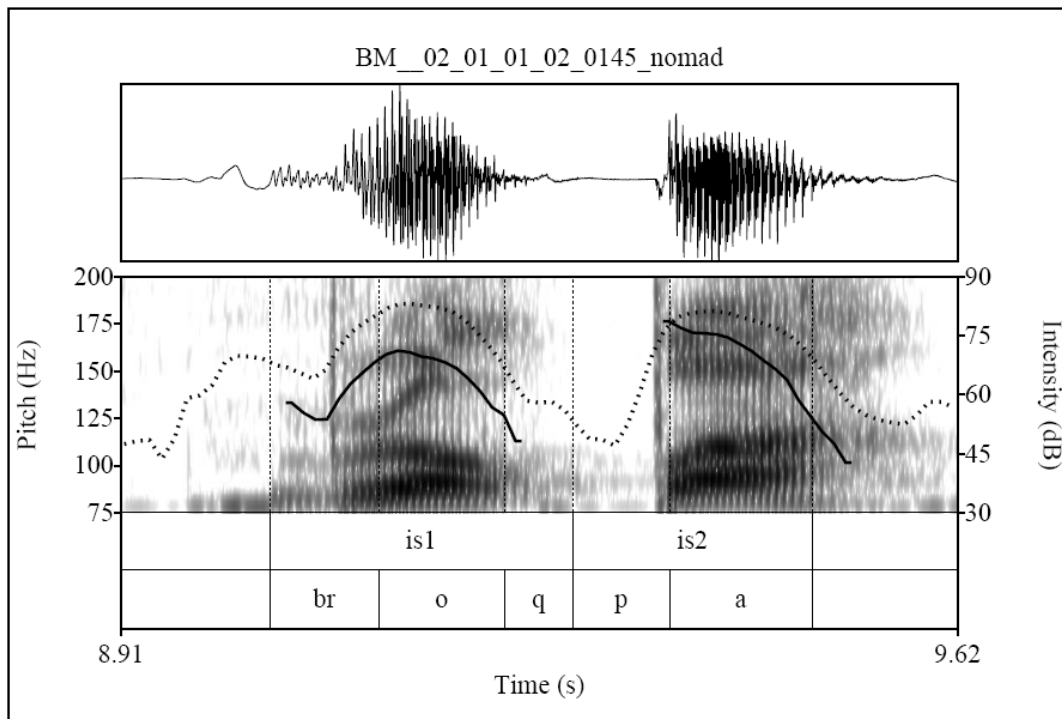
The points in the distribution plot above define a rough linear trend parallel to the dashed line, reflecting variability in the speaker's overall pitch level throughout the recording session, but consistency in the magnitude of the pitch contrast across syllables. That is, if the speaker began a word with a low pitch on $\sigma 1$, then the pitch on $\sigma 2$ would be higher, but still relatively low; if the speaker began a word with a high pitch on $\sigma 1$, then the pitch on $\sigma 2$ would be higher, and also relatively high. For instance, [ʔrya.'ɣan] *dga'.mkhan* 'friend' and [tʃiv.'ɕil] *?.dril?* 'bell' show nearly the same pitch increase across syllables – 21 Hz and 24 Hz respectively – yet fall at very

different positions on the graph: [ʔrya.'χan] at (σ_2 , σ_1) coordinates (150, 129), and [tʃv.'çil] at (194, 170).

One of the factors which influences this distribution is whether the final syllable of the word is open or closed. When σ_2 is open, pitch stays at a high level only briefly, and then arcs downward through the end of the utterance. The pitch trace for the noun [broq.'pa] 'brog.pa 'nomad, pastoralist' in Figure 4.9 below is an example.³⁸ Here, pitch is at ~170 Hz at the start of the vowel, but after ~40 msec near that level, begins to decline. The mean pitch over the medial 50% of the vowel is 163 Hz, and the slope over the vowel was measured as -47 Hz/100msec. Since the pitch in the σ_1 vowel was measured as 154 Hz, the difference in pitch across syllables is only 9 Hz. This noun is represented in the graph above by a point which falls fairly close to the dashed line.

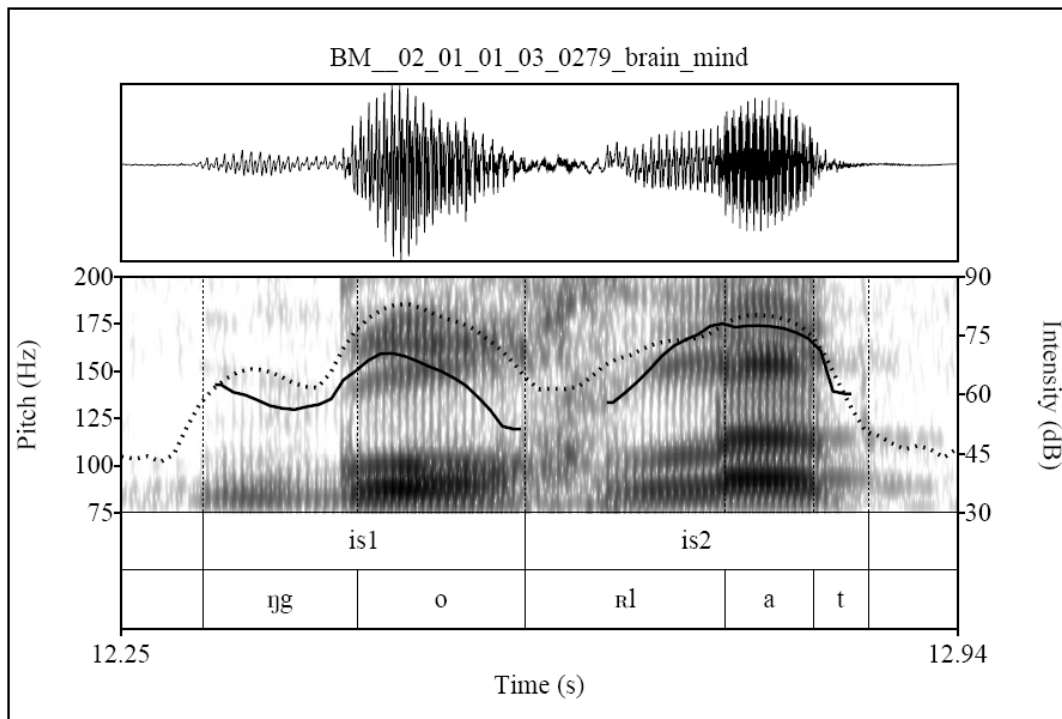
³⁸ In these figures, pitch is represented by a solid line, and intensity by a dashed line.

Figure 4.9 BM_01 / Nouns / Pitch trace: [broq.'pa] 'brog.pa' 'nomad, pastoralist'



In contrast, in the word [ŋgo.'ʌlat] *mgo.glad* 'brain, mind', σ_2 is closed. In this case, as illustrated in Figure 4.10 below, the pitch over the vowel remains high – averaging 174 Hz – and the utterance-final decline does not begin until the coda consonant is reached. Since pitch was measured for the vowel in σ_1 as 149 Hz, the difference in pitch across syllables is 25 Hz.

Figure 4.10 BM_01 / Nouns / Pitch trace: [ŋgo.'ɕlat] *mgo.glad* ‘brain, mind’

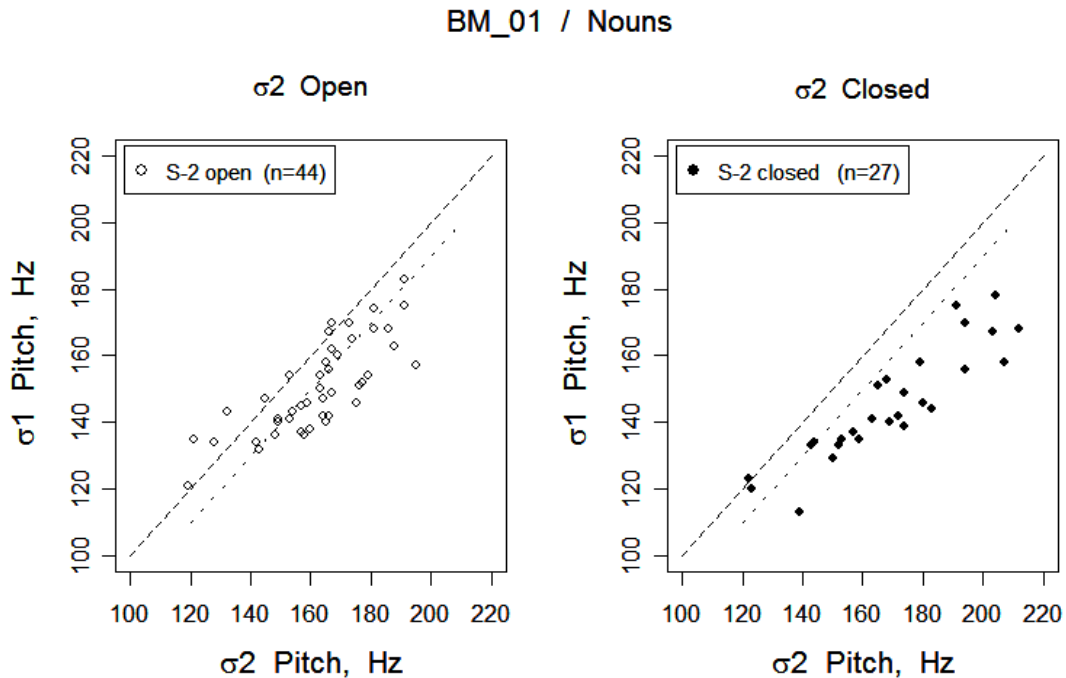


Based on whether σ_2 is open or closed, then, BM_01 nouns behave quite differently in terms of pitch contrast across syllables. The general pattern is illustrated in Figure 4.11 below. When σ_2 is closed, as shown in the plot on the right, nearly all nouns have a pitch contrast of more than 10 Hz. There were only a few exceptions – including [dʒʌŋ.'gɐl] ‘jungle’ and [k^ha.'ɕäɾɔ] ?? ‘mule’ – which plot close to the dashed line. (Pitch is quite low on both syllables of these words, as is intensity; they were probably instances of mumbling.) When σ_2 is open, as shown in the plot on the left, the difference in pitch across syllables is sometimes < 10 Hz, but is often smaller,

and sometimes pitch is even higher on σ_1 , as is the case with [sen.'ma] *sen.mo*

‘finger’, [thu.'lu] *thu.lu* ‘eweskin coat’, and several others.

Figure 4.11 BM_01 / Nouns: Pitch vs. σ_2 closure



Paired-sample t-tests comparing pitch measurements across syllables are summarized in Table 4.5 below. For the complete set of 72 nouns, the contrast in pitch is highly significant, with $p \gg 0.05$, with a mean difference of 16 Hz. If one were to repeatedly record random samples of nouns from speaker BM_01 and compare pitch across syllables for each group, there is a 95% probability that the mean difference would fall between 13 Hz and 19 Hz. When the nouns are considered separately, in terms of σ_2 closure, the pitch contrast is again significant, in both cases but of course the mean difference is much greater when σ_2 is closed (24 Hz) than

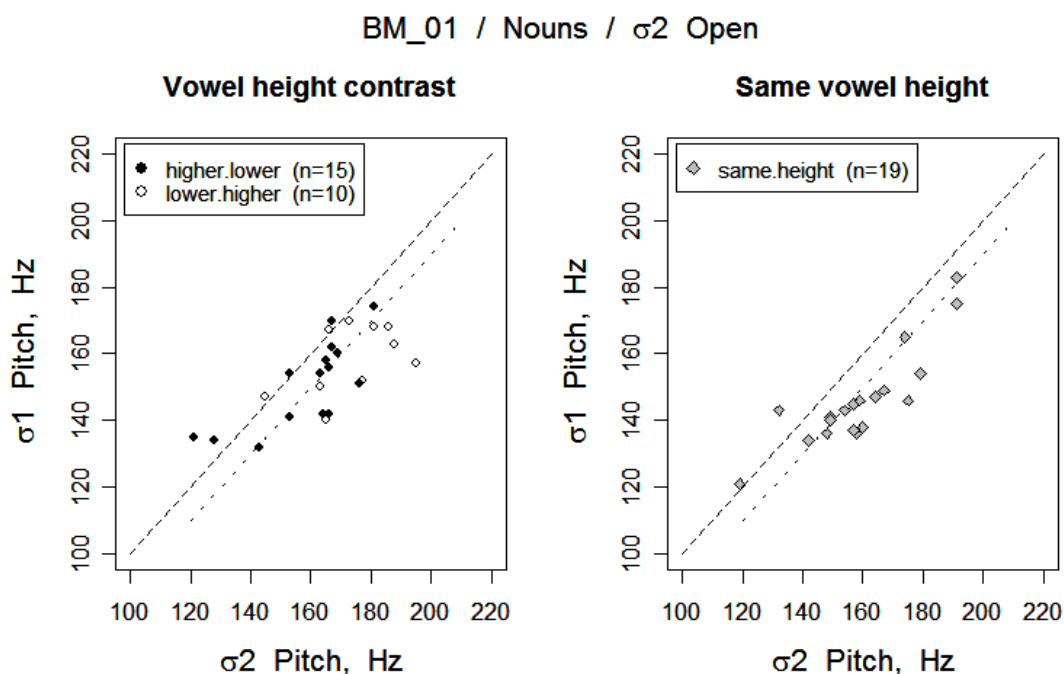
when σ_2 is open (11 Hz). In either case, pitch can be considered a statistically significant correlate of σ_2 stress.

Table 4.5 BM_01 / Nouns / Pitch: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
	all	11.0504	71	< 2.2e-16	16	13	19
BM_01	σ_2 open	7.1379	43	8.129e-09	11	8.5	15
	σ_2 closed	10.4619	26	8.211e-11	24	19	29

Since some of the σ_2 -open nouns show only a small increase in pitch across syllables, it is worth a closer look to determine whether the distribution is further influenced by contrasts in vowel height across syllables. As discussed in section 2.3.2.1, Lehiste (1970) has observed that high vowels tend to have an intrinsically high fundamental frequency, and low vowels tend to have an intrinsically low fundamental frequency. In Figure 4.12 below, the σ_2 -open nouns are plotted in terms of vowel height. If vowel height played a dominating role here, in the plot on the left we would expect to see a very clear distinction between [higher.lower] (such as [p^huŋ.'ma] *phying.pa* ‘wool felt’) and [lower.higher] nouns (such as [çoy.'vu] *shog.bu* ‘book, paper’), which does not appear to be the case.

Figure 4.12 BM_01 / Nouns / σ 2 Open: Pitch vs. vowel height



When vowel height is controlled for in the σ 2-open nouns (as in [tɕ^ho.'lo] *cho.lo* ‘dice’, the results of a paired-sample t-test, summarized in Table 4.6 below, are not much different than those for the full set of nouns with σ 2 open (see Table 4.5); the contrast in pitch is statistically significant, but sometimes not of very great magnitude.

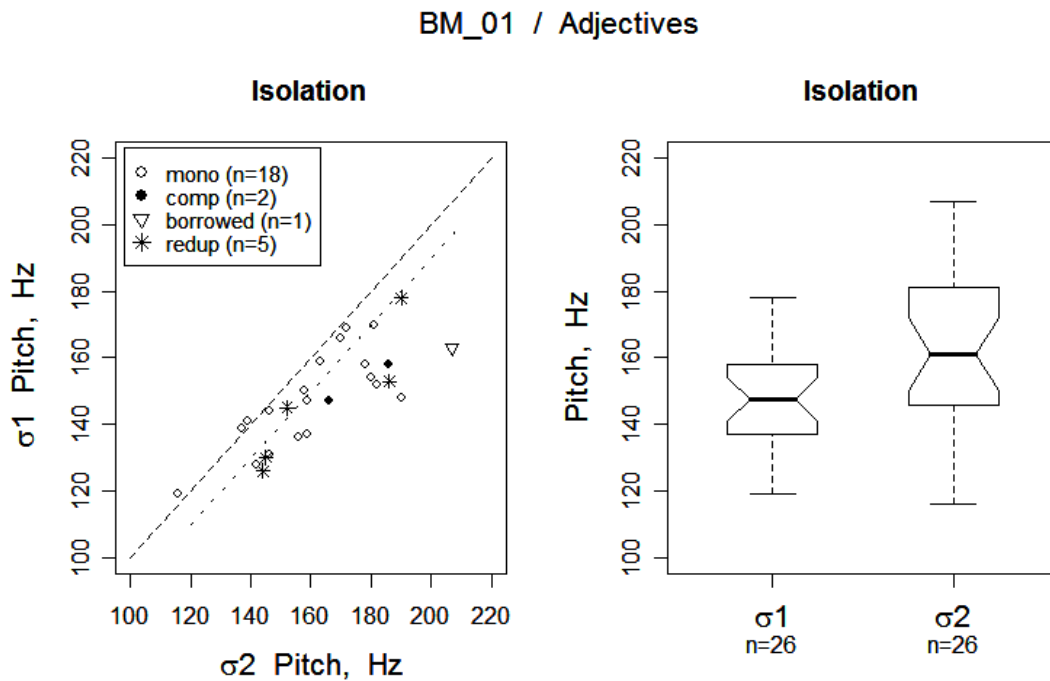
Table 4.6 BM_01 / Nouns / σ 2 Open /Pitch: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BM_01	same height	6.0334	18	1.053e-05	13	8.4	17

4.2.2.2 BM_01 Adjectives

Pitch is also a consistent correlate of stress in adjectives produced by speaker BM_01. In the distribution plot in Figure 4.13 below, nearly all points fall to the right of the dashed line. Again, though, for many of the monomorphemic forms, the increase in pitch across syllables is < 10 Hz.

Figure 4.13 BM_01 / Adjectives: Pitch



The overlap of the boxes and whiskers in the plot to the right above indicates that the aggregate of values measured on the two syllables overlap in range. The notches overlap slightly, as well, meaning that the median values for the two syllables are not significantly different.

However, when paired values are considered – i.e., when σ_1 and σ_2 pitch measurements from each word are compared – the contrast in pitch is found to be

consistent. As shown in Table 4.7 below, $p \ll 0.05$ and the mean difference across syllables is 15 Hz. If samples were repeatedly recorded and analyzed, there is a 95% probability each time that the mean pitch difference would fall between 10 and 21 Hz.

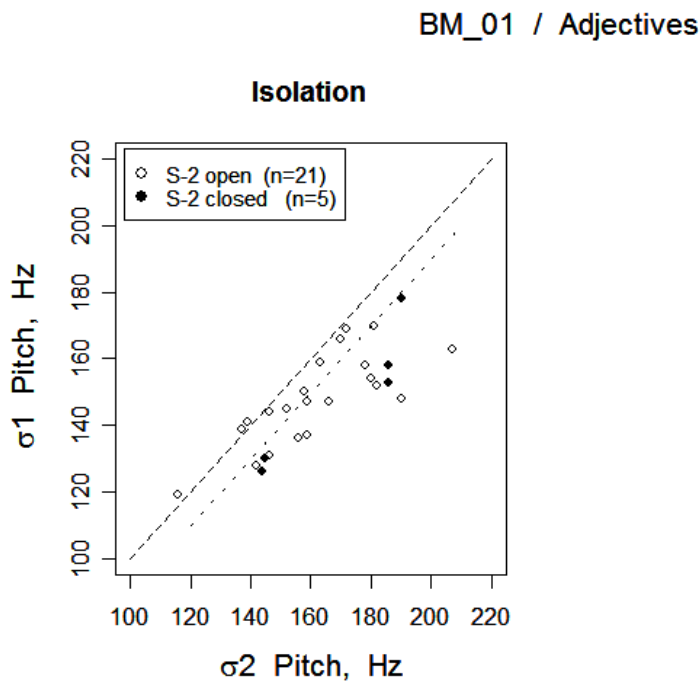
Table 4.7 BM_01 / Adjectives / Pitch: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BM_01	isolation	6.132	25	2.071e-06	15	10	21

What can be concluded from this analysis is that pitch is a statistically significant correlate of the perceived σ_2 stress pattern, but the contrast in pitch is not always particularly strong.

As was the case with nouns, some elements of the behavior of pitch in adjectives can be accounted for by considering whether σ_2 is open or closed. There are only five adjectives in which σ_2 is closed, but – as shown in Figure 4.14 below – the difference in pitch across syllables for all of them is > 10 Hz.

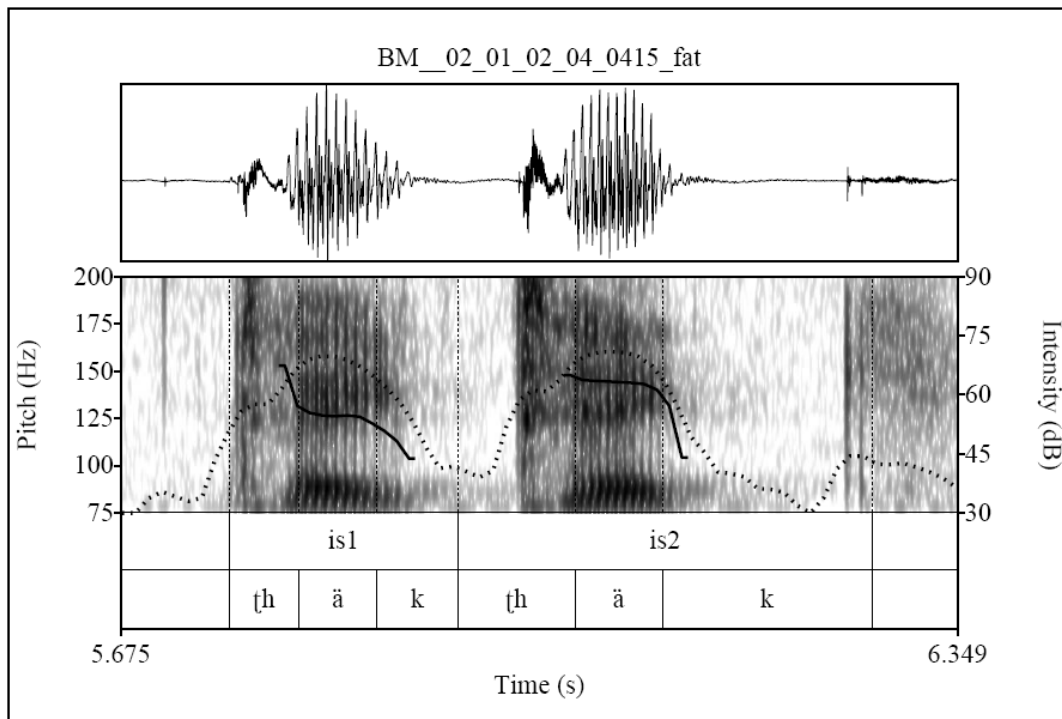
Figure 4.14 BM_01 / Adjectives: Pitch vs. $\sigma 2$ closure



Four of these are the reduplicated forms; the fifth is the color term [sɲo.'raŋ] *sngo.rang* ‘blue color’³⁹. In these cases, pitch is apparently elevated in the $\sigma 2$ -closed adjectives because the coda cuts off the terminal falling intonation contour. This is illustrated in the pitch trace for [tʰäk.'tʰäk] *mthag.mthag* ? ‘fat’, in Figure 4.15 below.

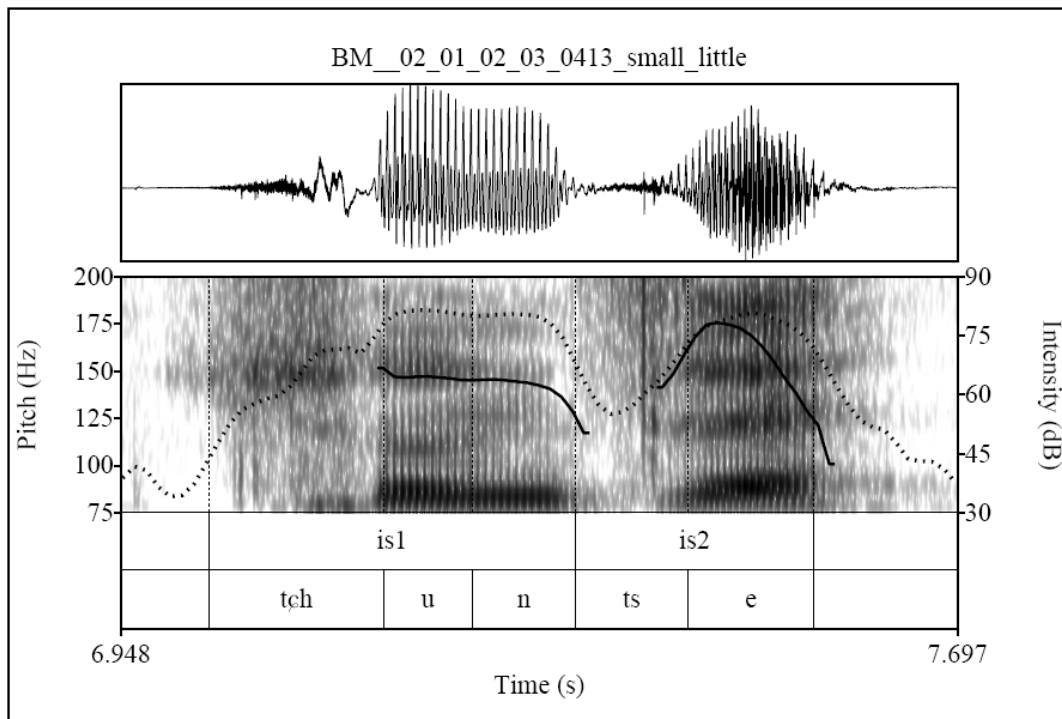
³⁹ *rang* ‘color’ is borrowed from Urdu. (Bielmeier p.c., 2008).

Figure 4.15 BM_01 / Adjectives / Pitch trace: [tʰäk.'tʰäk] *mthag.mthag?* ‘fat’



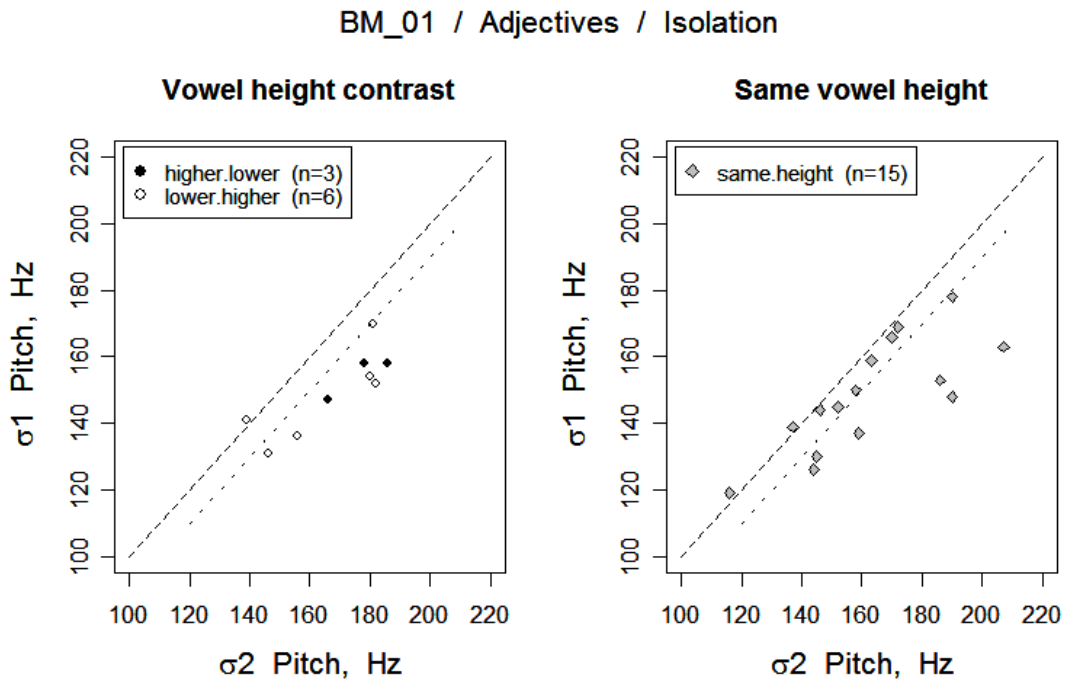
In contrast, when σ_2 is open, pitch reaches a peak value and then declines steadily, as illustrated by [tɕʰun.'tse] *chung.tshad?* ‘small, little’ in Figure 4.16 below.

Figure 4.16 BM_01 / Adjectives / Pitch trace: [tɕ^hun.'tse] *chung.tshad?* ‘small, little’



The intrinsic variation of pitch as a function of vowel height, on the other hand, does not appear to have a strong influence on pitch contrasts. In Figure 4.17 below, the distribution of points is about the same regardless of whether pitch is higher in one syllable or the other, or the same in both syllables.

Figure 4.17 BM_01 / Adjectives: Pitch vs. vowel height contrast



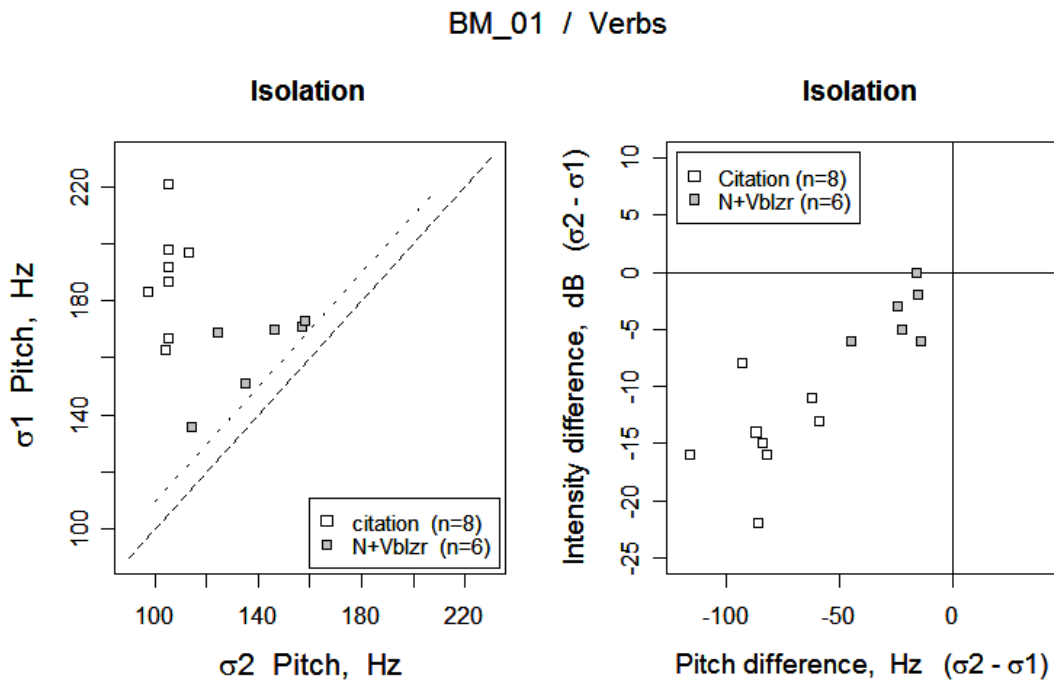
4.2.2.3 BM_01 Verbs

Pitch is also a consistent correlate of the $\sigma 1$ stress perceived on BM_01 verbs. In this case, there is sufficient evidence to demonstrate a distinction between morphological types.

In the graph on the left side of Figure 4.18 below, all verbs have a higher pitch on $\sigma 1$ and so plot to the left of the dashed line. The citation forms and N+Vblzr forms seem to define different patterns. The citation forms define a vertical band: pitch in $\sigma 1$ varies widely, ranging from ~160 Hz to ~220 Hz, while pitch in $\sigma 2$ always falls within a narrow range, between ~100 Hz and 115 Hz; this may be near the speaker's baseline pitch level. The distribution of the N+Vblzr forms, in contrast, roughly parallels the dashed "Pitch difference = 0" line. As discussed for BM_01 nouns

(Figure 4.8), this means that the relative contrast in pitch across syllables is kept fairly constant, even when the absolute values differ from word to word.

Figure 4.18 BM_01 / Verbs: Pitch



The distinction between citation forms and N+Vblzr forms is confirmed by the graph on the right above. Here, the difference in pitch across syllables is plotted against the difference in intensity across syllables. Intensity will be considered in greater detail in section 4.4.2, but I include this plot here to demonstrate that the two morphological types occupy very different acoustic spaces. This acoustic distinction is confirmed by the Welch t-tests summarized in Table 4.8 below. As shown, pitch declines, on average, a spectacular 84 Hz from σ_1 to σ_2 in the citation forms. In the N+Vblzr forms, the average pitch drop from σ_1 to σ_2 is 22 Hz, which is considerably less, though still highly salient. When these mean pitch differences are compared

using a Welch t-test, the fact that the p-value $\ll 0.05$ indicates that there is a 95% certainty that these means do *not* represent samples from a single, common population. The confidence limits indicate that, under repeated sampling, the mean difference between these mean differences will be between 78 Hz and 43 Hz.

Similar results are observed for intensity. For the citation forms, intensity drops 14 dB between σ_1 and σ_2 ; for N+Vblzr forms, the drop is ~ 4 dB. This difference in differences is highly significant, since $p \ll 0.05$. Given the results of these t-tests, BM_01 citation forms and N+Vblzr forms are treated separately in the graphs and analyses which follow.

Table 4.8 BM_01 / Verbs: t-tests comparing stress correlates vs. morphological category (Welch)

Stress correlate	t	DF *	p-value	Citation mean	N+Vblzr mean	95 % conf. limits	
						lower	upper
Pitch	-7.7129	11.857	5.869e-06	-84	-23	-78	-44
Intensity	-6.099	11.53	6.34e-05	-14	-3.7	-14	-6.9

* In the Welch two-sample test, the ‘degrees of freedom’ does not always turn out to be an integer.

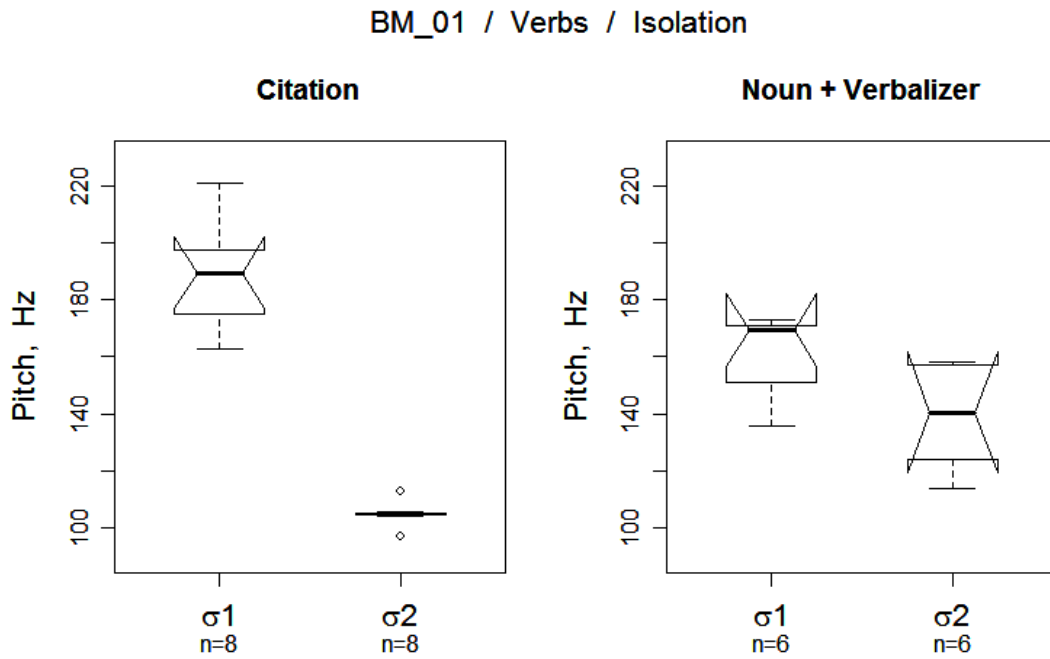
An explanation for the distinct acoustic behavior of these two types of verbs may lie in the difference in semantic content across syllables. In the citation forms, all of the important semantic information is carried in σ_1 (as was the case with monomorphemic nouns and adjectives). Since σ_2 is always a variant of [-pa] or [-ma], it can be muted or reduced. The only real role of this syllable, it seems, is to make clear to the listener that the word is structurally and phonologically a verb. To achieve this, σ_2 need only have a distinctly lower pitch and intensity than σ_1 . The

segmental content is less important than the acoustic contrast, which the speaker thus emphasizes. In the case of the N+Vblzr forms, both syllables have meaning, both syllables have important segmental content, and so both must be articulated fully. Since both syllables are thus produced with their own strong pitch and intensity, the contrast in these parameters across syllables is reduced and the points plot closer to the origin.

The contrasts in pitch across syllables for the two morphological types of verbs are illustrated by boxplots in Figure 4.19 below. For citation forms, pitch values do not overlap at all. As noted above, the eight pitch measurements on σ_2 fall within a very narrow range, which may represent this speaker's baseline pitch level.

In the N+Vblzr forms, in contrast, the aggregate of values measured independently on the two syllables do show some overlap, as do the notches representing the 95% confidence interval about the median. However, this does not reflect the contrast in values paired across the two syllables of individual words. (Besides, we already know from the graphs in Figure 4.18 that pitch is always higher in σ_1 .)

Figure 4.19 BM_01 / Verbs: Pitch contrasts vs. morphological category



As shown in the table below, paired-sample t-testing confirms that the difference in pitch across syllables for BM_01 N+Vblzr forms is indeed significant, with $p < 0.05$. Even the smallest probable mean decrease in pitch across syllables of just 10 Hz – the upper 95% confidence limit – likely constitutes a distinct enough pitch fall for a listener to be confident that this is a verb, rather than a noun or adjective. For the citation forms, too, $p \ll 0.05$, and the mean pitch difference is -84 Hz.

Table 4.9 BM_01 / Verbs / Pitch: Results of paired-sample t-tests (two-tailed)

Speaker	subset	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
BM_01	citation	-13.2523	7	3.258e-06	-84	-98	-69
	N+Vblzr	-68.70	5	0.005045	-23	-35	-10

4.3 Pitch slope

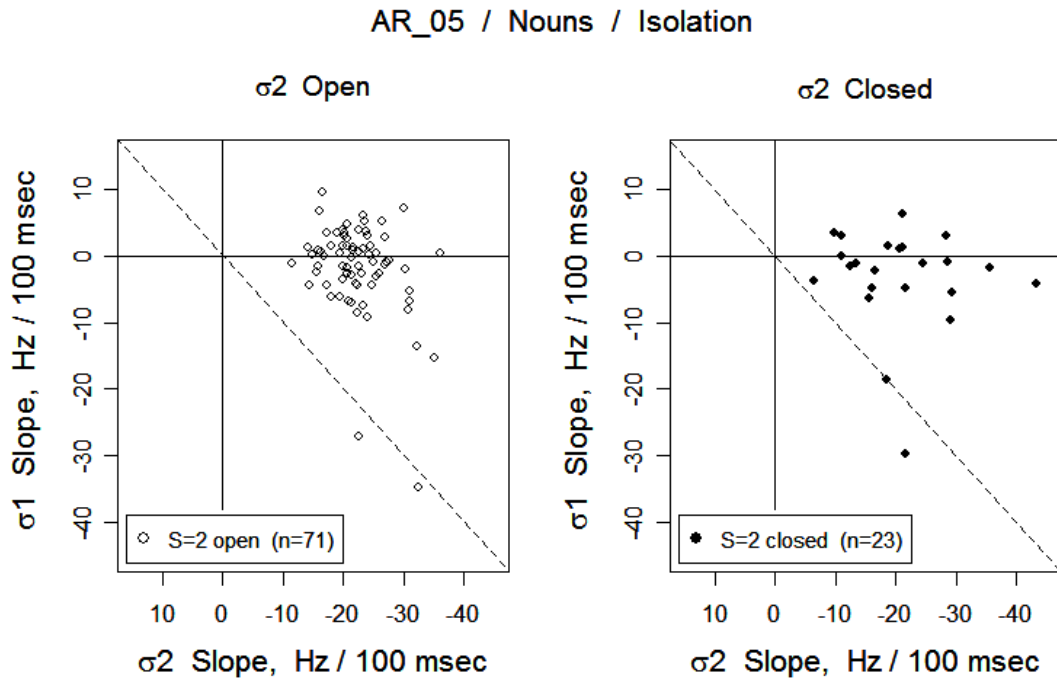
Pitch slope is not an acoustic correlate of stress for either Balti speaker, for words of any lexical category.

This fact can be most effectively established by previewing some of the data from Rebkong Amdo Tibetan, which is discussed in greater detail in Chapter 5. In Rebkong Amdo, pitch slope is a primary cue for σ_2 stress in nouns produced in isolation. Upon first hearing this dialect in the field, I was immediately struck by this characteristic.

Figure 4.20 below shows pitch slope measurements for isolation forms of nouns produced by Rebkong Amdo speaker AR_05, coded in terms of whether σ_2 is open or closed. The graph reveals a high degree of consistency, especially for the nouns with σ_2 open. What this shows is that the speaker manipulates pitch slope with some precision. In fact, the clustering of points suggests that the speaker is aiming towards a target: a flat or very gentle slope on σ_1 , and a slope of ~ -20 Hz/100msec in σ_2 .

When σ_2 is closed, some portion of the downward slope is borne by the coda, so there is greater variability in the duration and degree of slope which occurs within the σ_2 nuclear vowel.

Figure 4.20 AR_05 / Nouns / Isolation: Pitch slope



Thus when pitch slope is a robust correlate of stress – as it is in Rebkong Amdo – there is a distinctive and consistent relationship between the slopes on the two syllables. In words which plot above the x-axis in Figure 4.20, pitch in σ_1 slopes upward, and pitch in σ_2 slopes downward. In words which plot below the x-axis – but still to the right of the dashed line – pitch in σ_1 slopes downward, and pitch in σ_2 slopes downward even more steeply. In either case, pitch in σ_2 is “more downward”.

In Balti, there are, indeed, differences in pitch slope across the two syllables of a word, and these differences are often statistically significant and in many cases probably also perceptually significant. However, they are not contextually significant. For speaker BSh_03, the pitch slope pattern is not distinct from the pitch pattern specific to each lexical category. Thus the plots illustrating pitch slope measurements for this speaker – presented in section 4.3.1 below – bear little resemblance to the plots for Rebkong Amdo speaker AR_05, above. For speaker BM_01, the pitch slope pattern is not distinct from the intonation pattern, which is the same in all lexical categories. Plots illustrating pitch slope measurements are actually quite similar to those for speaker AR_05, though they diverge in some respects.

Since pitch slope for the two Balti speakers is somewhat variable and is influenced by a number of factors, the discussion below is more descriptive than analytical. Results for speaker BSh_03 are presented in section 4.3.1, followed by results for speaker BM_01 in section 4.3.2.

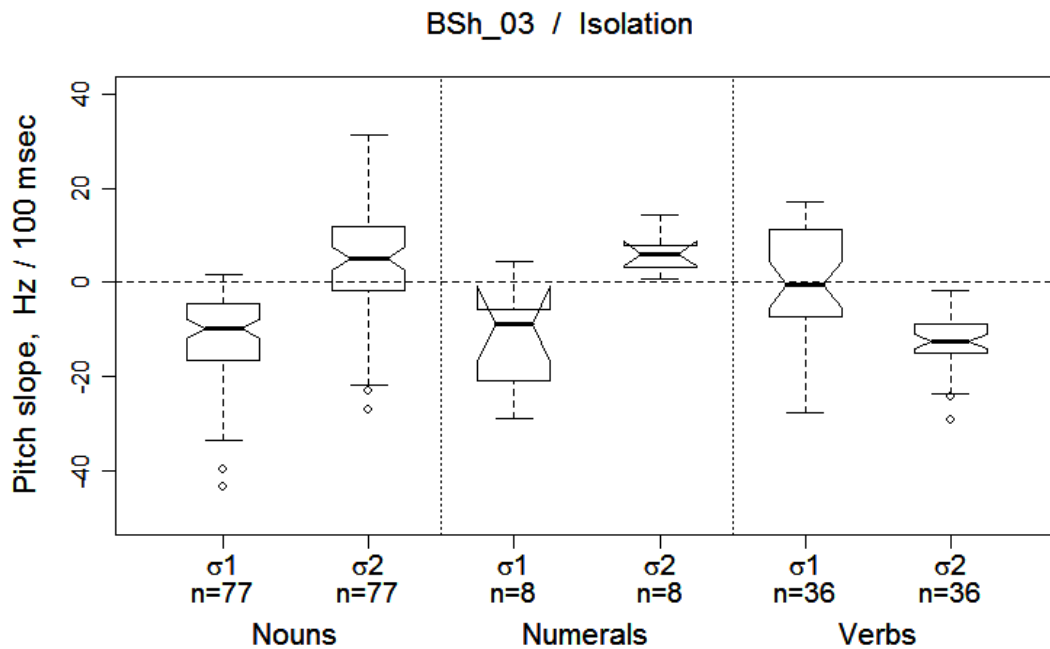
4.3.1 Pitch slope for speaker BSh_03

Pitch slope contrasts for the isolation forms of BSh_03 nouns, numerals, and verbs are illustrated in the box-and-whisker plots in Figure 4.21 below. (Adjectives are not included here, but are addressed separately in section 4.3.1.2.) It is important here not to focus only on the trends suggested by the median values, but to bear in mind the distribution represented: each hinge of the box represents 25% of the measurements, as does each whisker. It is also important to recall – as discussed in section 2.5.1.1 – that when a box-and-whisker plot indicates that values on the two

syllables of a set of words are different, then indeed, they are different. But when a box-and-whisker plot does not indicate a clear difference, this means that the range of the aggregate of values on each syllable is similar, but a more careful examination of paired values is required.

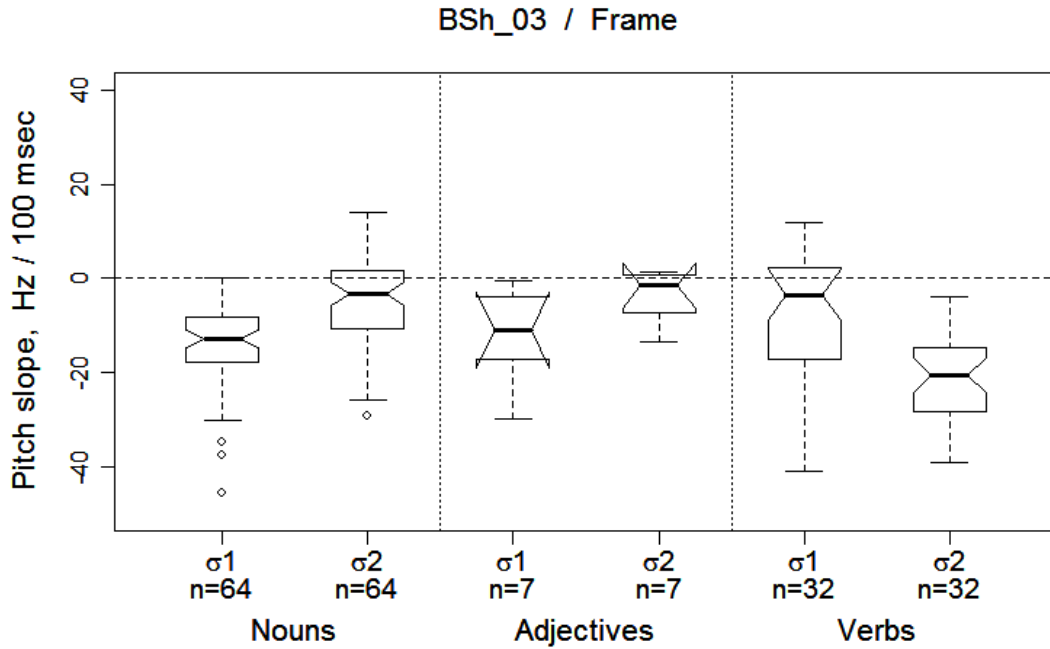
It is clear here that non-verbs and verbs exhibit distinctly different patterns. For nouns, pitch on σ_1 almost always slopes downward, while pitch on σ_2 usually slopes upward but sometimes slopes downward. For numerals, pitch on σ_1 is usually downward, and pitch on σ_2 is always upward. In contrast, for verbs the pitch on σ_1 slopes upward half the time and downward half the time; the median value indicates a level pitch. For σ_2 , pitch always slopes downward.

Figure 4.21 BSh_03 / Isolation: Pitch slope contrasts



The BSh_03 frame forms show similar but more subdued patterns, as illustrated in Figure 4.22 below.

Figure 4.22 BSh_03 / Frame: Pitch slope contrasts



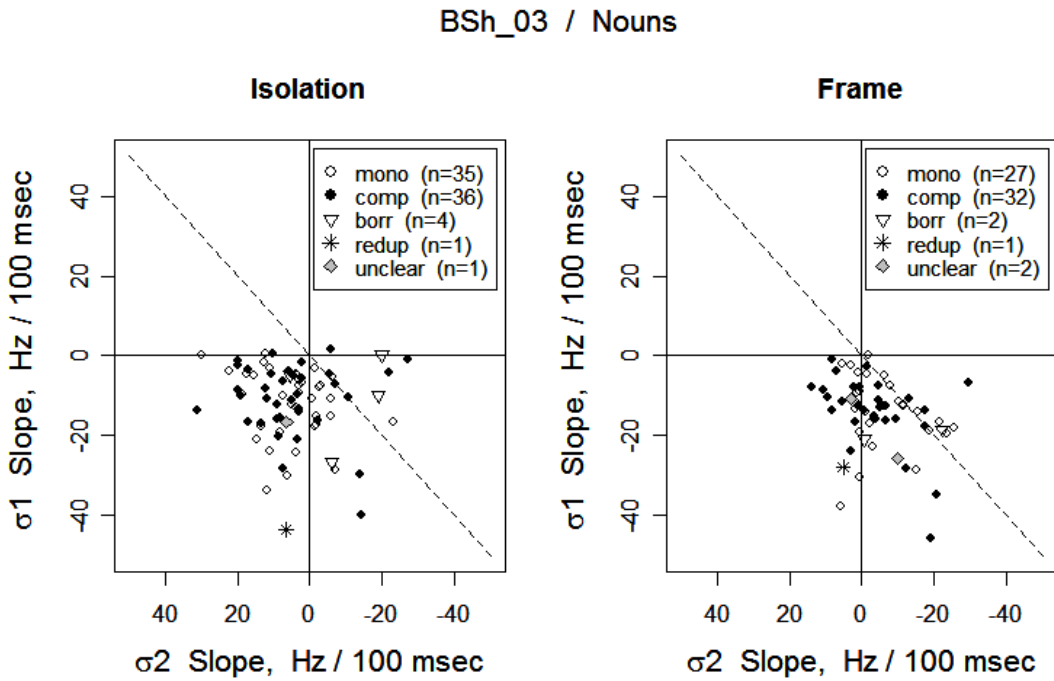
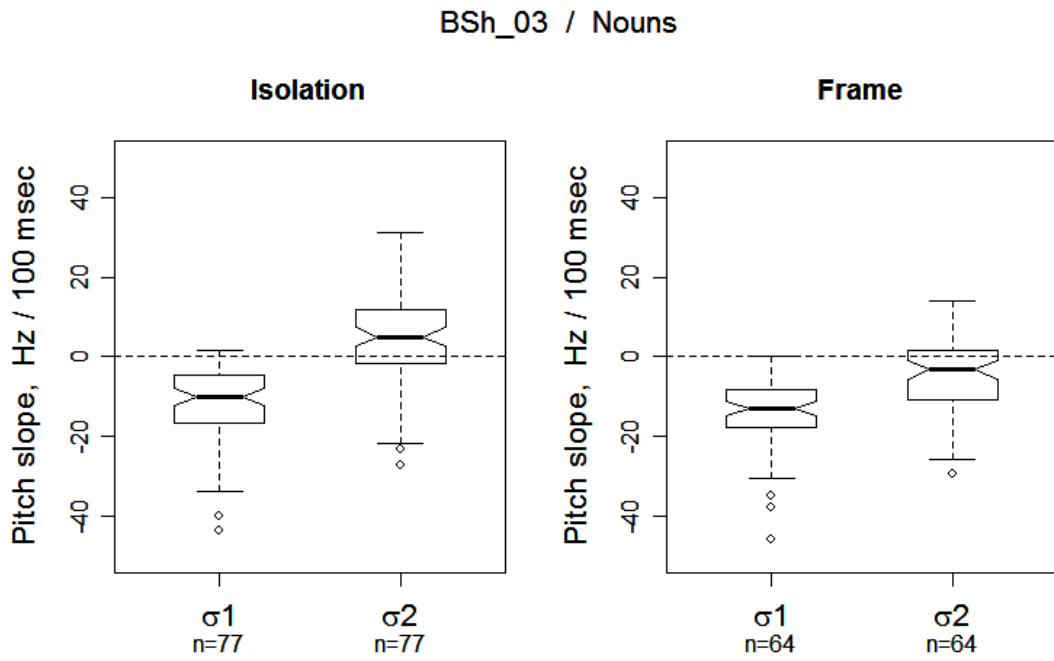
As demonstrated for each lexical category in sections 4.3.1.1 through 4.3.1.4, the pitch slope patterns observed are a by-product of the pitch contrasts across syllables, and of the utterance level intonation patterns. There is no independent correlation between pitch slope and stress. In the sections below, I describe the observed patterns and attempt to explain their origins.

4.3.1.1 BSh_03 Nouns

Pitch slope measurements for isolation and frame forms of BSh_03 nouns are illustrated in Figure 4.23 below. The solid lines which mark the median values in the

box-and-whisker plots helpfully depict the dominant trends. For the isolation forms, in general, pitch slopes downward in σ_1 , and slopes upward in σ_2 . For the frame forms, in general, pitch slopes downward in σ_1 , and also slopes downward in σ_2 , but not as steeply. The whiskers represent 25% of the values measured on each syllable. For σ_2 in the isolation forms, the whisker extending below the “0” line represents the minority of cases in which the pitch slope is downward. For σ_2 in the frame forms, the whisker extending above the “0” line represents the minority of cases in which the pitch slope is upward.

Figure 4.23 BSh_03 / Nouns: Pitch slope



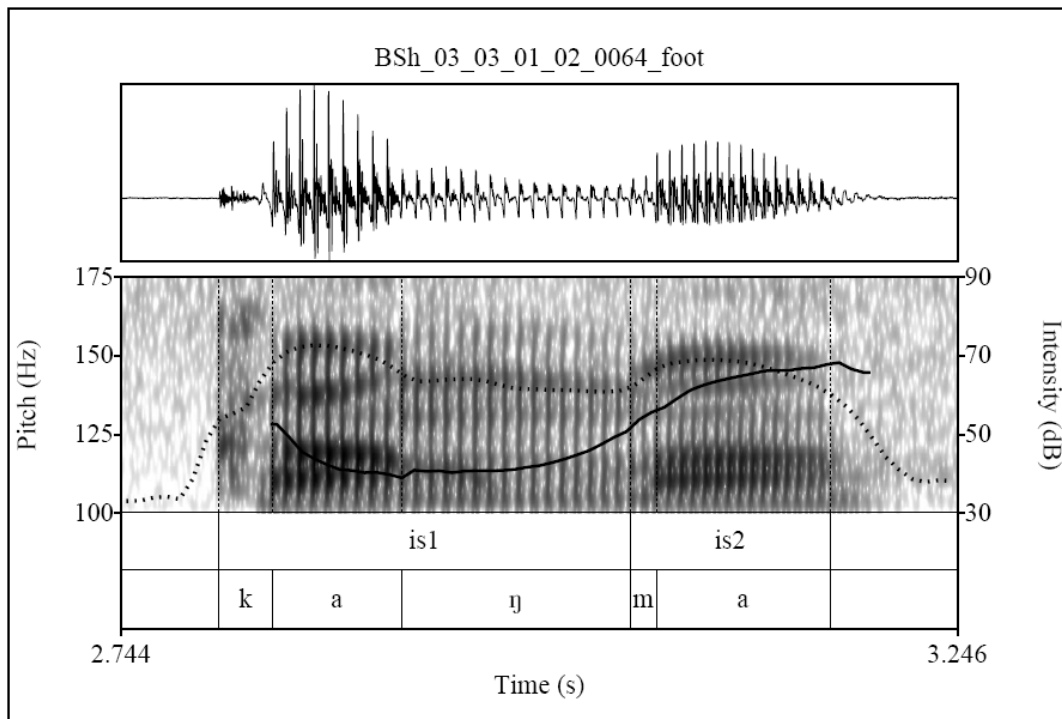
In the bottom part of the figure, all points fall very close to or below the x-axis, indicating that the σ_1 pitch contour is nearly flat or slopes downward. Points falling to the left of the y-axis have an upward slope on σ_2 ; points falling to the right of the y-axis have a downward slope on σ_2 . The contrast with the pattern for Amdo is immediately clear: in Figure 4.20, most points fall to the right of the dashed line; here, most points fall to the left of the dashed line. This reflects the fact that the vector representing pitch on σ_2 is “more upward” than that representing σ_1 .

It is clear from these plots that the two syllables do indeed exhibit different slope patterns. (If there were no difference in behavior across syllables, all points would plot along or close to the dashed line – as some among the frame forms do.) However, this difference does not reflect a correlation with the syntagmatic stress contrast. Rather, for the isolation forms, the slope pattern is a natural outcome of the pattern of pitch contrasts across syllables. We already know (section 4.2.1.1) that pitch is a robust correlate of σ_2 stress in BSh_03 nouns: the mean increase in pitch from σ_1 to σ_2 is 23 Hz for isolation forms, and 20 Hz for frame forms. (The medians are 23 Hz and 19 Hz, respectively.) For the frame forms, the slope pattern is a natural outcome of the interaction of this lexical pitch pattern with utterance-level intonation.

Thus for both isolation and frame forms, pitch slope is an epiphenomenon of other acoustic features. It lacks contextual significance; this is why it is not an acoustic correlate of stress.

For the isolation forms, the dependence of the pitch slope pattern on the pitch pattern can be illustrated by the case of the monomorphemic noun [kaŋ.'ma] *rkang.ma* ‘foot, leg’. The pitch trace for this word is shown in Figure 4.28 below. Here, the mean pitch on σ_1 is 115 Hz, the mean pitch on σ_2 is 143 Hz, and the σ_2 - σ_1 difference is $(143 - 115) = 28$ Hz. This difference is quite close to the speaker’s “target” difference of 23 Hz (the mean / median). Since the consonants at the syllable boundary are voiced sonorants, the pitch contour can be followed continuously from one syllable to the next, and we can see that the upward pitch slope of the vowel in σ_2 is simply part of the increase in pitch across syllables. The rate of rise in the σ_2 vowel is 13 Hz/100msec. Since the full span of the vowel is 104 msec, the pitch increase over the whole vowel is almost exactly 13 Hz.

Figure 4.24 BSh_03 / Noun / Pitch trace: [kaŋ.'ma] *rkang.ma* ‘foot, leg’

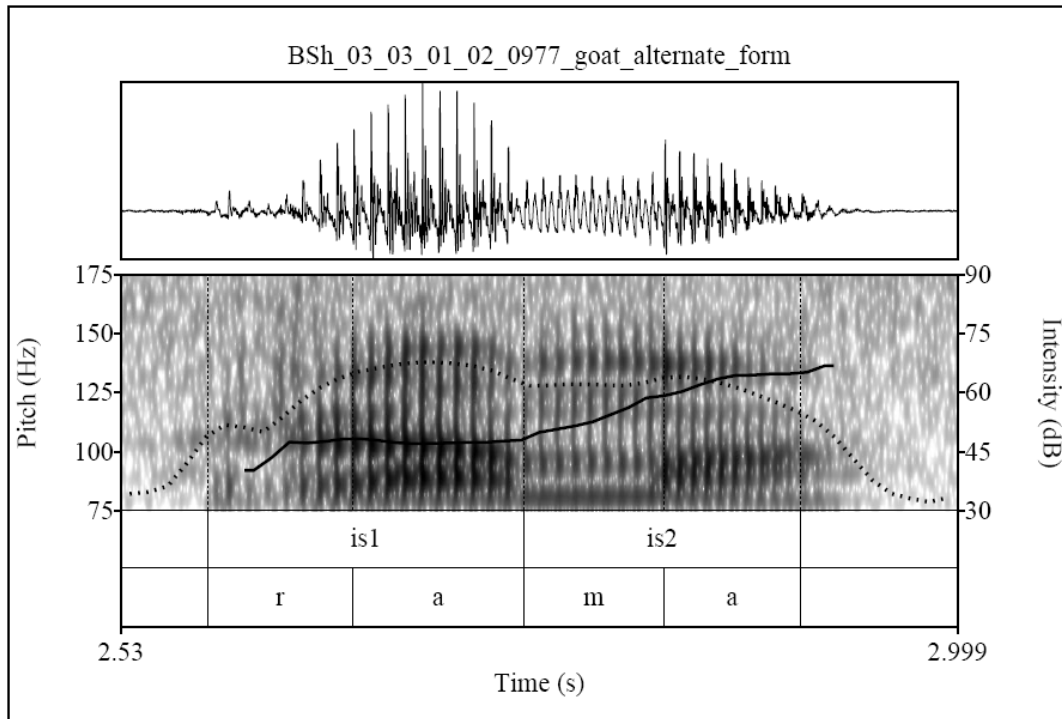


For the vowel in σ_1 , pitch slope was measured as -18 Hz/100msec. (The duration of the vowel here is 77 msec, so the actual drop in pitch is proportional to that span.) This is a typical example: as noted above, pitch is nearly level or slopes downward on σ_1 for *all* BSh_03 nouns produced in isolation. I believe this downward slope occurs because, once the speaker has begun to produce a word, he immediately adjusts his pitch towards a target level, as necessary. This target level must be one which will comfortably allow the speaker to achieve the target pitch increase of ~ 23 Hz across syllables. A value of ~108 Hz may be regarded as the target value for σ_1 ; this is the median of the σ_1 pitch measurements plotted in Figure 4.3. In [kaŋ.'ma] *rkang.pa* ‘foot, leg’, the pitch at the very beginning of the vowel is 126 Hz. The

speaker makes a quick downward adjustment to achieve a mean pitch value of 115 Hz across the medial 50% of the vowel. (The mean pitch value across the stable, medial portion of the rhyme is 113 Hz.) Thus the downward pitch slope on σ_1 – like the upward pitch slope on σ_2 – has nothing to do with any kind of independent correlation between pitch slope and stress, but reflects instead the role of pitch as a cue for stress. The reason pitch never slopes upward on the σ_1 vowel is that the target pitch is low – a pitch level that will be in contrast to the high target of σ_2 .

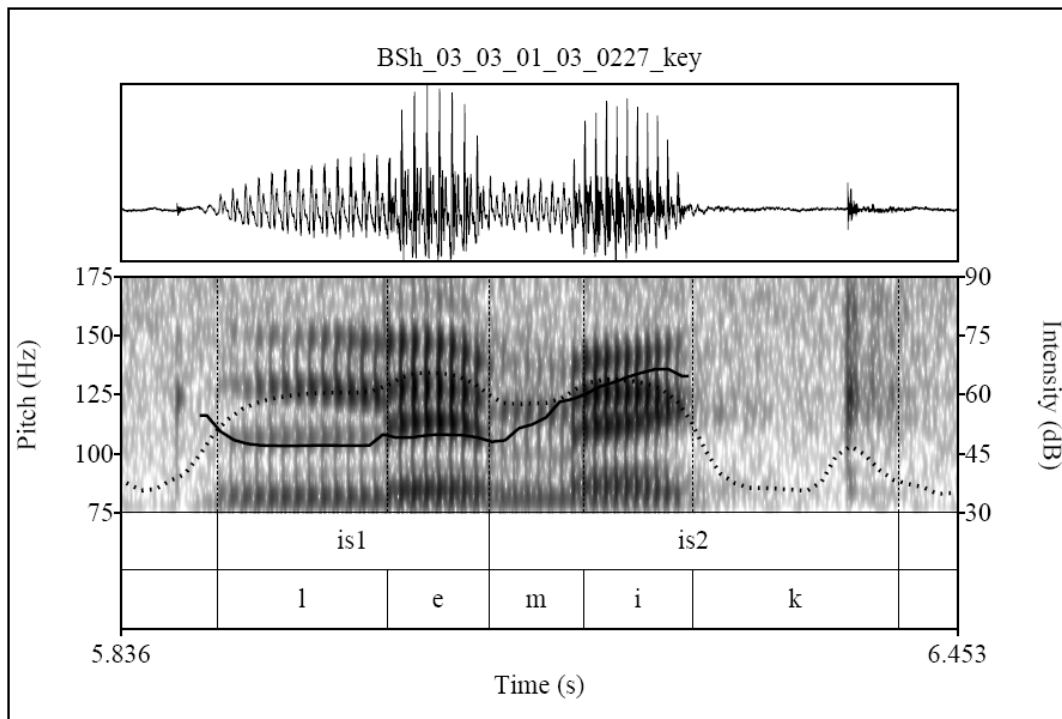
The pitch trace for the word [ra.'ma] *ra.ma* 'goat' supports this hypothesis. As illustrated in Figure 4.25 below, in this case the σ_1 slope is only 0.52 Hz/100msec. Since the speaker started the utterance at a very low pitch – 92 Hz at the very beginning – he needed only to make a quick upward adjustment towards the target, and from there maintained a steady pitch of 104 Hz across the onset and nucleus. The pitch slope for σ_2 was measured as 12 Hz/100 msec, and the difference in slopes for this word is approximately $12 - 0.52 = 11.5$ Hz/100 msec.

Figure 4.25 BSh_03 / Noun / Pitch trace: [ra.'ma] *ra.ma* 'goat'



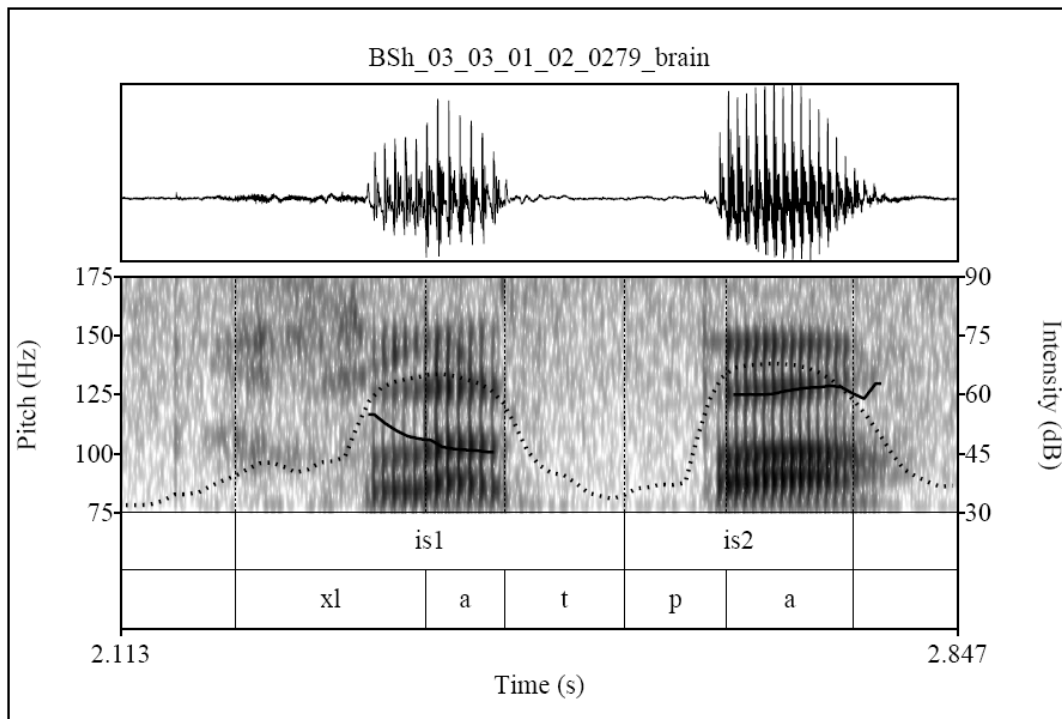
The patterns described above occur just as readily in nouns where σ_2 is closed, as illustrated by the pitch trace for the compound form [le.'mik] *lde.mig* 'key' in Figure 4.26 below.

Figure 4.26 BSh_03 / Noun / Pitch trace: [le.'mik] *lde.mig* 'key'



When the σ_2 onset is not pitch-bearing – e.g., a voiceless stop – a slightly different pattern is sometimes encountered. For example, in [xlat.'pa] *klad.pa* / *glad.pa* 'brain' in Figure 4.27 below, the pitch trace is not continuous. After the stops at the syllable boundary, pitch does not climb gradually but begins and remains at a relatively high level across the vowel in σ_2 (125 Hz) with a gentle slope (2.9 Hz/100msec).

Figure 4.27 BSh_03 / Noun / Pitch trace: [xlat.'pa] *klad.pa* / *glad.pa* 'brain'

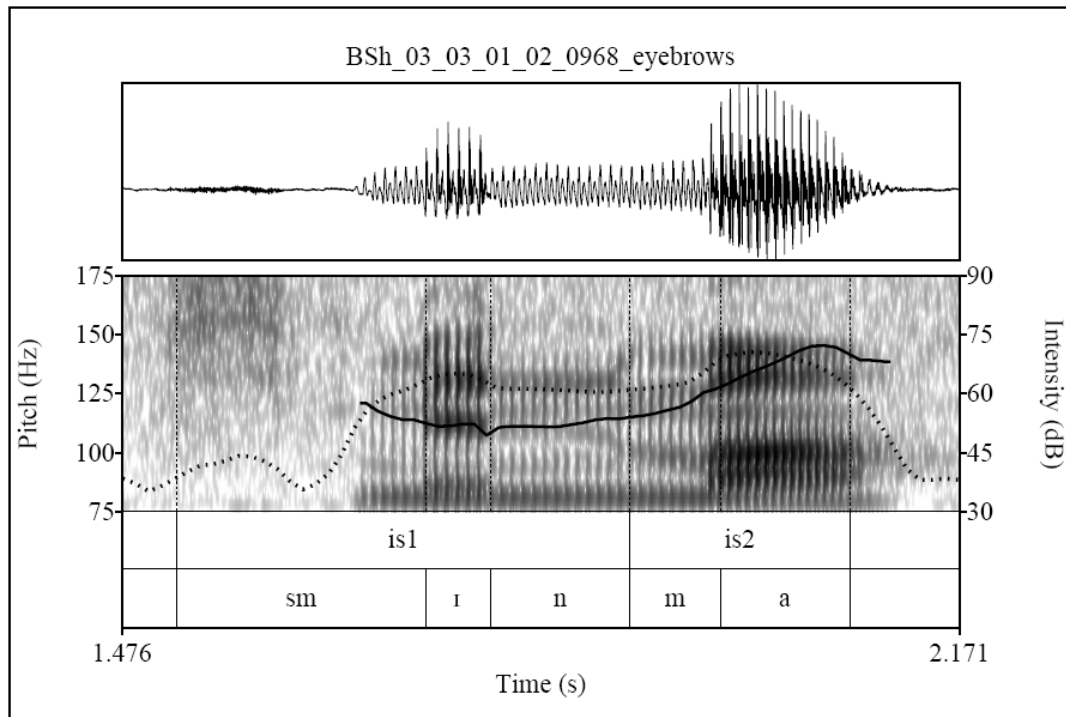


In all of the pitch traces provided above, it is interesting to note the absence of the terminal falling intonation contour that one would expect to see in words produced in isolation. I believe this absence reflects the speaker’s deliberate care to convey the low-high pitch contrast which, as discussed in section 4.2.1, is the primary correlate of σ_2 stress.⁴⁰

⁴⁰ I observed this same phenomenon with another Balti speaker, who – like speaker BSh_03 – was highly educated and very serious about providing “accurate” data. In that case, I worked with the speaker – BSk_02, from the town of Skardu – by first eliciting, discussing, and transcribing my word list. Recording was then completed in a separate, “more efficient” session. This recording seemed self-conscious and unnatural to me, as I noted the lack of a falling intonation contour at the end of words produced in isolation. With speaker BSk_02 it was more exaggerated than here, and it was noticeably different than the speaker’s more casual production the first time through the word list. Having decided that this recording was “not representative”, I chose not to analyze it – but I now realize that it may reflect the careful enunciation of someone who has been educated, which in South Asian schools often entails formal drilling, repetition, and recitation.

Nonetheless, as noted above, about one third of the isolation forms of nouns recorded from BSh_03 *do* show a downward slope on σ_2 . These are instances in which the speaker was less meticulous about avoiding a falling final contour. In the monomorphemic noun [smn.'ma] *smn.ma* ‘eyebrows’ shown in Figure 4.28 below, the speaker maintained a rising pitch slope – measured as 13 Hz/100 msec – nearly to the end of the vowel, and then apparently let it go.

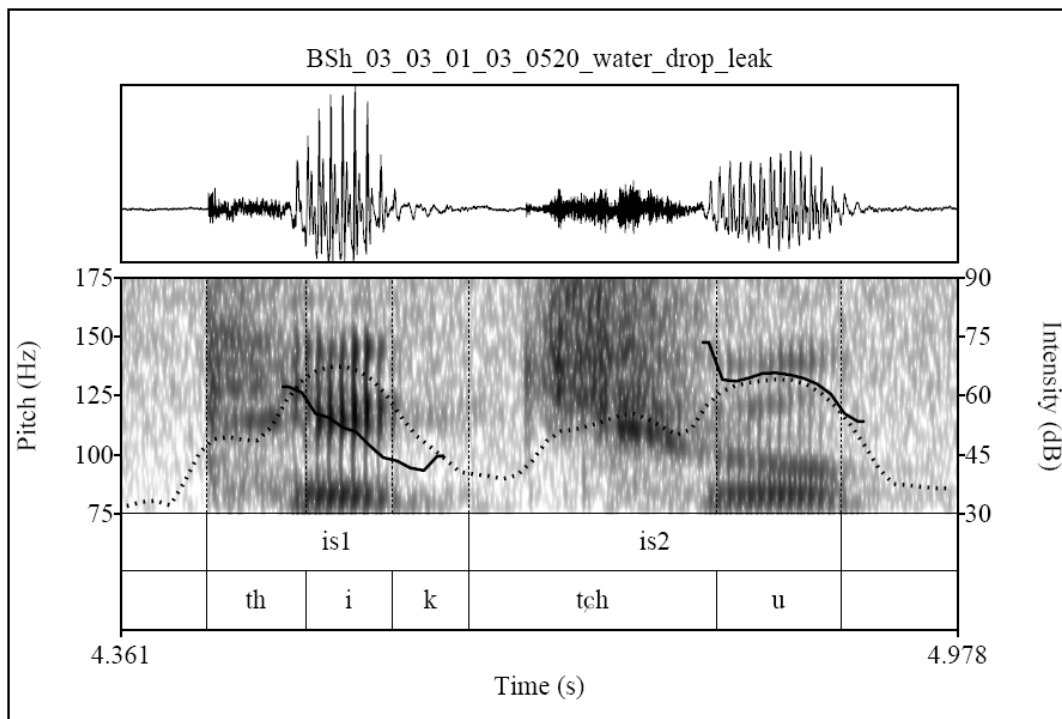
Figure 4.28 BSh_03 / Noun / Pitch trace: [smn.'ma] *smn.ma* ‘eyebrows’



The intonation patterns of these two speakers stand in contrast to the very informal, less careful, and sometimes mumbled responses from speaker BM_01, which, as shown in 4.3.2, are characterized by a final falling intonation contour. This speaker was older and uneducated, and I think this may account for the difference in intonation patterns.

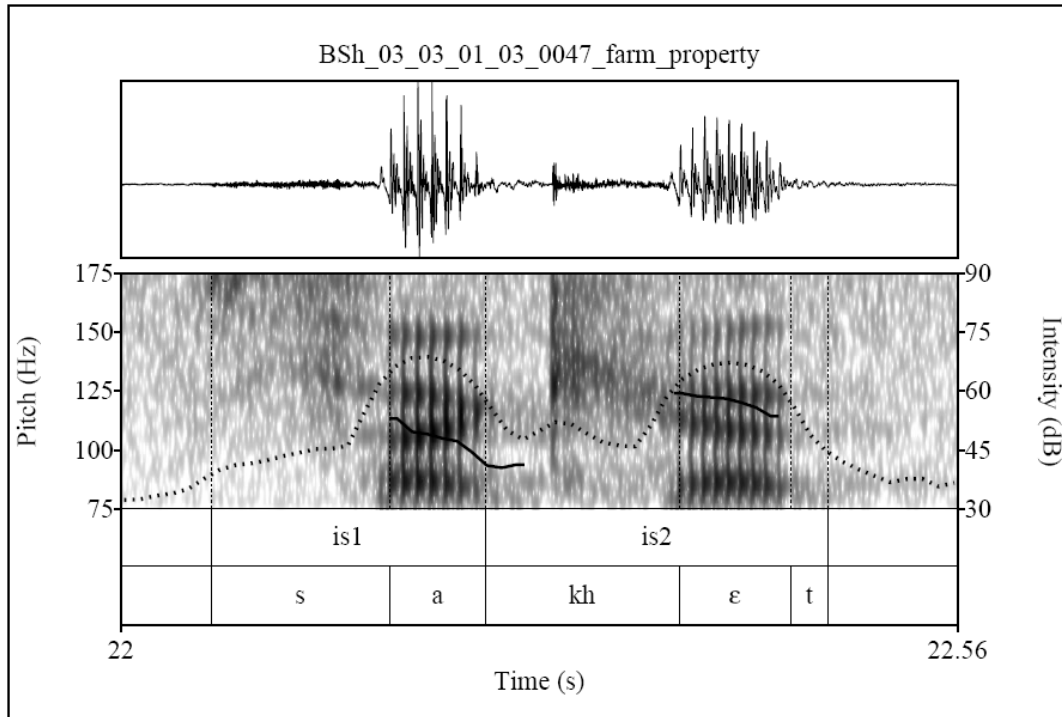
The pitch trace for [tʰig.ʰtɕʰu] *thig.chu* ‘water drop, leak’ in Figure 4.29 below shows the decline beginning earlier. Here, pitch on σ_2 is fairly constant over the central span, at 134 Hz. Near the end of the vowel – which is near the end of the utterance – pitch gently declines. The overall slope for this vowel was measured as -14 Hz/100 msec.

Figure 4.29 BSh_03 / Noun / Pitch trace: [tʰik.ʰtɕʰu] *thig.chu* ‘water drop, leak’



A third illustration is provided by the pitch trace for [sa.ʰkʰɛt] *sa.khyad* ‘farm property, agricultural field’ in Figure 4.30 below. Here, the pitch slope on σ_1 is -30 Hz/100 msec, and the pitch slope on σ_2 is -14 Hz/100 msec. The difference in slopes is calculated as $-14 - (-30) = 16$ Hz/100 msec.

Figure 4.30 BSh_03 / Noun / Pitch trace: [sa.'k^hεt] *sa.khyad* ‘farm property, agricultural field’



This slope difference of 16 Hz/100 msec happens to be the same as the mean slope difference calculated across all paired syllables for the full set of 77 nouns, as determined by a paired-sample t-test summarized in Table 4.10 below. As shown, this difference is highly significant ($p \ll 0.05$), and the 95% confidence interval defines a fairly narrow range of probable mean differences. That is, if we repeatedly recorded and analyzed samples of nouns produced in isolation by speaker BSh_03, there is a 95 % probability that we would observe approximately the same relationship between pitch slopes across syllables.

Table 4.10 BSh_03 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

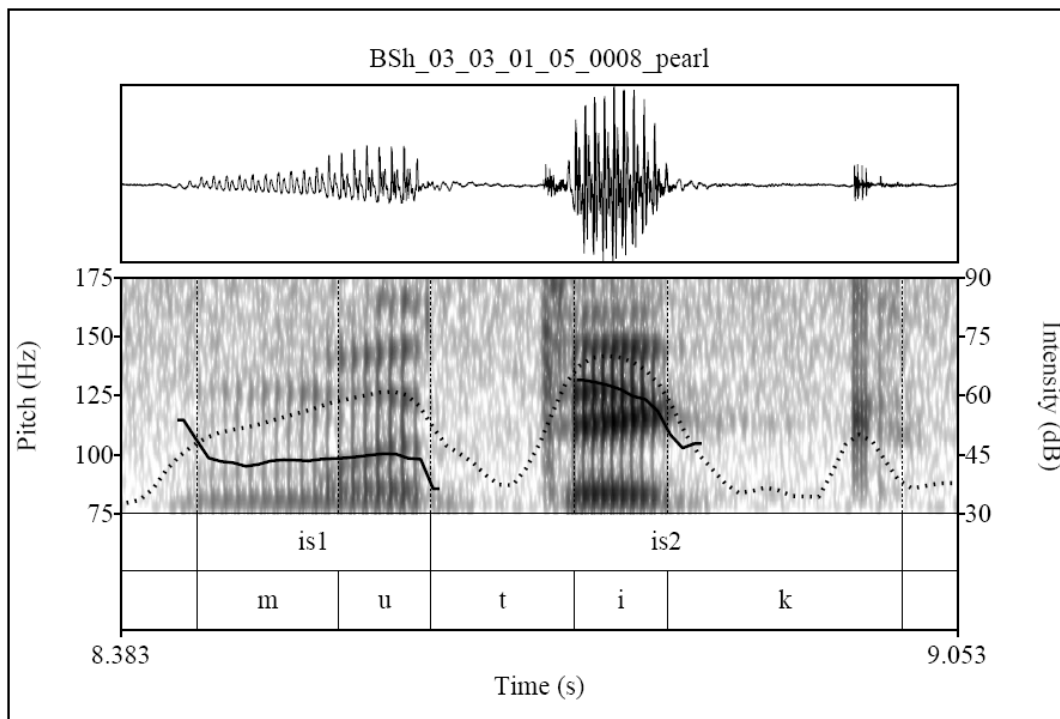
Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BSh_03	Isolation	9.9917	76	1.722e-15	16	13	20

As discussed in section 2.4, a positive pitch slope difference can – in theory – mean one of three things; either: (a) σ_1 slopes downward, and σ_2 slopes upward; (b) σ_1 slopes downward, and σ_2 also slopes downward, but not as steeply; or (c) σ_1 slopes upward, and σ_2 also slopes upward, but more steeply. (These configurations all fall to the left of the dashed line in Figure 4.23; graphically, the vector representing σ_2 is always “above” the vector representing σ_1 – the opposite of what is observed in Amdo nouns.) In practice, only (a) and (b) occur here, because there are no cases in which the pitch on σ_1 slopes upward, as illustrated in Figure 4.23. These are the two situations illustrated in the examples above, (a) being more frequent in this sample than (b). (Configuration (a) is the one represented by the median values in the box-and-whisker plot in Figure 4.23.)

Among the cases where the speaker did not attempt to avoid a falling final intonation contour were a few in which comparison of pitch slope across syllables yielded a negative difference. In [mu.'tik] *mu.tig* ‘pearl’ shown in Figure 4.31 below, pitch is virtually level on σ_1 , sloping at 0.42 Hz/100msec. In σ_2 , pitch slopes downward at -20 Hz/100msec, so the slope difference is $(-20) - (0.42) = -20.42$ Hz/100msec. The difference in pitch across syllables is still quite close to the target:

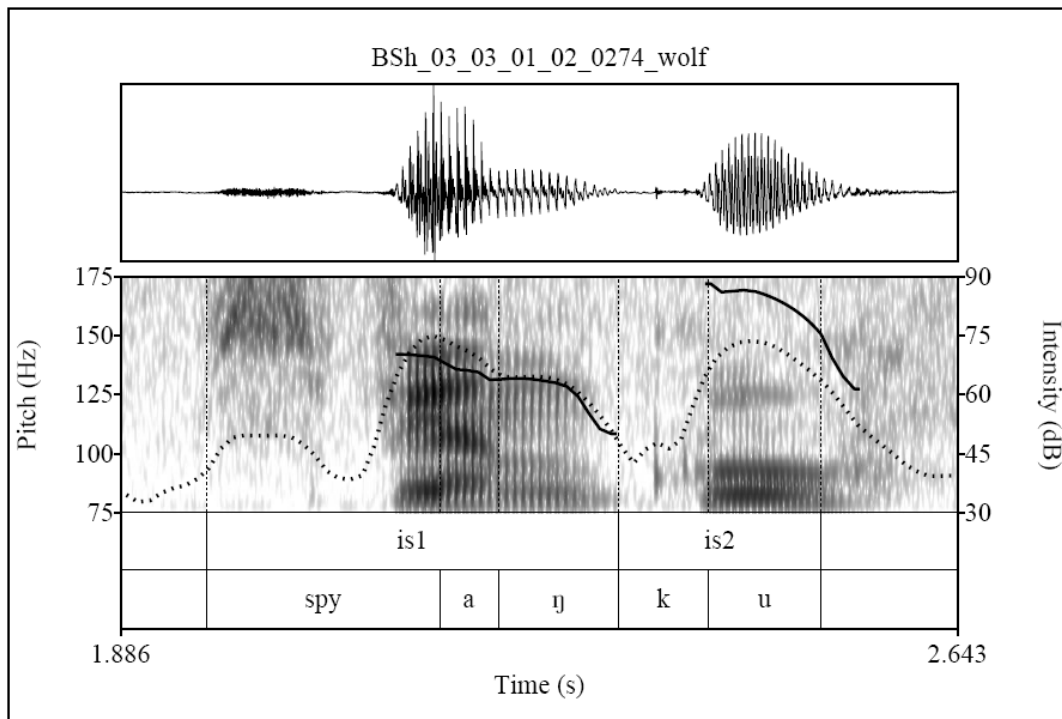
the average pitch measured across the medial 50% of the vowel was 100 Hz for σ_1 and 127 Hz for σ_2 mean, for an increase across syllables of 27 Hz.

Figure 4.31 BSh_03 / Noun / Pitch trace: [mu.'tik] *mu.tig* ‘pearl’



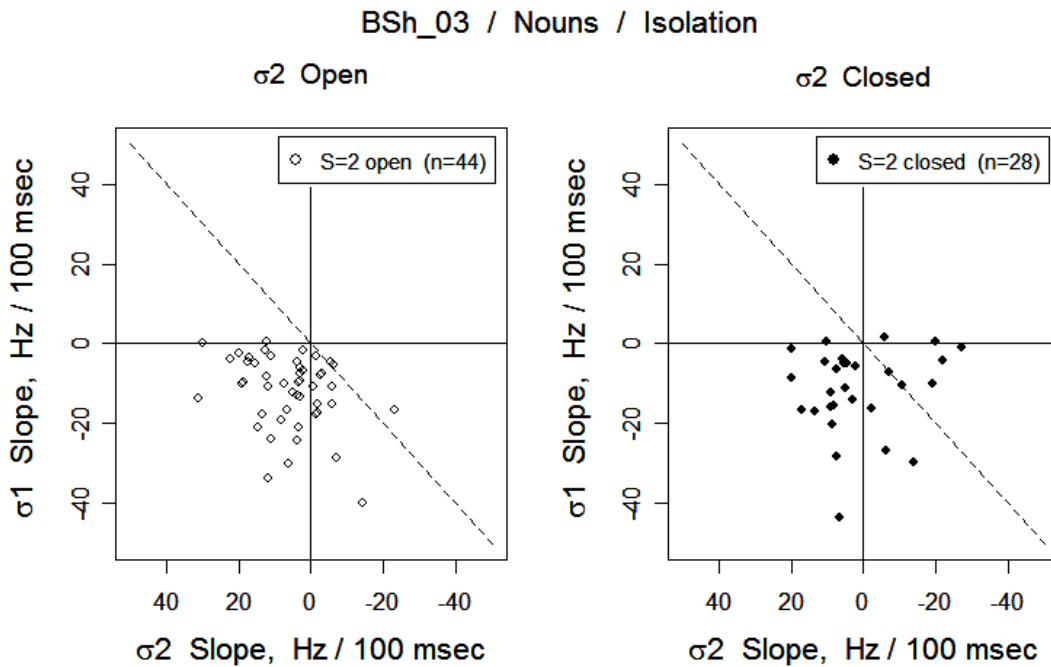
[spyɑŋ.'ku] *spyang.ku* ‘wolf’ is the only noun with σ_2 open which has a negative pitch slope difference. As illustrated in Figure 4.32 below, pitch slopes downward fairly steadily in σ_1 (at -16 Hz/100msec), while σ_2 exhibits a late and steep fall, averaging -23 Hz/100msec. The slope difference is calculated as $(-23) - (-16) = -7$ Hz/100msec. The pitch difference across syllables here is 167 Hz – 135 Hz = 32 Hz.

Figure 4.32 BSh_03 / Noun / Pitch trace: [spyaŋ.'ku] *spyang.ku* 'wolf'



In Figure 4.33 below, pitch slope measurements for isolation forms are plotted in terms of whether σ_2 is open or closed. By and large, this factor makes no difference here – it becomes relevant in other cases discussed below. The points with σ_2 closed which fall to the right of the dashed line are those with a final falling intonation contour, as discussed above.

Figure 4.33 BSh_03 / Nouns / Isolation: Pitch slope vs. $\sigma 2$ closure

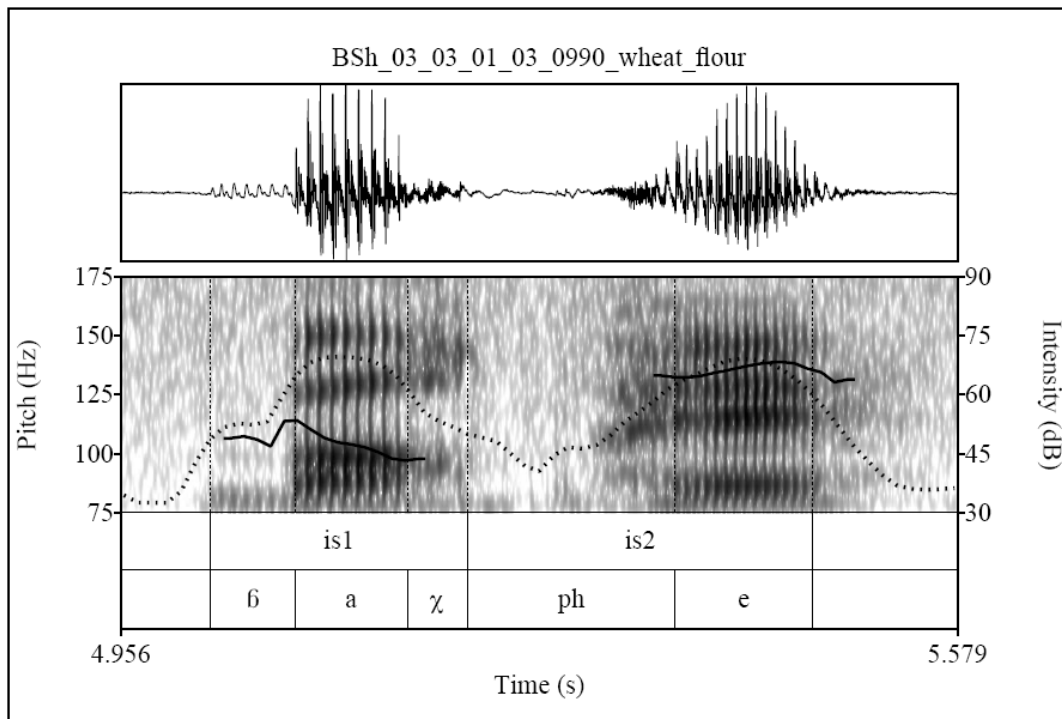


The pitch slope patterns exhibited by BSh_03 nouns produced in the sentence frame reflect the interaction of the lexical pitch pattern with the phrasal intonation pattern. The upward pitch slope observed in $\sigma 2$ of the isolation forms is altered, in the frame forms, because of a declining intonation over the latter part of the utterance.

Isolation and frame forms of the word [baχ.'p^he] / [mbax.'p^he] *bag.phye*

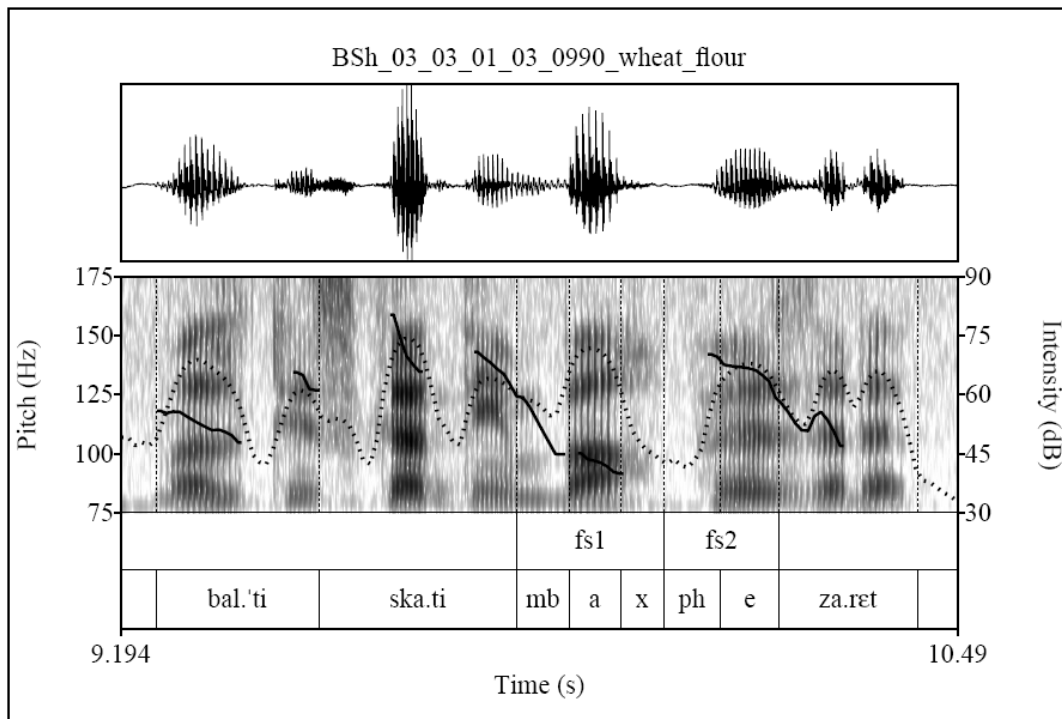
‘wheat flour’ provide an illustration. The isolation form, as shown in Figure 4.34 below, has a downward-sloping pitch on $\sigma 1$ (averaging 104 Hz, sloping at -21 Hz/100msec) and an upward-sloping pitch on $\sigma 2$ (averaging 137 Hz, sloping at 3.4 Hz/100msec). The difference in slope across syllables is 24 Hz/100msec. This word is typical of the patterns described above and illustrated in Figure 4.23.

Figure 4.34 BSh_03 / Noun / Isolation / Pitch trace: [bax.'p^he] *bag.phye* 'wheat flour'



When produced within the sentence frame, the target word becomes the focus. In Figure 4.35 below, pitch increases in steps from the beginning of the sentence until it approaches the target. At this point, pitch must drop in order to achieve the low-high pitch pattern on the target word needed to convey stress. This accounts for the downward slope on the vowel in σ_1 (averaging 97 Hz, sloping at -14 Hz/100msec). Pitch then leaps upward from σ_1 to σ_2 , but immediately begins to decline to the end of the utterance (averaging 136 Hz, sloping at -17 Hz/100msec). This intonation-driven contour accounts for the downward pitch slope in σ_2 . The difference in pitch slopes is calculated as -3 Hz/100msec – i.e., downward a bit more steeply in σ_2 than in σ_1 .

Figure 4.35 BSh_03 / Noun / Frame / Pitch trace: [mbax.'p^he] *bag.phye* ‘wheat flour’

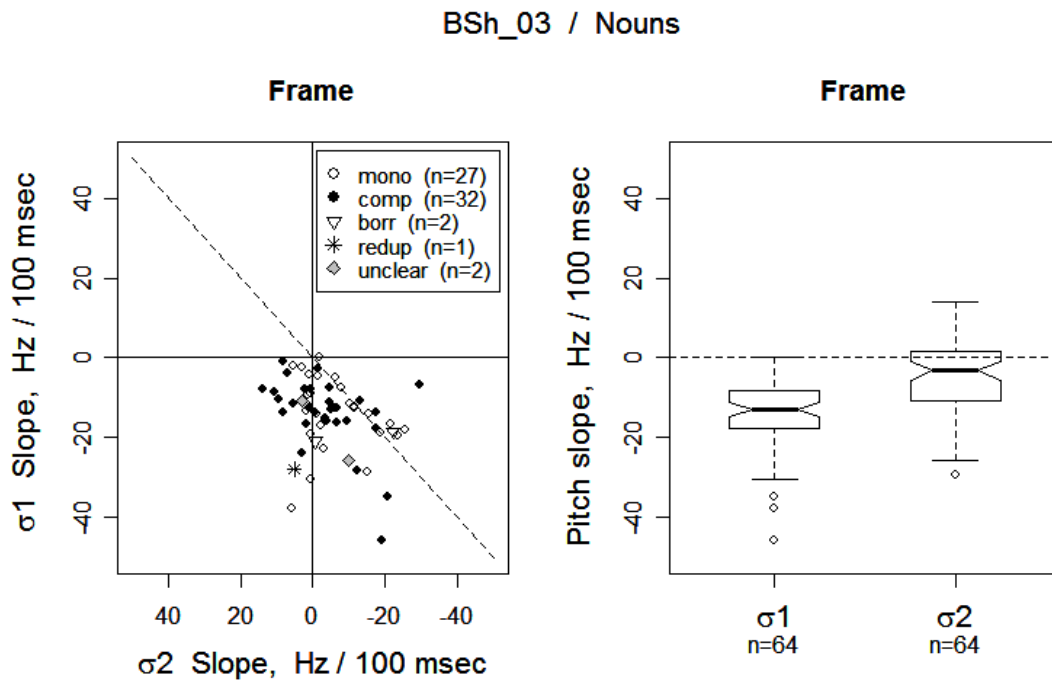


Pitch slope measurements for all BSh_03 nouns produced in the sentence frame are plotted in Figure 4.36 below. (These were also shown in Figure 4.23, for comparison with the isolation forms.) In general, as illustrated in the box-and-whisker plot, pitch slopes downward in σ_1 , and in most cases also slopes downward in σ_2 , though not as steeply – exactly the case illustrated just above. In about one third of the nouns, pitch slopes upward in σ_2 , as indicated by the whisker extending above the “0” line.

In the distribution plot, all nouns fall close to or to the left of the dashed line with the exception of the compound [χ mul.'t ϵ^h u] *rngul.chu* ‘sweat, perspiration’ at (σ_2, σ_1) coordinates (-30, -7). For the majority of nouns, $(\text{slope}_{\sigma_2} - \text{slope}_{\sigma_1}) > 0$ – i.e.,

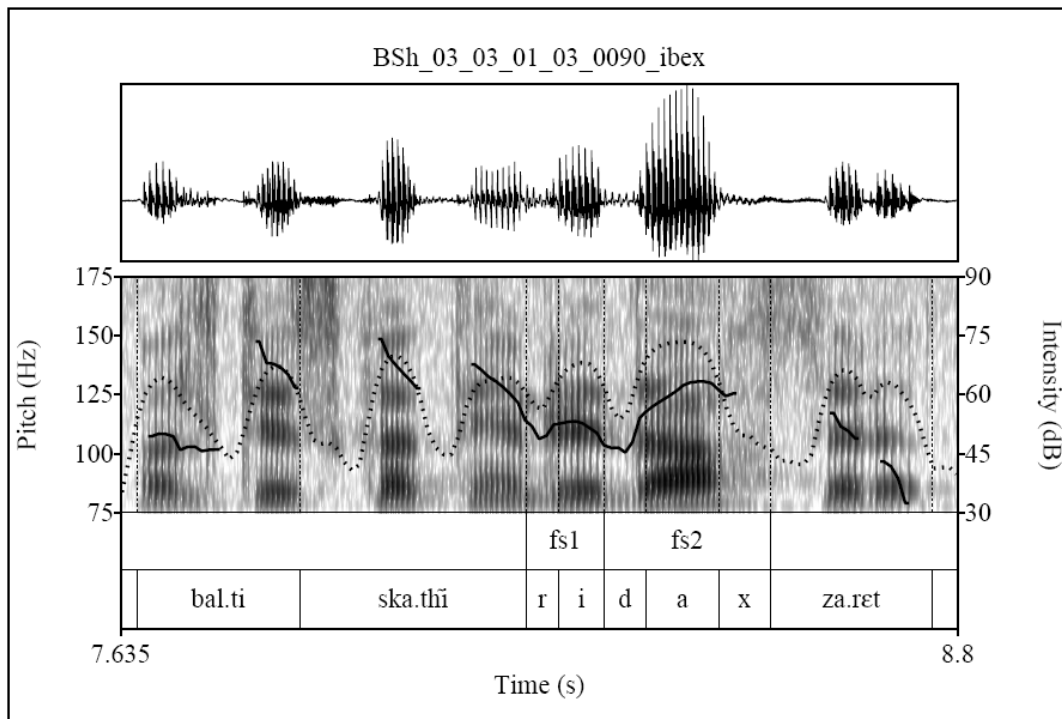
they fall to the left of the dashed line – so again, either (a) $\sigma 1$ slopes downward, and $\sigma 2$ slopes upward; or (b) $\sigma 1$ slopes downward, and $\sigma 2$ also slopes downward, but not as steeply.

Figure 4.36 BSh_03 / Nouns / Frame: Pitch slope



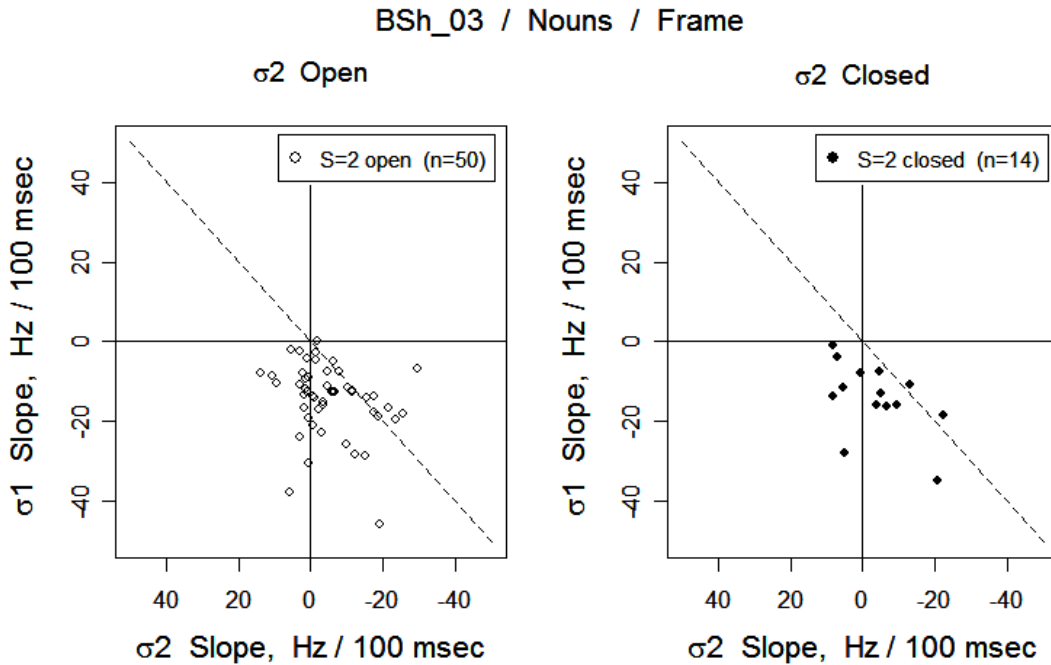
The compound noun [ri.'dax] *ri.dwags* ‘ibex’ is an example of (a). In Figure 4.37 below, the pitch on $\sigma 1$ is low and slopes downward (averaging 113 Hz, sloping at -14 Hz/100msec), while the pitch on $\sigma 2$ is high and slopes upward (averaging 128 Hz, sloping at, 8 Hz/100msec).

Figure 4.37 BSh_03 / Noun / Frame / Pitch trace: [ri.'dax] *ri.dwags* 'ibex'



Pitch slope measurements for the BSh_03 frame forms are plotted in Figure 4.38 below in terms of whether σ_2 is open or closed. As was the case for the isolation forms (Figure 4.33), there is no distinction here.

Figure 4.38 BSh_03 / Nouns / Frame: Pitch slope vs. $\sigma 2$ closure



A paired-sample t-test, summarized in Table 4.11 below, shows that the difference in slope across syllables for frame forms is statistically highly significant ($p \ll 0.05$), with a mean difference of 9.6 Hz/100 msec.

Table 4.11 BSh_03 / Nouns / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed)

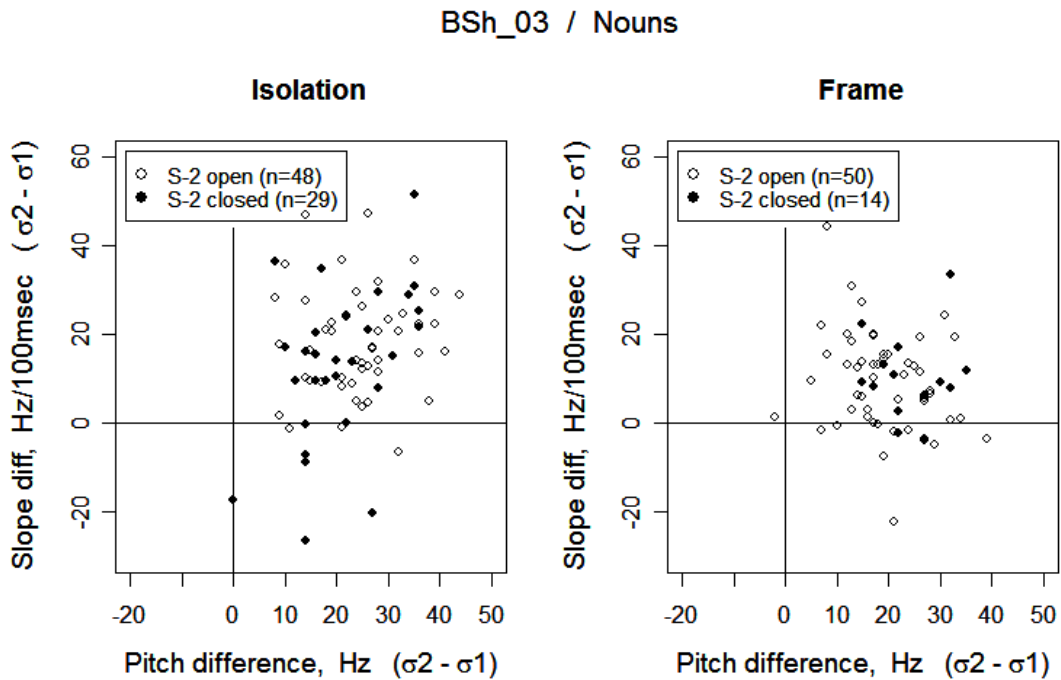
Speaker	Setting	t	DF	p-value	Mean diff (Hz100msec)	95 % conf. limits	
						lower	upper
BSh_03	Frame	7.1029	63	1.322e-09	9.6	6.9	12

While the slope contrast across syllables may be statistically significant, as with the isolation forms, it is not contextually significant. What this analysis has demonstrated is that the pitch slope patterns observed can be explained in terms of

pitch, for the isolation forms, and in terms of the interaction of pitch and intonation, for the frame forms. The observed patterns are simply a byproduct of other factors. Thus pitch slope cannot be considered an acoustic correlate of stress for BSh_03 nouns.

The relationship between pitch and pitch slope for isolation and frame forms is captured in Figure 4.39 below. All points fall to the right of the y-axis: the pitch difference is positive, meaning that pitch rises from σ_1 to σ_2 . Most of the points also fall above the x-axis; the pitch slope difference is positive, too, meaning that the slope on σ_2 is “more upward”, as described above. Thus the correspondence between pitch and pitch slope observed in both isolation and frame forms is conveyed by the fact that most of the nouns lie in the upper right quadrant. The upward pitch slope in σ_2 is not contextually significant because it is a mere continuation of the increase in pitch across syllables.

Figure 4.39 BSh_03 / Nouns: Pitch difference vs. pitch slope difference



This pattern is very different from what is observed in nouns produced by Balti speaker BM_01, where the utterance-final falling intonation contour dominates pitch slope even in the isolation forms, and from what is observed in nouns produced by the Rebkong Amdo speakers, where pitch slope is a strong correlate of σ_2 stress in non-verbs.

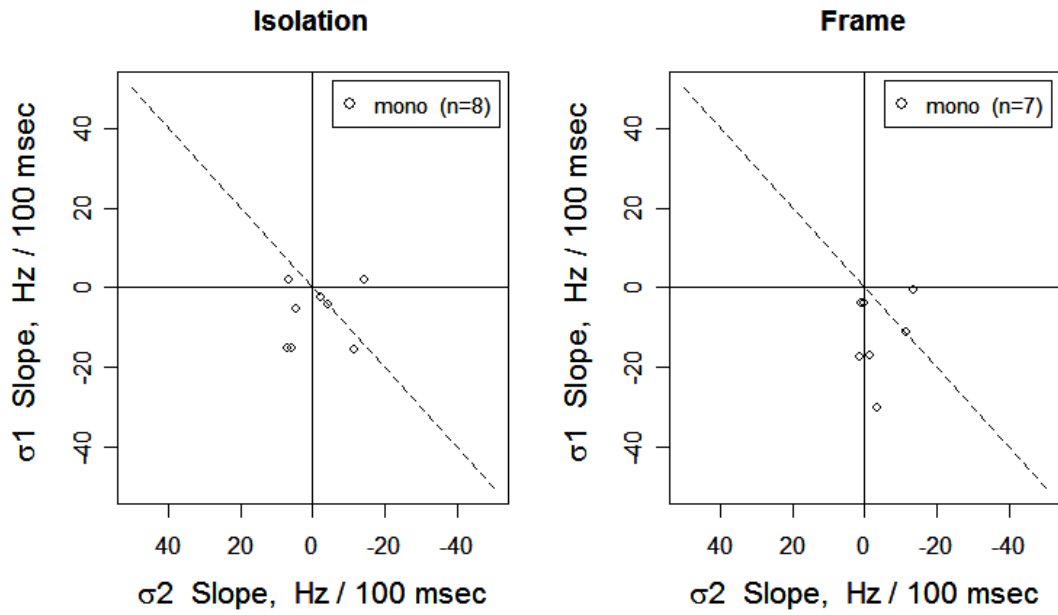
4.3.1.2 BSh_03 Adjectives

Pitch slope patterns for BSh_03 adjectives are similar to those described for nouns. Measurements for isolation and frame forms are plotted in Figure 4.40 below. Though there are very few tokens, it is clear that all of them fall close to or below the x-axis, meaning that pitch slopes downward on σ_1 . Half of the isolation forms have an upward slope on σ_2 , and half have a downward slope on σ_2 . In the frame forms,

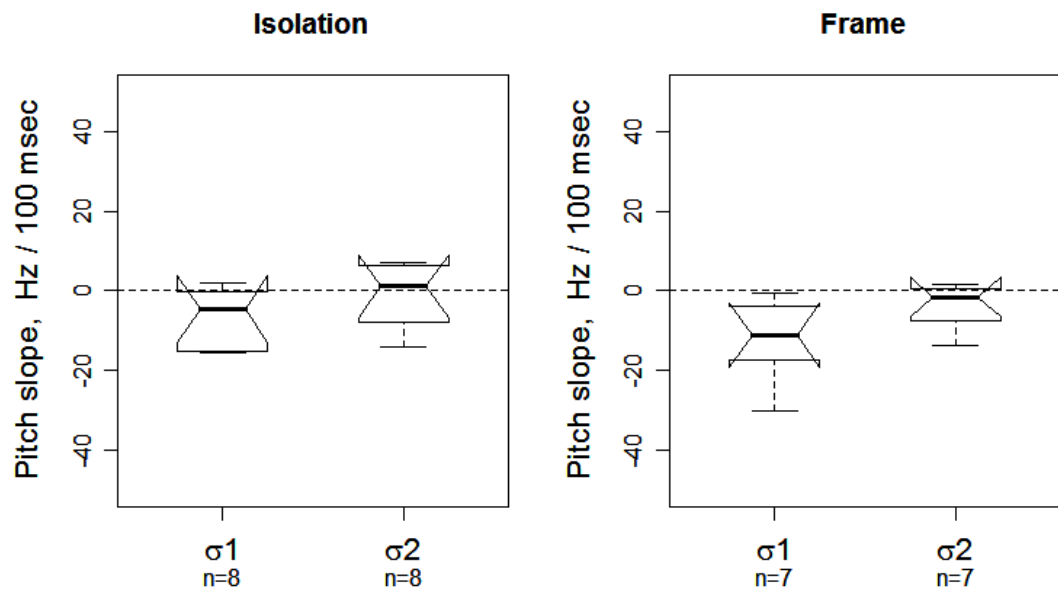
pitch is either flat on σ_2 , or slopes downward. The one word which falls to the right of the dashed line in both settings is [kar.'p^ho] *dkar.po* 'white'.

Figure 4.40 BSh_03 / Adjectives: Pitch slope

BSh_03 / Adjectives

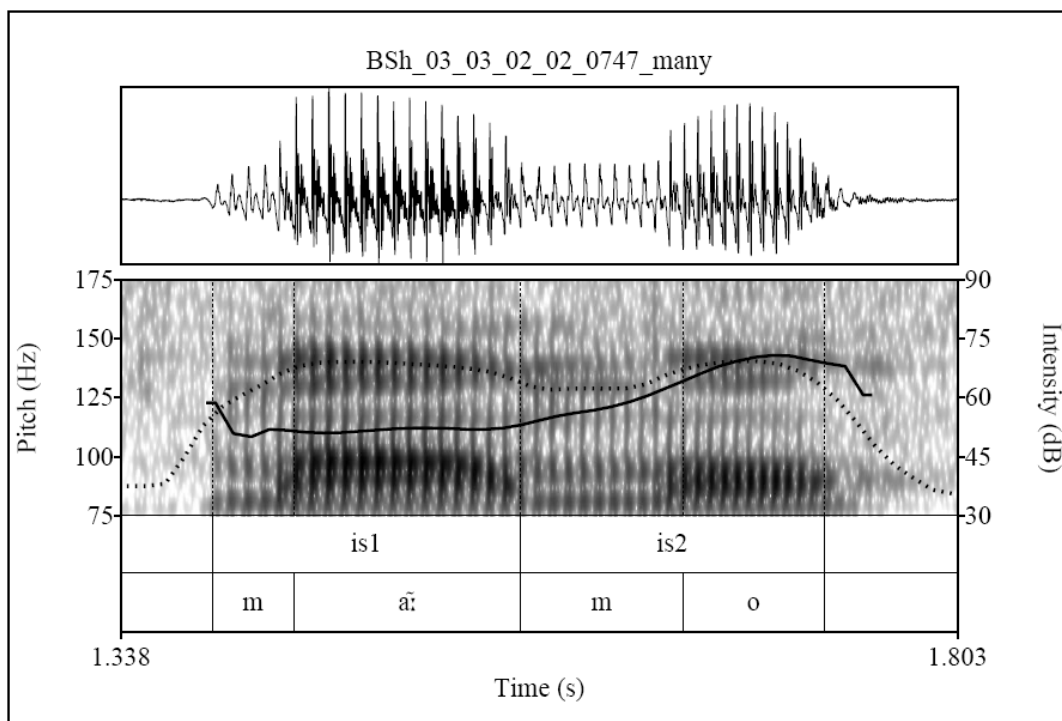


BSh_03 / Adjectives



A representative example is [mã:.'mo] *mang.mo* ‘many’. In isolation, as shown in Figure 4.41 below, the pitch slope on σ_1 is a barely discernible 2 Hz/100msec; the slope on σ_2 is 7 Hz/100msec. The difference in pitch slope is calculated as 5 Hz/100msec. As was the case with the majority of nouns, the speaker is careful to maintain the high pitch on σ_2 ; there is only a minor utterance-final decline.

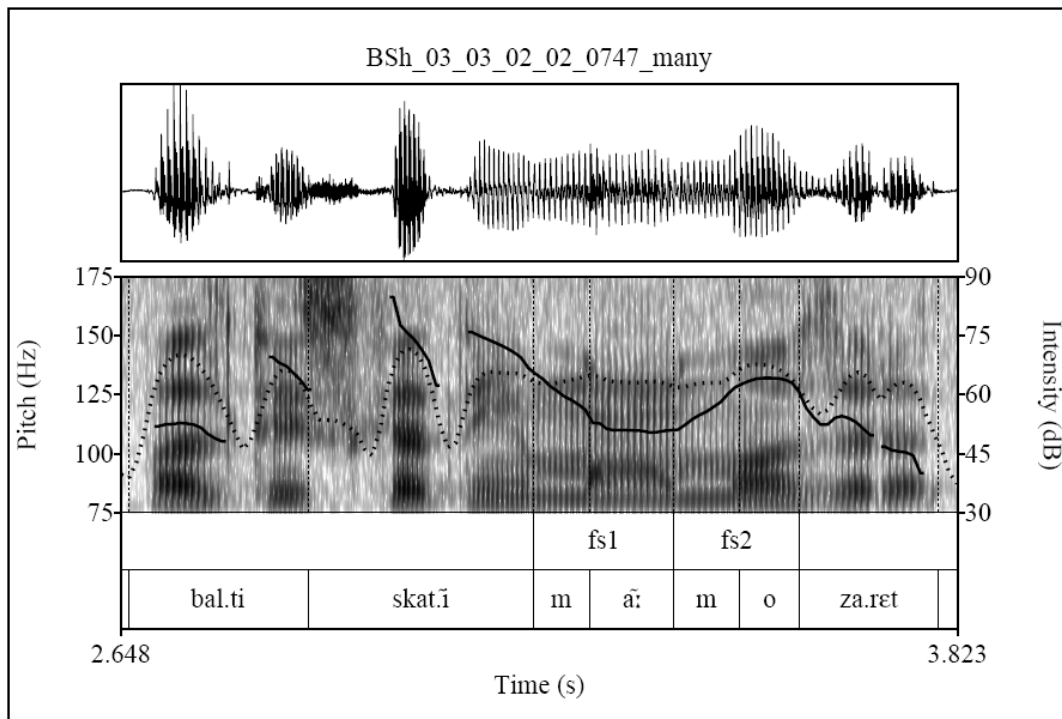
Figure 4.41 BSh_03 / Adjective / Isolation / Pitch trace: [mã:.'mo] *mang.mo* ‘many’



When the same word is produced in the sentence frame, lexical pitch interacts with phrasal intonation. From the beginning of the utterance, pitch rises in steps to a peak before the target word, and then declines to the low pitch required on σ_1 as part of the stress pattern. The slight downward pitch slope on σ_1 (-3.9 Hz/100msec)

reflects this descent. Pitch rises across the consonant at the syllable boundary. For the vowel in σ_2 , the pitch slope was measured as 0.19 Hz/100msec. The difference in pitch slope is 4.1 Hz/100msec. This is essentially the same as the difference determined for the isolation form (4.5 Hz/100msec).

Figure 4.42 BSh_03 / Adjective / Frame / Pitch trace: [mã:.'mo] *mang.mo* ‘many’



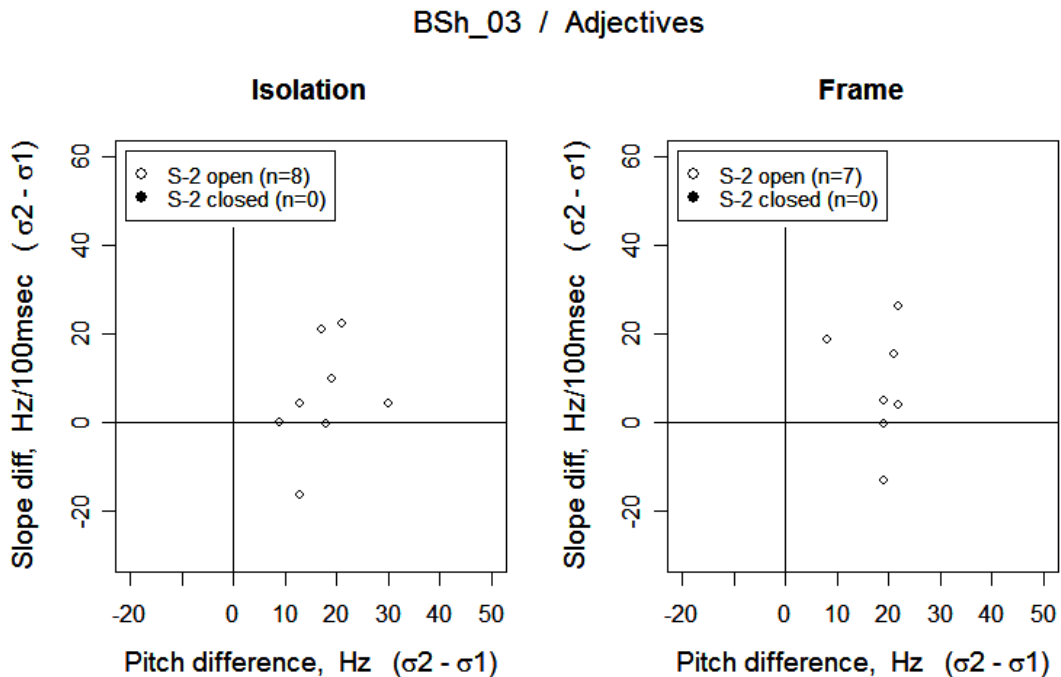
Paired-sample t-tests comparing pitch slope across syllables for isolation and frame forms are summarized in Table 4.12 below. For adjectives in both settings, $p > 0.05$ and the 95% confidence intervals about the mean differences include zero. These tests thus indicate that the difference in slope is not statistically significant, which is not surprising given the small sample size.

Table 4.12 BSh_03 / Adjectives / Pitch slope: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BSh_03	Isolation	1.2918	7	0.2374	5.6	-4.7	16
	Frame	1.6059	6	0.1594	8.1	-4.3	20

Figure 4.43 below shows the relationship between pitch and pitch slope for BSh_03 adjectives – all of which are monomorphemic and have an open σ_2 . As was the case with nouns (Figure 4.39), most of the points fall in the upper right quadrant. Pitch on σ_2 is higher than pitch on σ_1 , and the vector representing slope on σ_2 is likewise “above” that representing slope on σ_1 , reflecting the general rise across syllables.

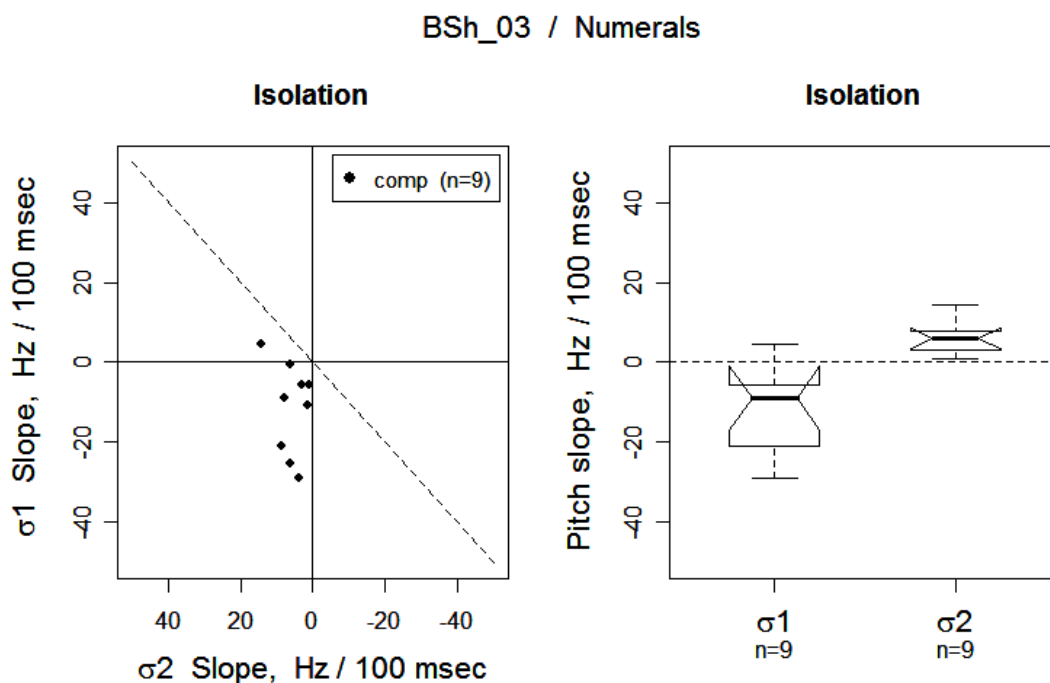
Figure 4.43 BSh_03 / Adjectives: Pitch difference vs. pitch slope difference



4.3.1.3 BSh_03 Numerals

As illustrated in Figure 4.44 below, BSh_03 numerals exhibit a high degree of acoustic consistency in terms of pitch slope – as was also the case for pitch (Figure 4.5). The speaker’s production was likely influenced by the facts that the numerals were recorded one after the other, and they constitute a closed semantic set. In σ_1 , pitch slopes downward for all numerals except [t̥u.'ruk] *bcu.drug* ‘sixteen’. In σ_2 , all of the numerals have an upward slope; that is, the speaker consistently avoided an utterance-final falling intonation contour.

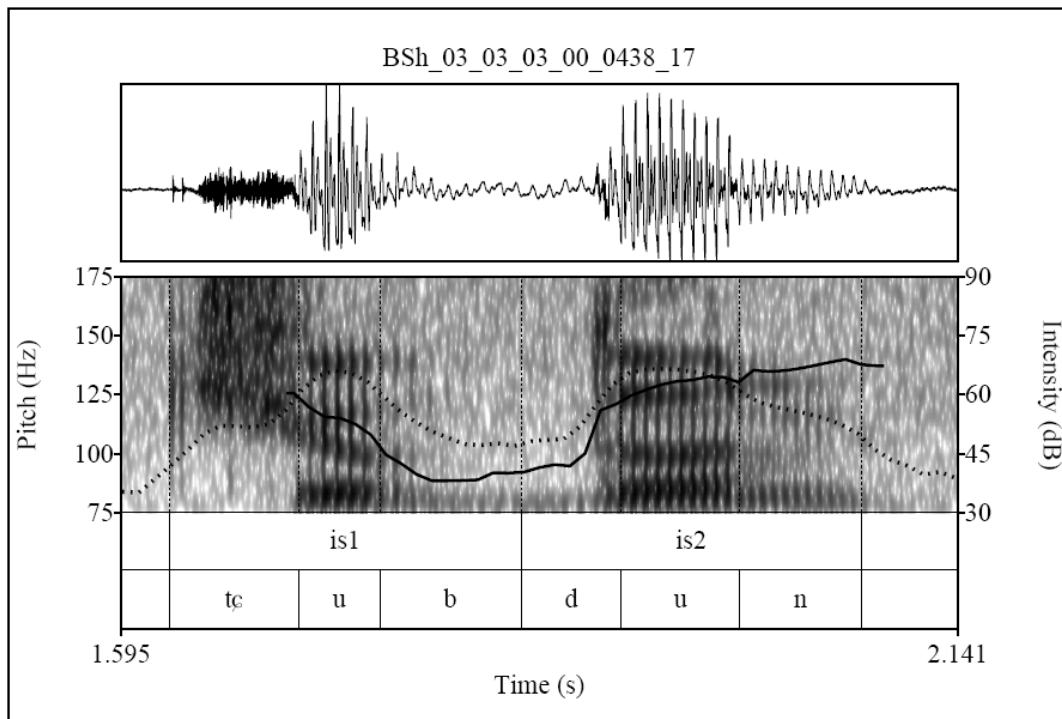
Figure 4.44 BSh_03 / Numerals: Pitch slope



The numeral [t̥ub.'dun] *bcu.bdun* ‘seventeen’ is representative. As shown in Figure 4.45 below, in the vowel in σ_1 pitch slopes downward at a rate of -21

Hz/100msec; the mean pitch over the medial 50% of the vowel is 115 Hz. In σ_2 , the speaker maintains a slightly upward-sloping pitch – 8.8 Hz/100msec – achieving a mean pitch of 131 Hz. The difference in pitch slope is calculated as $(8.8) - (-21) = 30$ Hz/100msec.

Figure 4.45 BSh_03 / Numerals / Pitch trace: [tɕub.'dun] *bcu.bdun* ‘seventeen’



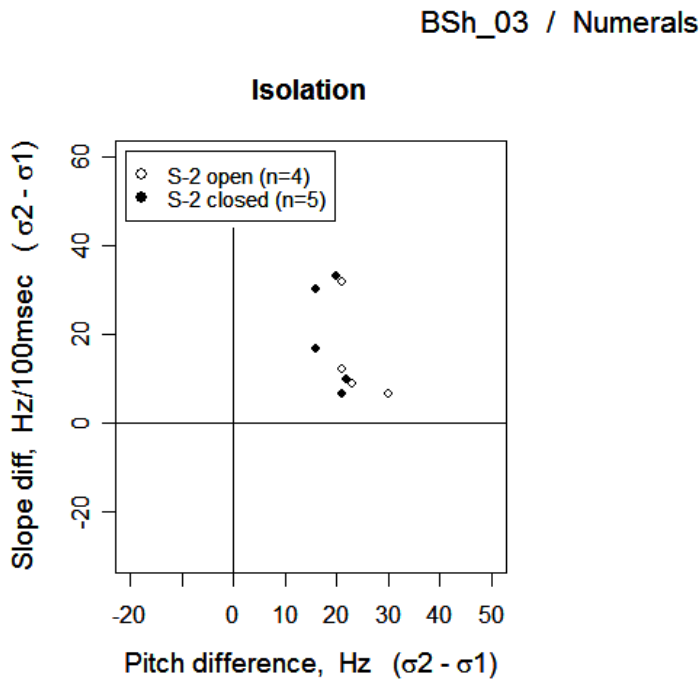
Given the consistency of this small sample, the difference in pitch slope across syllables was found to be statistically significant: $p < 0.05$. The mean difference in slope across syllables is 17 Hz/100msec.

Table 4.13 BSh_03 / Numerals / Pitch slope: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BSh_03	Isolation	4.6738	8	0.001595	17	8.7	26

In Figure 4.46 below, where pitch slope difference is plotted against pitch difference, all of the numerals fall in the upper right quadrant, as was the case for most of the other non-verbs discussed above. Again, this reflects the fact that the pitch slope pattern is governed by the pitch contrast across syllables; since pitch on σ_2 is higher than pitch on σ_1 , the vector representing slope on σ_2 is likewise “higher” than the vector representing slope on σ_1 .

Figure 4.46 BSh_03 / Numerals: Pitch difference vs. pitch slope difference

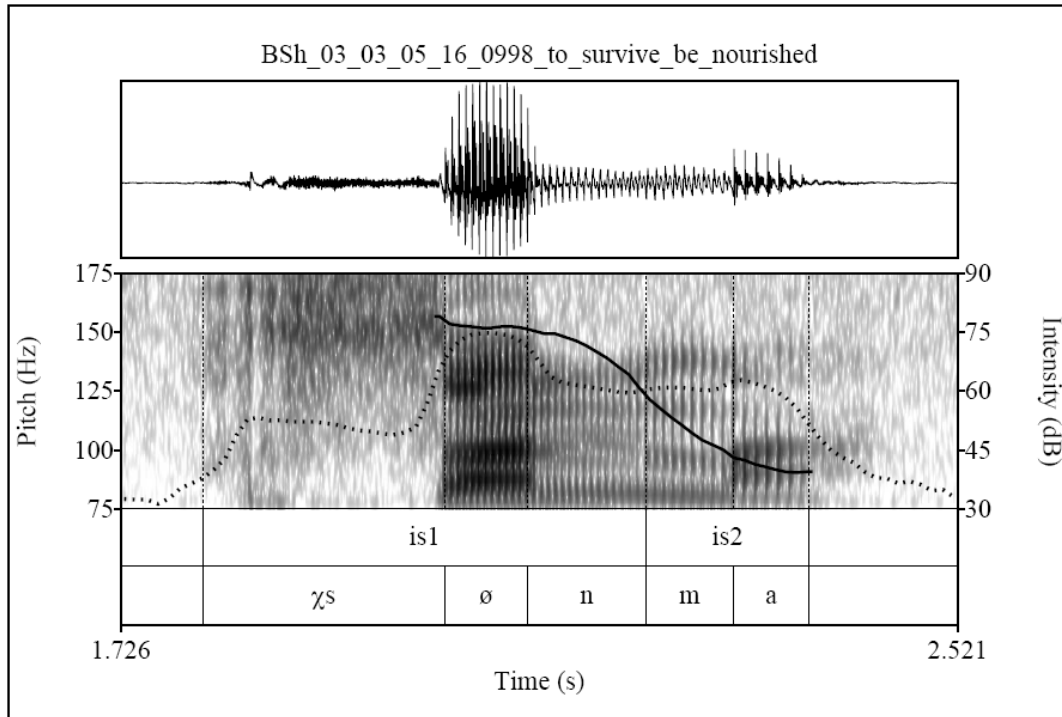


4.3.1.4 BSh_03 Verbs

In BSh_03 verbs, too, the observed pitch slope patterns are entirely attributable to other factors: to the contrast in pitch across syllables, for isolation forms, and to the interaction of pitch and intonation, for frame forms. Of course, since verbs are stressed on σ_1 rather than σ_2 , the patterns are considerably different than those described for nouns, adjectives, and numerals.

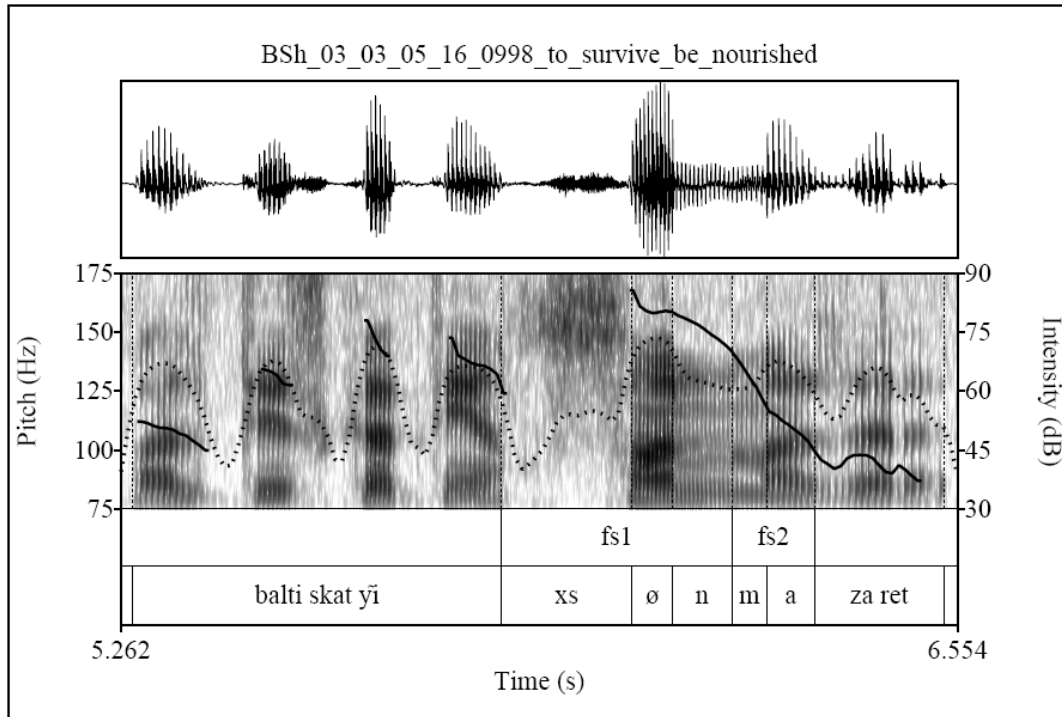
As an example, the pitch trace for the isolation form of the verb [$\chi s\text{ø}n.ma$] *gson.ma* ‘to survive, to be nourished’ is provided in Figure 4.47 below. The pitch on σ_1 is high and fairly level (averaging 152 Hz, sloping at -2.3 Hz/100msec). Since the consonants at the syllable boundary are both nasals, the declining pitch trajectory can be observed continuously to the end of the utterance. The pitch on σ_2 (averaging 92 Hz, sloping at -12 Hz/100msec) is simply the natural continuation of this trend. The difference in slopes is $(-12) - (-2.3) = -9.7$ Hz/100msec. The negative number reflects the fact that pitch slopes downward more steeply in σ_2 than in σ_1 ; in this sense, slope is more prominent in the *unstressed* syllable.

Figure 4.47 Pitch trace for BSh_03 ['χsøn.ma] *gson.ma* ‘to survive, to be nourished’ (isolation)



The pattern is even more dramatic when the word is embedded in the sentence frame, as it is reinforced by the larger-scale intonation pattern. The pitch trace for the sentence ['bal.ti 'skat.yĩ 'χsøn.ma za.ret] – ‘In [the] Balti language [we] ‘to survive’ say’ – is shown in Figure 4.48 below. The pitch increases successively on each syllable nucleus, reaches a peak on the stressed syllable (σ_1) of the target word, and then declines to a soft and creaky close. In the stressed syllable – the first syllable of the target word – the average pitch is 159 Hz, and the pitch slope is -3.1 Hz/100msec. Pitch then falls to an average of 110 Hz on σ_2 , with a slope of -25 Hz/100 msec. The difference in pitch slope is $(-25) - (-3.1) = -22$ Hz /100msec.

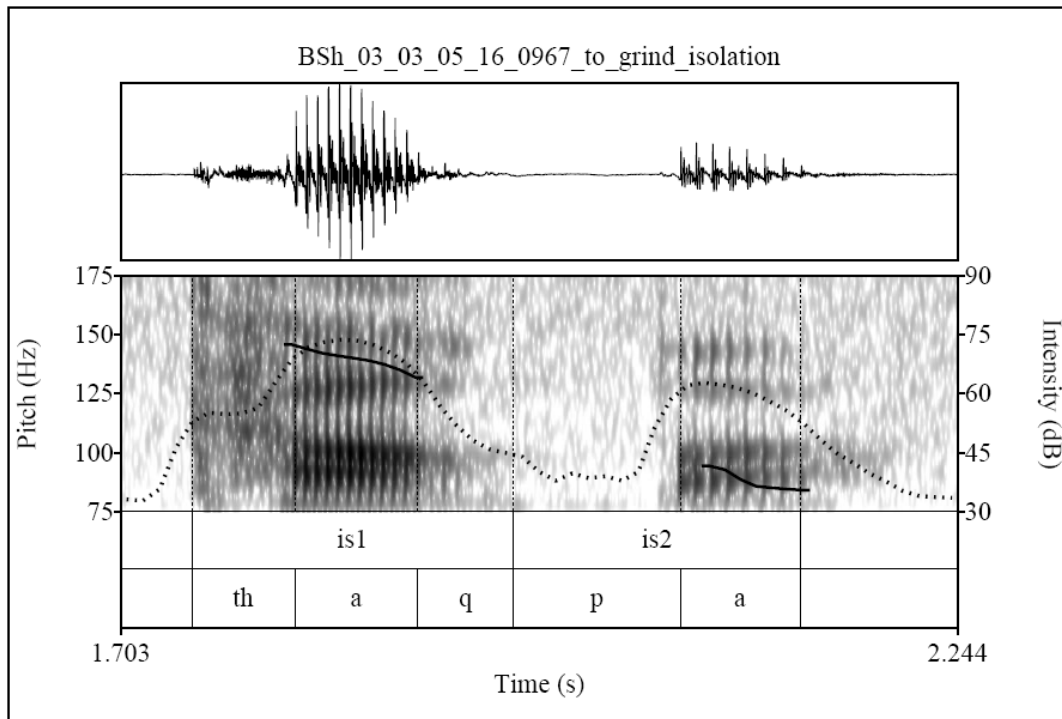
Figure 4.48 Pitch trace for BSh_03 ['xsøn.ma] *gson.ma* 'to survive, to be nourished' (frame)



A similar trend can be observed in Figure 4.49 below for the verb [^haq.pa] *'thag.pa* 'to grind', even though here the pitch trace is not continuous since the consonants at the syllable boundary are both stops. (The exact boundary between [q] and [p] here is somewhat arbitrary, since there are no reliable acoustic indicators of the change in place of articulation in the middle of the stop closure.) Again, in σ_1 the pitch is high and slopes downward (140 Hz, -14 Hz/100msec), and in σ_2 is lower and slopes downward a bit more steeply (90 Hz, -19 Hz/100msec). Again, this pitch slope pattern is completely unrelated to stress, and is simply a reflection of the general

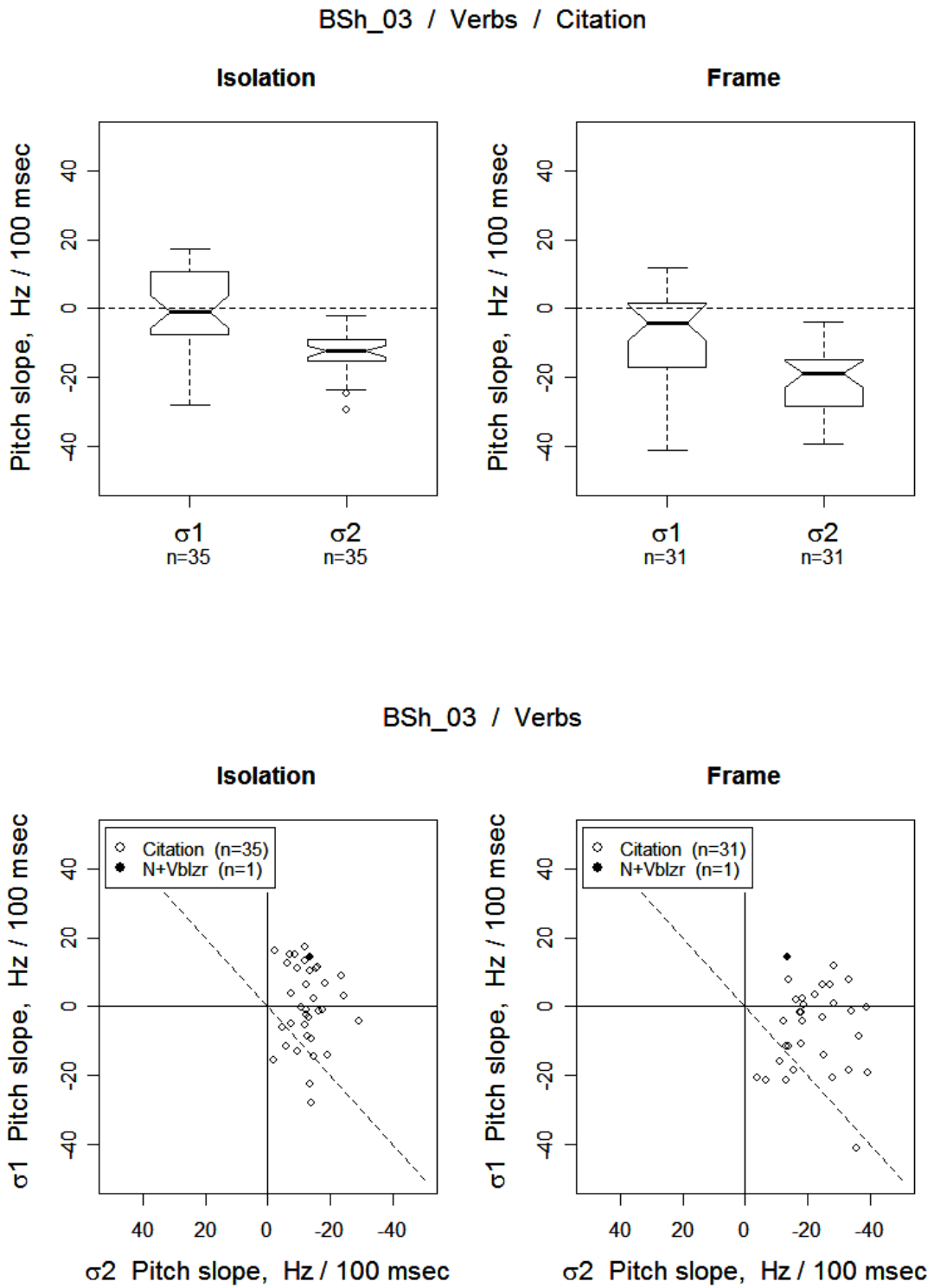
decline in fundamental frequency across the word. And this decline occurs because pitch is itself such a strong acoustic cue for stress.

Figure 4.49 Pitch trace for BSh_03 [t^haq.pa] 'thag.pa 'to grind' (isolation)



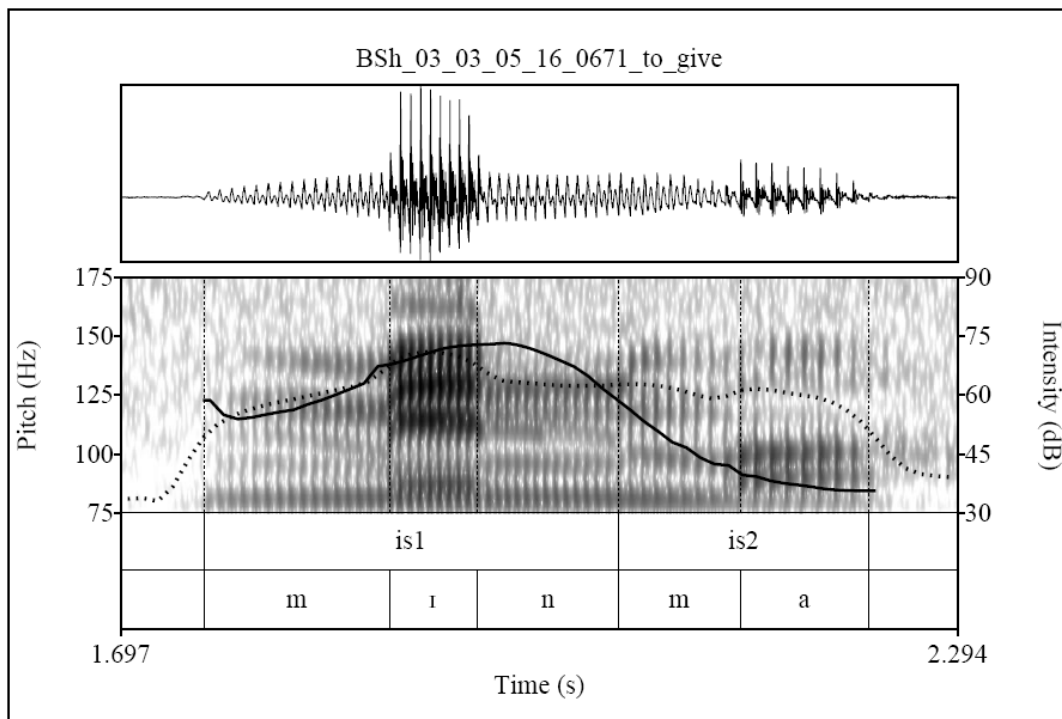
Pitch slope measurements for all BSh_03 verbs are plotted in Figure 4.50 below. The median values in the box-and-whisker plots reflect the prevailing trends: in σ_1 , pitch slope is sometimes upward and sometimes downward; in σ_2 , pitch slope is always downward. These patterns are in stark contrast to those observed in the non-verbs, illustrated in Figure 4.23, Figure 4.40, and Figure 4.44.

Figure 4.50 BSh_03 / Verbs: Pitch slope



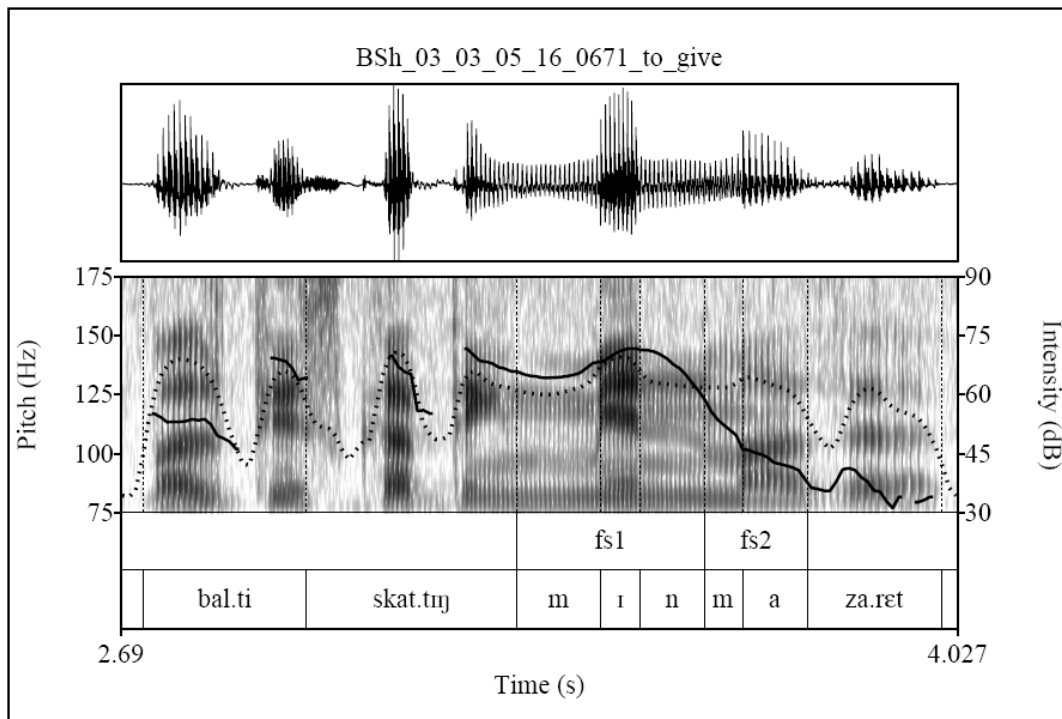
In the distribution plot above, all points fall to the right of the y-axis, since slope is always downward on σ_2 ; they fall both above and below the x-axis. The points above the x-axis represent verbs in which pitch slopes upward in σ_1 . The citation form [ʼmm.ma] *sbyin.ma?* ‘to give’ is one such case. As illustrated in Figure 4.51 below, pitch slopes upward in σ_1 (averaging 144 Hz, sloping at 13 Hz/100msec), and downward in σ_2 (averaging 87 Hz, sloping at -11 Hz/100msec). The slope difference is $(-11) - (13) = -24$ Hz/100msec.

Figure 4.51 BSh_03 / Verb / Isolation / Pitch trace: [ʼmm.ma] *sbyin.ma?* ‘to give’



The same pattern occurs in the frame form, as shown in Figure 4.52 below. Here, pitch on σ_1 averages 143 Hz and slopes at 8.0 Hz/100msec, and on σ_2 averages 97 Hz and slopes at -14 Hz/100msec. The difference in pitch across syllables is $(-14) - (8) = -22$ Hz/100msec.

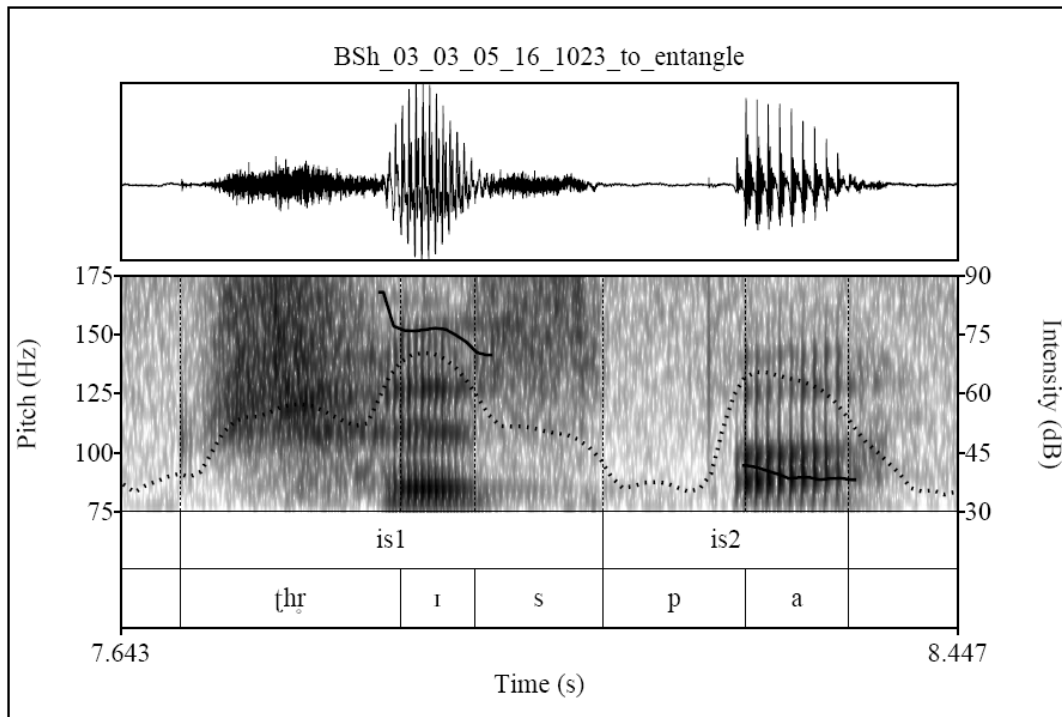
Figure 4.52 BSh_03 / Verb / Frame / Pitch trace: ['mɪn.ma] *sbyin.ma?* 'to give'



There are a few verbs in Figure 4.50 which plot to the left of the dashed line. In these cases, pitch slopes downward more steeply in σ_1 than in σ_2 . An example is [tʰɾɪs.pa] *'khris.pa* 'to entangle', shown in Figure 4.53 below. Here, pitch on σ_1 averages 152 Hz, and slopes downward at -11 Hz/100msec; pitch on σ_2 averages 90

Hz, and slopes downward at -5.8 Hz/100msec. The difference in slopes is $(-5.8) - (-11) = 5.2$ Hz/100msec.

Figure 4.53 BSh_03 / Verb / Pitch trace: [t^h_ris.pa] 'khris.pa' 'to entangle'



Pitch slope measurements for all of the BSh_03 citation forms were compared across syllables using paired-sample t-tests, summarized in Table 4.14 below. For both isolation and frame forms the contrast in slope is statistically highly significant ($p \ll 0.05$), with mean differences of -12 Hz/100msec and -15 Hz/100msec, respectively. The negative slope difference indicates that, graphically, the vector representing the slope in σ_2 is “below” the vector representing slope in σ_1 . (That is, the slope is more downward and thus more prominent on the syllable which is *not* stressed.) Theoretically, this relationship could have three possible configurations: (a)

σ_1 slopes downward, and σ_2 also slopes downward, but more steeply; (b) σ_1 slopes upward, and σ_2 slopes downward; or (c) σ_1 slopes upward, and σ_2 also slopes upward, but not as steeply. Again, though, only (a) and (b) occur, since there are no cases in which the pitch slope on σ_2 is downward, as shown in Figure 4.50 on page 225.

Table 4.14 BSh_03 / Verbs / Isolation / Pitch slope: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz100msec)	95 % conf. limits	
						lower	upper
BSh_03*	Isol.	-5.8758	34	1.248e-06	-12	-17	-8.2
	Frame	-5.3184	30	9.48e-06	-15	-21	-9.3

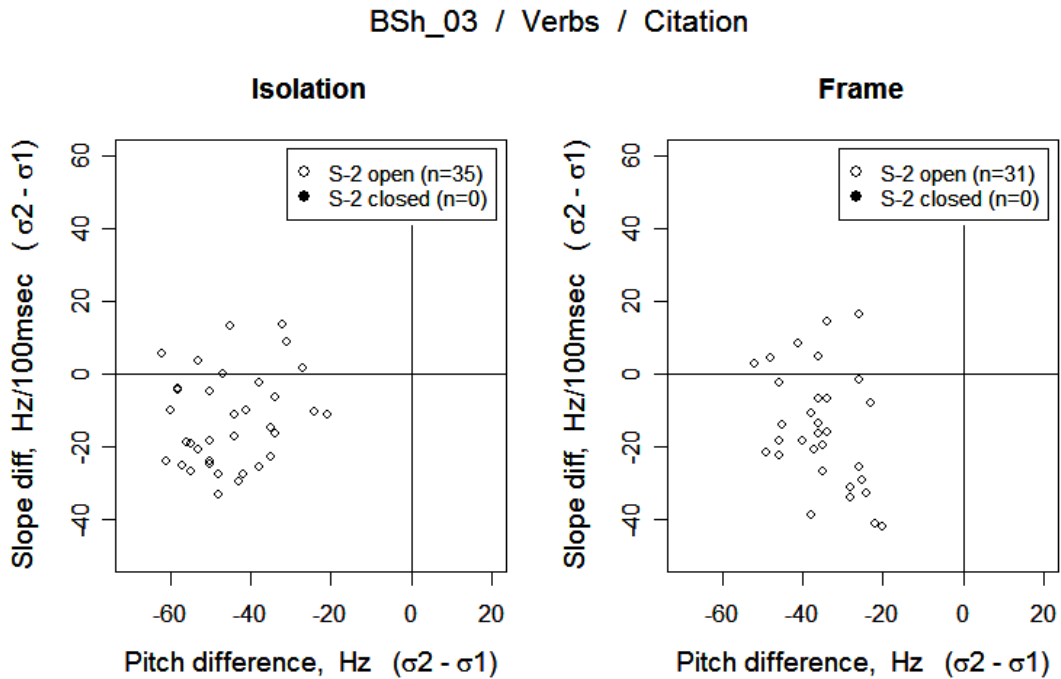
* The lone BSh_03 N+Verbalizer form was excluded.

The relationship between the pitch difference across syllables and the pitch slope difference across syllables is summarized in Figure 4.54 below. (All of the verbs have an open σ_2 since they are all citation forms, where σ_2 is [-pa].)

In contrast to the pattern observed in nouns (Figure 4.39) – as well as adjectives and numerals – here, all points fall to the left of the y-axis; the pitch difference is negative, meaning that pitch falls from σ_1 to σ_2 . Most of the points also fall below the x-axis; the pitch slope difference is negative, too, meaning that the slope on σ_2 is “more downward”, as described above. The points which plot above the x-axis are those which pattern like [tʰɾis.pa] 'khris.pa 'to entangle', shown in

Figure 4.53. The fact that the majority of points lie in the lower left quadrant reflects the correlation between pitch and pitch slope.

Figure 4.54 BSh_03 / Verbs: Pitch difference vs. pitch slope difference

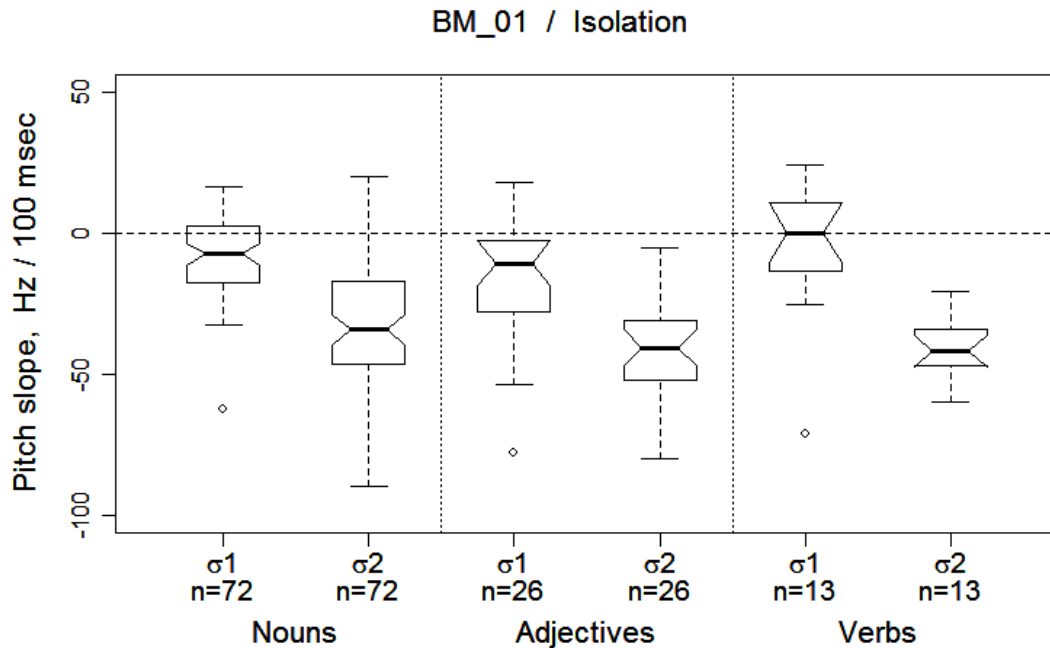


4.3.2 Pitch slope for speaker BM_01

Pitch slope contrasts for nouns, adjectives, and verbs produced by speaker BM_01 are illustrated in the box-and-whisker plot in Figure 4.55 below. The pattern of contrasts in slope across syllables is more or less the same for the three lexical categories. For nouns and adjectives, the dominant trend – represented by the median values – is that pitch in σ_1 slopes downward, and pitch in σ_2 also slopes downward, but more steeply. For verbs, pitch in σ_1 is either upward or downward, and in σ_2

always slopes downward. In general, then, slope is “more downward” in σ_2 than in σ_1 ; the vector representing σ_2 is “below” the vector representing σ_1 .

Figure 4.55 BM_01 / Isolation: Pitch slope contrasts



This is very different than the pattern observed in isolation and frame forms produced by speaker BSh_03. In that case (Figure 4.21, Figure 4.22), the contrast in slope across syllables was governed by the contrast in pitch across syllables. Since non-verbs have a low-high pitch contrast, the pitch slope is “more upward” across syllables, rather than “more downward”, as here. BSh_03 verbs showed the opposite slope pattern, because they have the opposite pitch pattern.

This pattern is *not* much different, however, from those observed in isolation and frame forms produced by Rebkong Amdo speakers AR_04 and AR_05, discussed in Chapter 5. In that case, for the non-verbs, the contrast in slope across syllables

represents a robust correlate of stress. In this case, I attribute the very similar patterns to the effects of an utterance-level intonation pattern. The difference between the two scenarios – pitch slope as a correlate of stress *vs.* pitch slope as an effect of intonation – might be illuminated if I had recordings of BM_01 words embedded within a sentence frame. My prediction is that, in such hypothetical BM_01 frame forms, pitch would *not* slope downward in σ_2 , since it would no longer be at the end of the utterance. In Rebkong Amdo, even the frame forms exhibit a distinctive downward slope in σ_2 , because it is essential in conveying the stress pattern. Unfortunately, I was only able to record isolation forms from speaker BM_01, so I cannot test this argument at present.

The lack of recorded frame forms from speaker BM_01 is unfortunate, but not fatal. If one's objective is to identify an acoustic distinction between the two cases, one may be disappointed; if one's objective is to document the acoustic signature of the sharp, falling pitch slope perceived on stressed syllables in Rebkong Amdo, then this can certainly be done, as I show in section 5.3.

A primary difference between BM_01 and Rebkong Amdo, of course, is that in Rebkong Amdo, I *hear* a dramatic, “punchy” fall in pitch in σ_2 of non-verbs, and in BM_01 I do not.

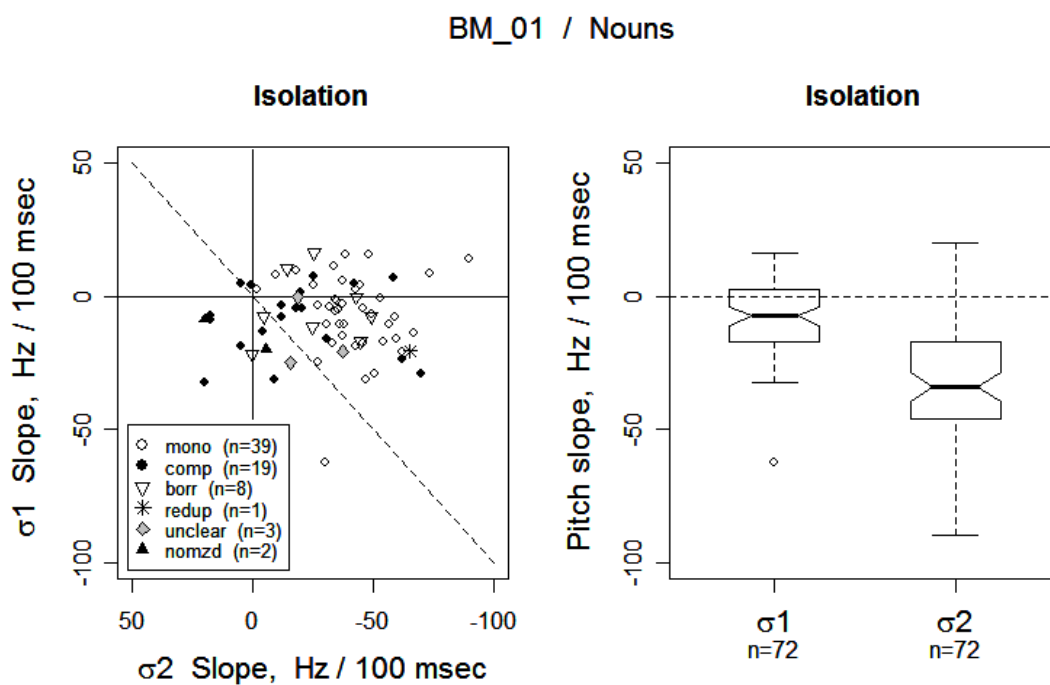
Details of pitch slope for BM_01 nouns, adjectives, and verbs are discussed in sub-sections 4.3.2.1, 4.3.2.2, and 4.3.2.3, respectively.

4.3.2.1 BM_01 Nouns

As noted, pitch slope patterns for nouns produced by speaker BM_01 are very different from those described for speaker BSh_03 (Figure 4.23). Here, slope is governed not by a rising pitch across syllables, but by a terminal falling intonation contour.

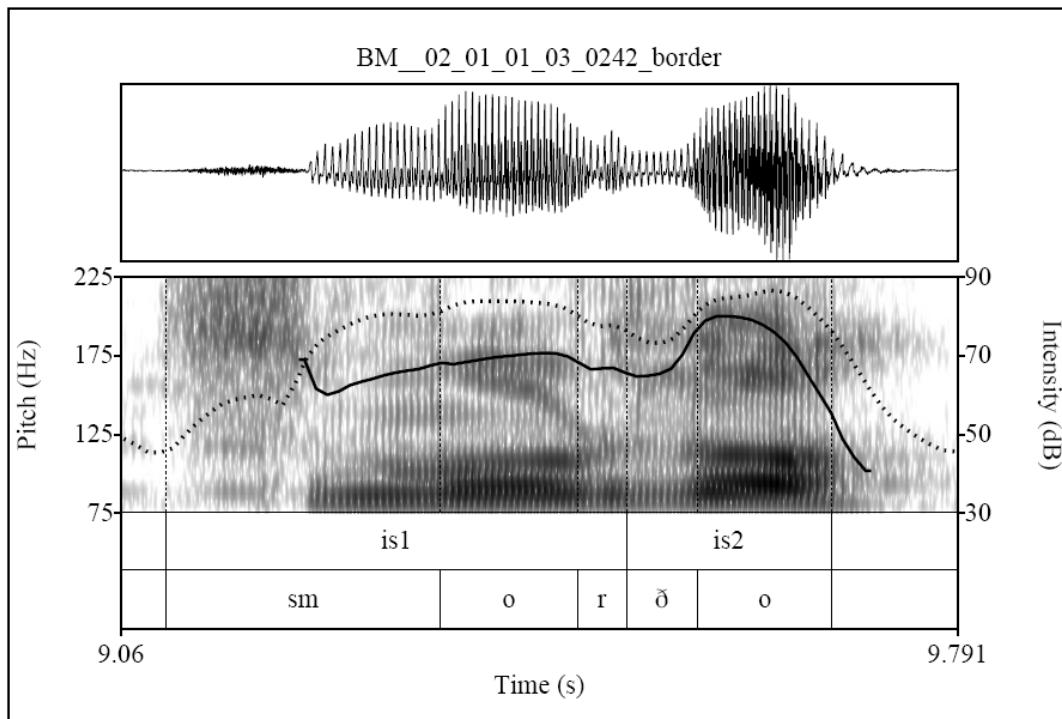
As shown in Figure 4.56 below, most of the BM_01 nouns fall to the right of the dashed line, indicating that pitch slope is “more downward” in σ_2 , i.e., that the vector representing σ_2 is always “below” the vector representing σ_1 .

Figure 4.56 BM_01 / Nouns: Pitch slope



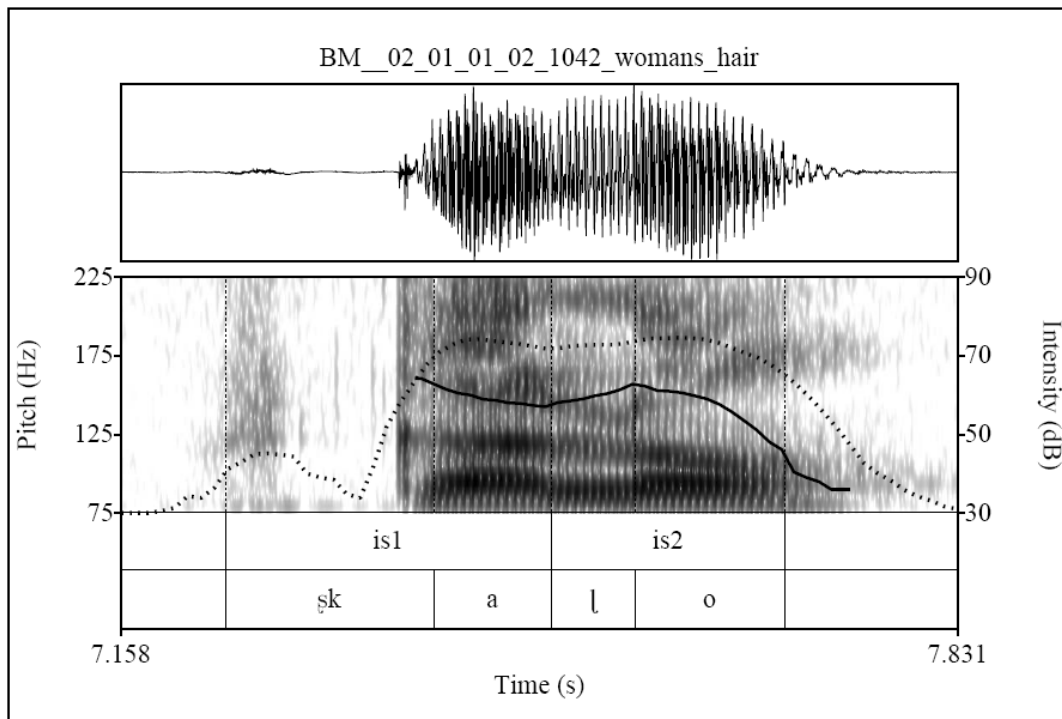
To illustrate, the pitch trace for the word [smor.'ðo] *mu.rdo* ‘border’ is provided in Figure 4.57 below. Here, pitch in σ_1 slopes upward at 6.8 Hz/100msec, but in σ_2 slopes downward, at -59 Hz/100msec.

Figure 4.57 BM_01 / Noun / Isolation / Pitch trace: [smor.'ðo] *mu.rdo* ‘border’



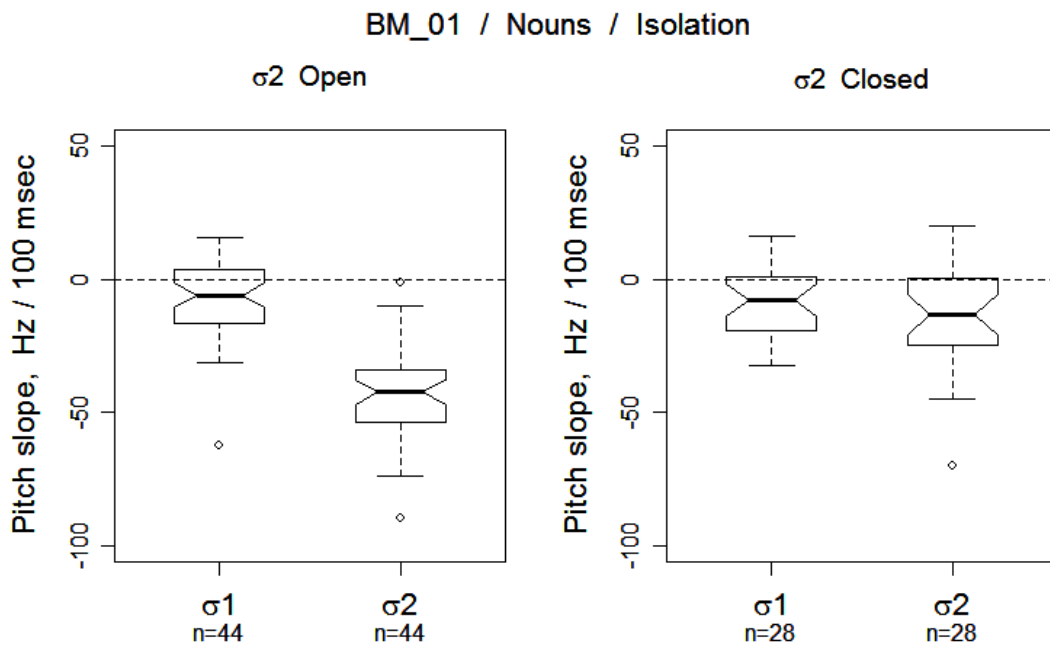
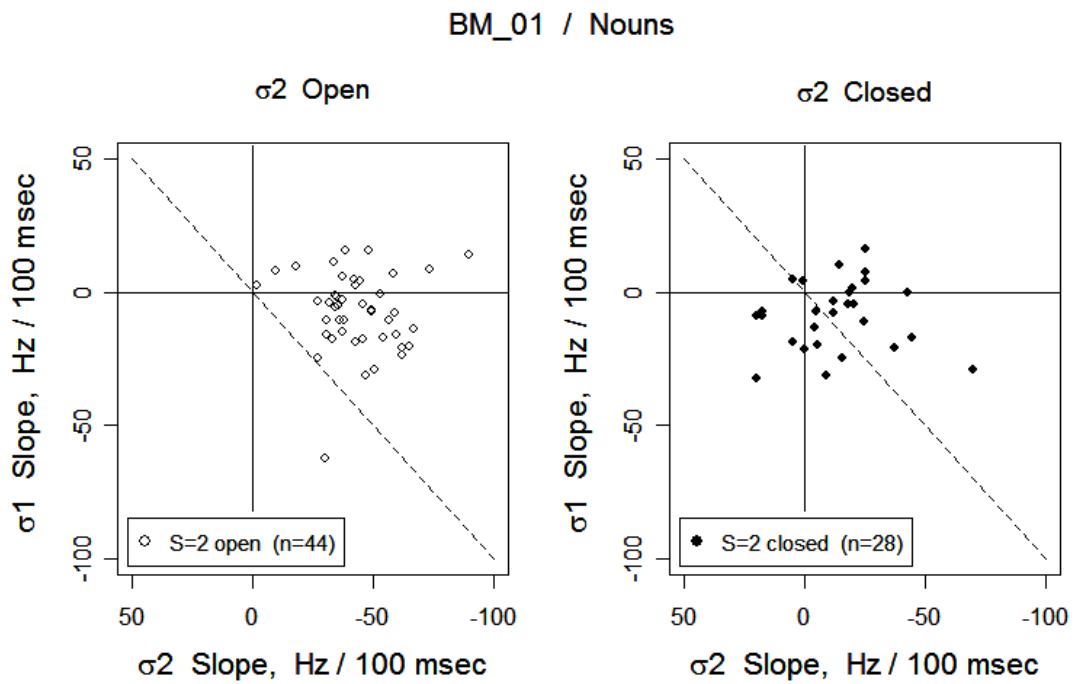
In the case of [ʃka.'lo] *skra.lo* ‘woman’s hair’ shown in Figure 4.58 below, pitch in σ_1 slopes downward at -11 Hz/100msec, and in σ_2 slopes downward even more steeply, at -31 Hz/100msec. Thus in this and the preceding example, the vector representing σ_2 is “below” the vector representing σ_1 .

Figure 4.58 BM_01 / Noun / Isolation / Pitch trace: [ska.'lo] *skra.lo* ‘woman’s hair’



In contrast, there were some nouns in the plot in Figure 4.56 which fall to the left of the dashed line. With only one exception, it turns out that these are words in which σ_2 is closed, as illustrated in Figure 4.59 below. When σ_2 is open, all nouns plot to the right of the dashed line. (This is actually similar to the pattern observed in Rebkong Amdo nouns, as illustrated in Figure 4.20 on page 186.) When σ_2 is closed, points are evenly distributed to either side of the dashed line.

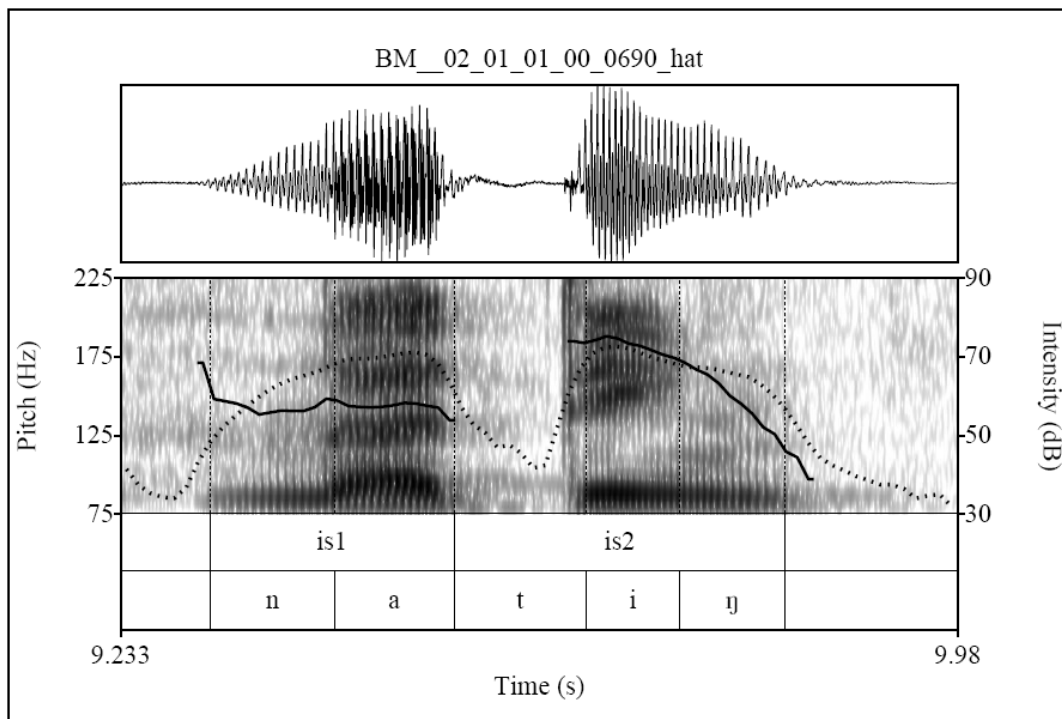
Figure 4.59 BM_01 / Nouns: Pitch slope vs. $\sigma 2$ closure



In the box-and-whisker plots at the bottom of the figure, where σ_2 is open, there is a clear difference in the slope patterns across syllables. Pitch slopes “more downward” in σ_2 than in σ_1 , as described above. The nouns with a closed σ_2 show almost complete overlap of the notches, boxes, and whiskers; there is no evidence of any consistent pattern in pitch slope contrast across the two syllables.

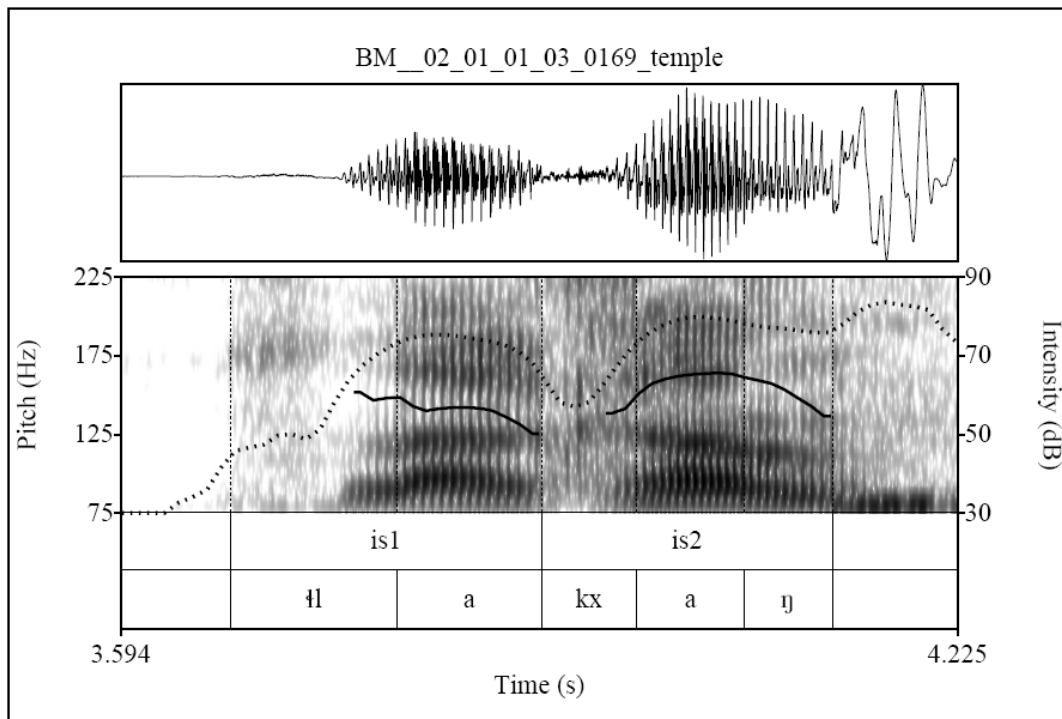
Figure 4.60 below [na.'tɪŋ] ?? ‘hat’ illustrates a noun with σ_2 closed which behaves just like the nouns with σ_2 open – i.e., it falls to the right of the dashed line. In σ_1 , pitch is nearly level, with a slope of -0.39 Hz/100msec. In σ_2 , pitch slopes downward at -19 Hz/100msec in the vowel; this downward slope– a final falling intonation contour – continues and even steepens in the sonorant coda.

Figure 4.60 BM_01 / Noun / Isolation / Pitch trace: [na.'tɪŋ] ?? ‘hat’



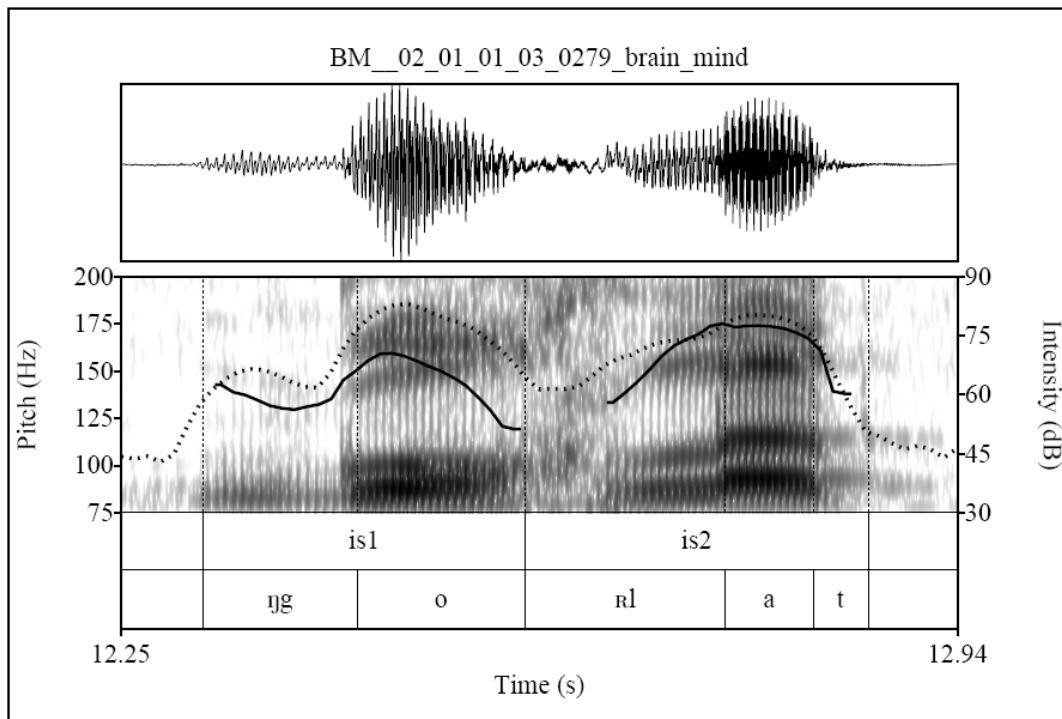
However, not all words in which σ_2 is closed by a sonorant coda exhibit this falling intonation contour, as illustrated by [ɬla.'kʰaŋ] *lha.khang* ‘temple’ in Figure 4.61 below. This word plots to the left of the dashed line in Figure 4.59.

Figure 4.61 BM_01 / Noun / Isolation / Pitch trace: [ɬla.'kʰaŋ] *lha.khang* ‘temple’



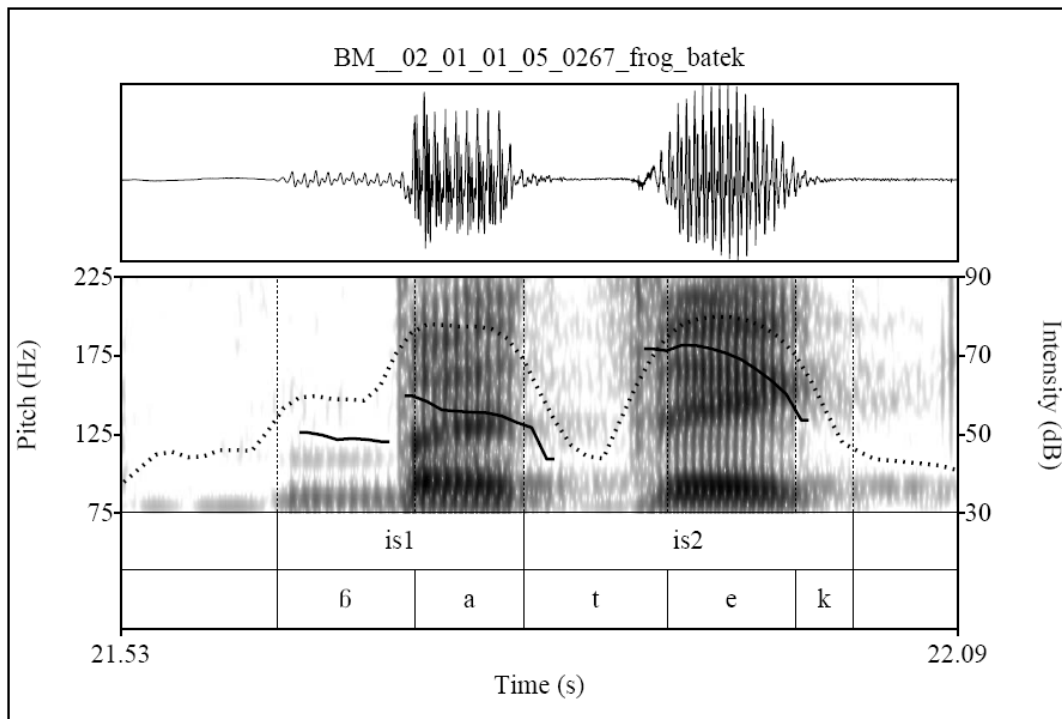
In some cases, as in [ŋgo.'ɾlat] *mgo.glad* ‘brain, mind’ shown in Figure 4.62 below, a voiceless stop coda in σ_2 appears to block the final falling intonation. In σ_1 pitch slopes downward at -31 Hz/100msec, and in σ_2 slopes downward at only -9 Hz/100msec. That is, σ_1 slopes “more downward” than σ_2 , so the point representing this word also falls to the left of the dashed line in Figure 4.59.

Figure 4.62 BM_01 / Noun / Isolation / Pitch trace: [ŋgo.'ɾlat] *mgo.glad* 'brain, mind'



On the other hand, a voiceless stop coda in σ_2 does not always block a falling final intonation contour, as illustrated by the pitch trace for [ʙa.'tek^h] ?? 'frog' in Figure 4.63 below. This word plots to the right of the dashed line in Figure 4.59.

Figure 4.63 BM_01 / Noun / Isolation / Pitch trace: [ba.'tek^h] ?? ‘frog’



We saw in Figure 4.59 that nouns with a closed σ_2 fall to either side of the dashed line. What these last four examples demonstrate is that this truly represents a random distribution, that pitch may randomly slope “more downward” in one syllable or else the other. When σ_2 is open, on the other hand, pitch is always more downward in σ_1 .

These patterns are reflected in paired-sample t-tests, which are summarized in Table 4.15 below. When σ_2 is open, the contrast in pitch slope across syllables is statistically highly significant ($p \ll 0.05$), with a mean difference in slope of -36 Hz/100msec. When σ_2 is closed, the difference in pitch slope across syllables is *not* statistically significant, with $p > 0.05$.

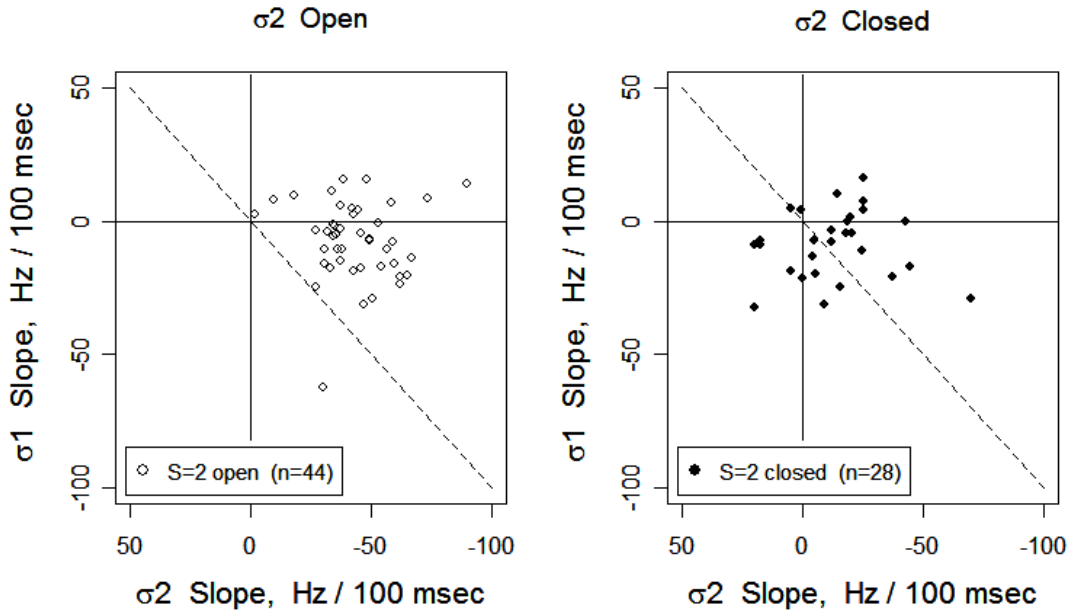
Table 4.15 BM_01 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BM_01	σ 2 open	-10.9838	43	4.637e-14	-36	-42	-29
	σ 2 closed	-0.9073	27	0.3723	-4.2	-14	5.4

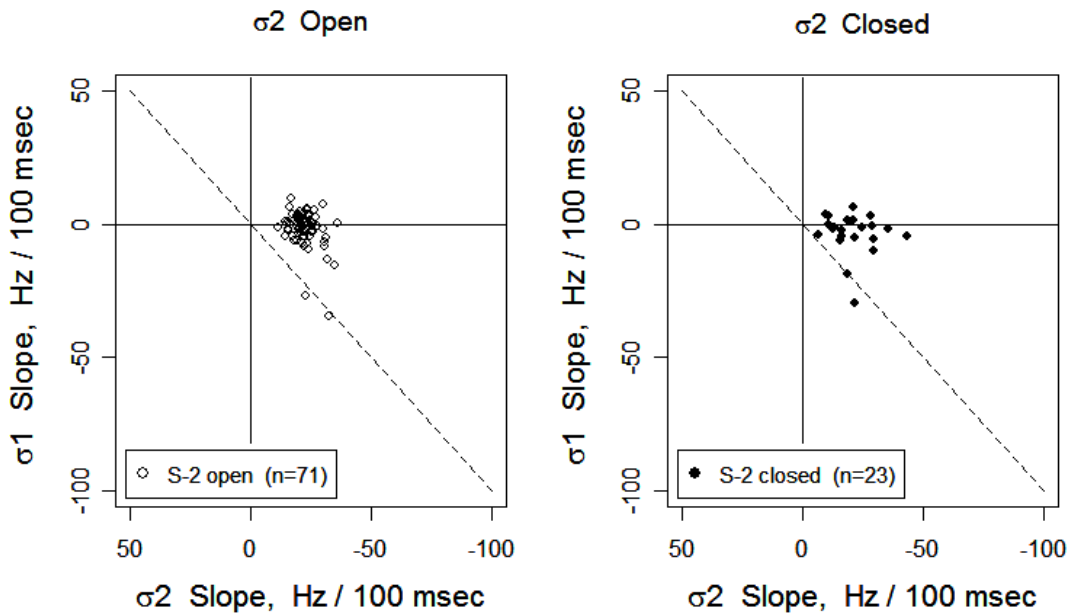
This distinction in the behavior of BM_01 nouns on the basis of σ 2 closure is actually different than what is observed in Rebkong Amdo, where pitch slope is a robust correlate of stress. In Figure 4.64 below, pitch slope in BM_01 nouns plotted in terms of σ 2 closure is compared to pitch slope in AR_05 nouns. The top part of this figure is a duplicate of Figure 4.59; the bottom part is a duplicate of Figure 4.20, except now compressed to the same scale on the x- and y-axes as for BM_01.

Figure 4.64 BM_01 / Nouns: Pitch slope vs. σ_2 closure, compared to Rebkong Amdo

BM_01 / Nouns



AR_05 / Nouns / Isolation



As shown, the slope patterns exhibited by BM_01 nouns with σ_2 open are not substantively different from those of speaker AR_05. In both cases, all points fall to the right of the dashed line, with a “more downward” slope in σ_2 . For speaker BM_01, this pattern is the casual consequence of an unimpeded utterance-final falling intonation; for Rebkong Amdo speaker AR_05, it reflects the speaker’s fairly precise manipulation of pitch slope in order to convey stress. Of course, there is a vast difference in the range of variability for the two speakers. This may reflect the overall inconsistency of the recording with speaker BM_01. On the other hand, it may be more significant: perhaps it results from the fact that speaker AR_05 aims to achieve a target slope contrast, while BM_01 is instead vaguely allowing pitch to taper off.

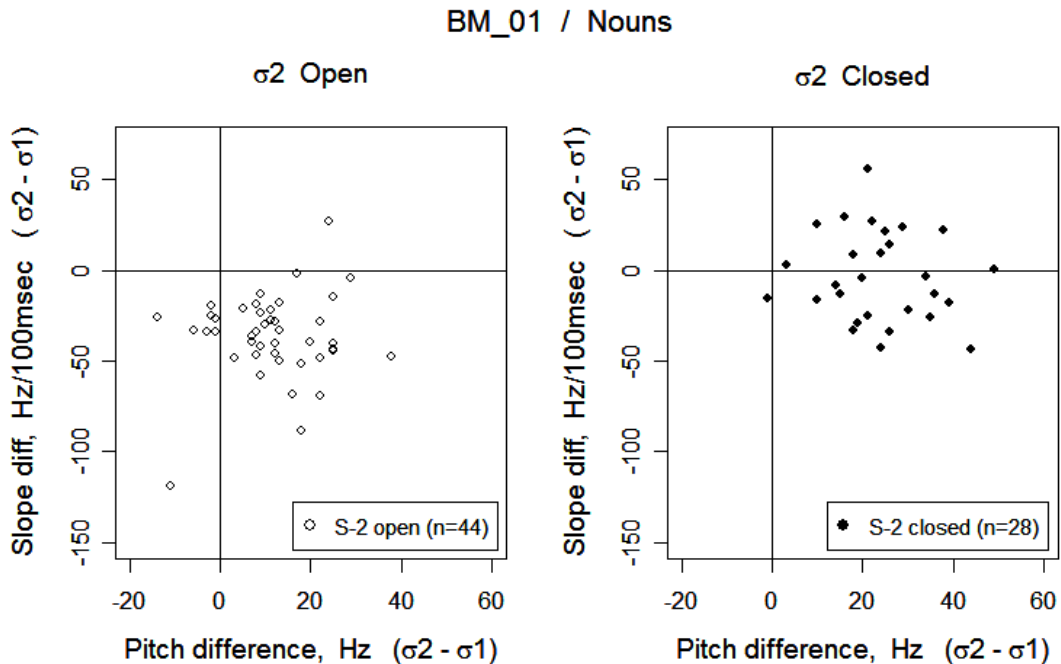
The random slope contrasts exhibited by BM_01 nouns with σ_2 closed are simply not observed in the nouns produced by speaker AR_05. In Rebkong Amdo, a robust and consistent contrast in pitch slope is of primary importance in conveying a robust and consistent stress contrast; thus pitch always slopes “more downward” in σ_2 , regardless of syllable closure. This divergence in the behavior of nouns with a closed σ_2 may actually be evidence of an acoustic distinction between pitch slope as a reflex of intonation, and pitch slope as a cue for stress, even without the crucial frame forms, as discussed previously.

Finally, the relationship between pitch and pitch slope for BM_01 nouns is illustrated in Figure 4.65 below. Almost all points fall to the right of the y-axis; as demonstrated in section 4.2.2.1, pitch generally rises from σ_1 to σ_2 , so the pitch

difference is positive. When σ_2 is closed, the pitch contrast is sometimes elevated, by virtue of the fact that much of the utterance-final declination is borne by a sonorant coda. The distribution with respect to the x-axis reflects the slope patterns discussed above: when σ_2 is open, slope is always more downward in σ_2 ; when σ_2 is closed, the slope contrast across syllables is random.

This pattern is very different from what was observed in Figure 4.39 (page 212) for speaker BSh_03. There, most points fell in the upper right quadrant, reflecting the fact that pitch slope was governed by the pitch contrast across syllables. Here, pitch slope is dominated by final intonation. This pattern is also different from what is observed in Rebkong Amdo, where pitch slope actually contrasts with pitch.

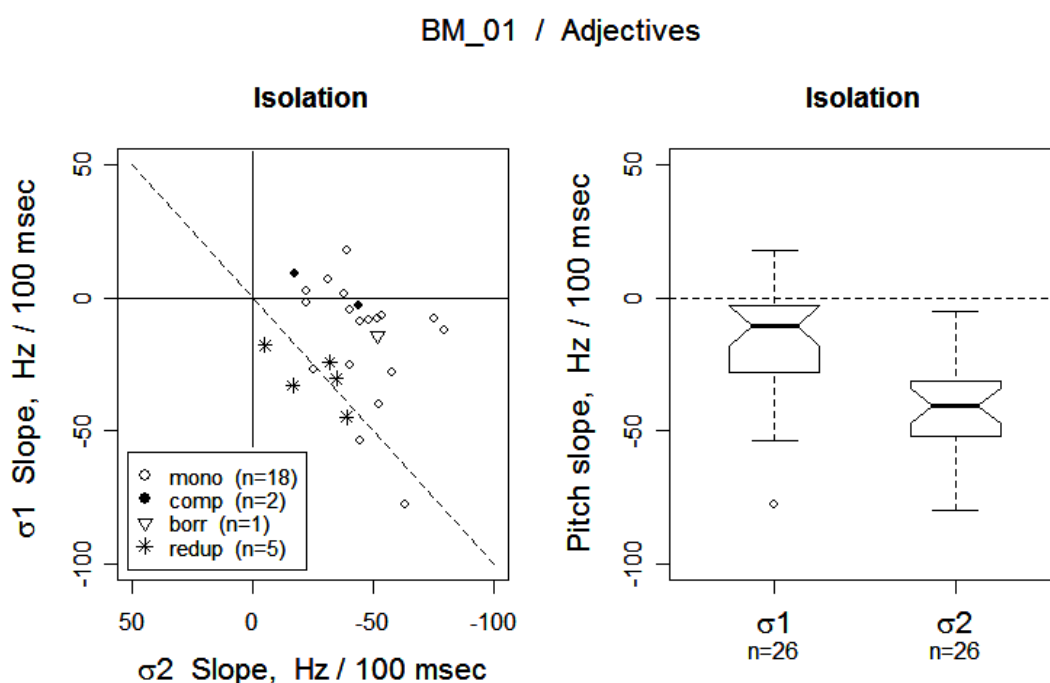
Figure 4.65 BM_01 / Nouns: Pitch difference vs. pitch slope difference, by σ_2 closure



4.3.2.2 BM_01 Adjectives

The pitch slope patterns observed in BM_01 adjectives are very similar to those described above for nouns. As shown in Figure 4.66 below, most adjectives plot to the right of the dashed line, indicating that the pitch slope is “more downward” in σ_2 than in σ_1 . It is interesting to note that the five reduplicated forms are distributed close to and to either side of the dashed line, indicating random slope contrasts for this control group.

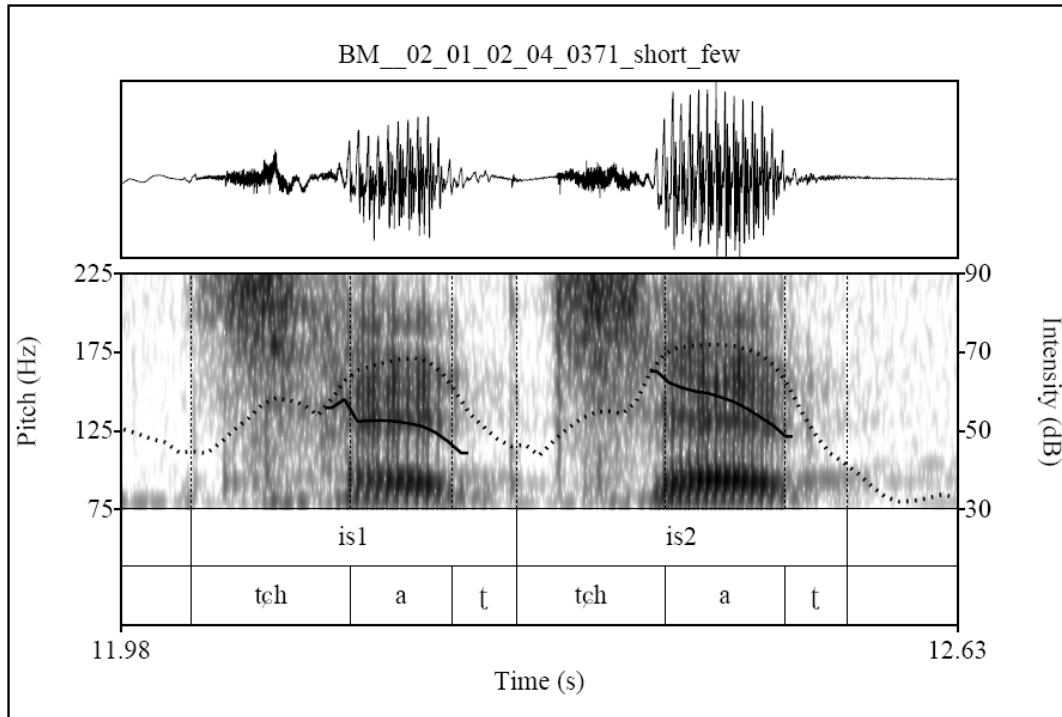
Figure 4.66 BM_01 / Adjectives / Isolation: Pitch slope



As illustration, the reduplicated form [tʰat.tʰat] *chad.chad* ‘short, few’ is shown in Figure 4.67 below. In σ_1 , pitch slopes downward at -24 Hz/100msec, and in

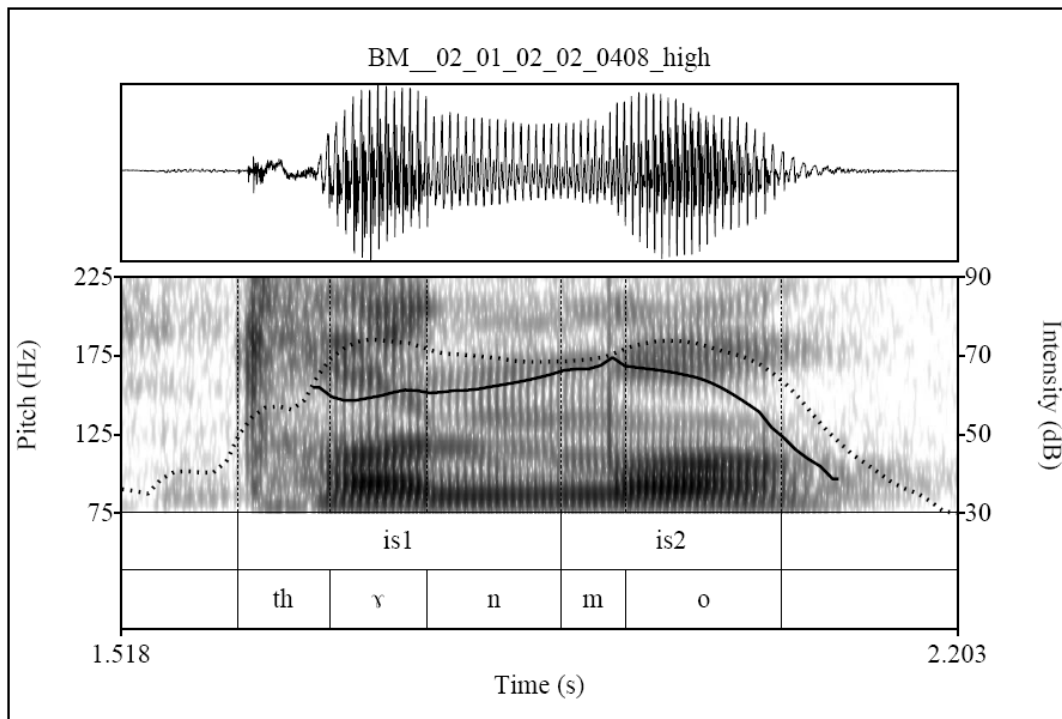
σ_2 slopes downward more steeply, at -32 Hz/100msec. This adjective plots to the right of the dashed line in Figure 4.66.

Figure 4.67 BM_01 / Adjectives / Isolation / Pitch trace: [tɕ^hat.'tɕ^hat] *chad.chad* ‘short, few’



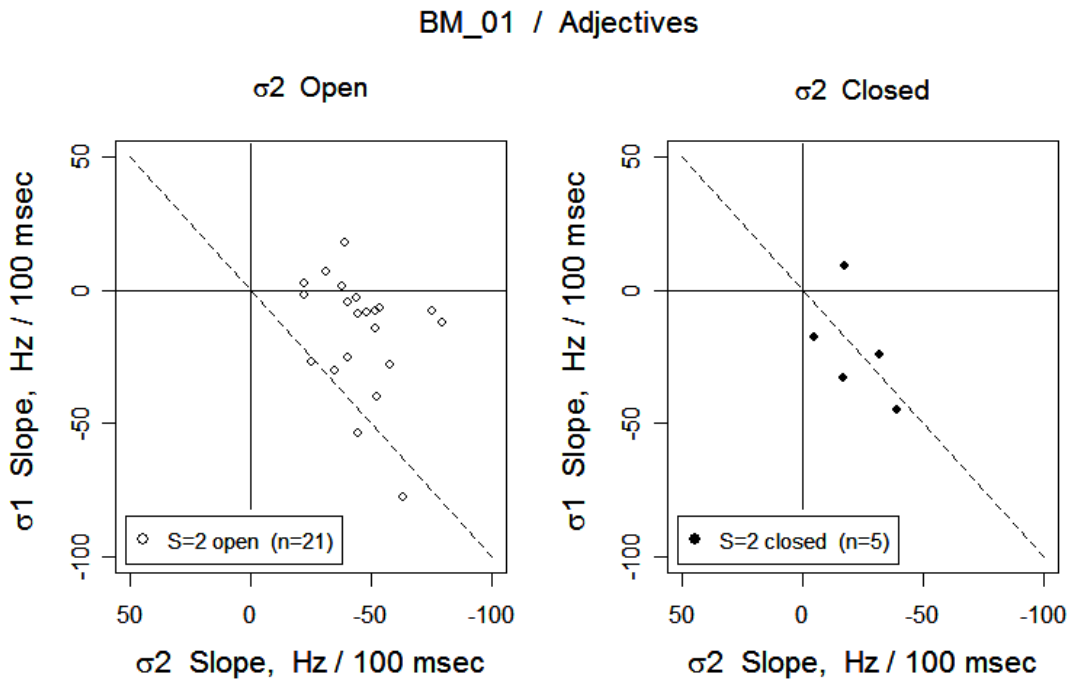
In the monomorphemic adjective [t^hʌn.'mo] *mthon.mo* ‘high’, the pitch slope is slightly upward in σ_1 , at 6.9 Hz/100msec, and downward quite steeply in σ_2 at -31 Hz/100msec. The difference in pitch slope is $(-31) - (6.9) = -38$ Hz/100msec. This point plots well to the right of the dashed line, and above the x-axis.

Figure 4.68 BM_01 / Adjectives / Isolation / Pitch trace: [tʰʌn.'mo] *mthon.mo* 'high'



Adjectives are plotted in terms of the closure of σ_2 in Figure 4.69 below. As was the case with nouns, when σ_2 is open, nearly all points fall to the right of the dashed line. When σ_2 is closed, points are distributed to either side of the line.

Figure 4.69 BM_01 / Adjectives: Pitch slope vs. $\sigma 2$ closure



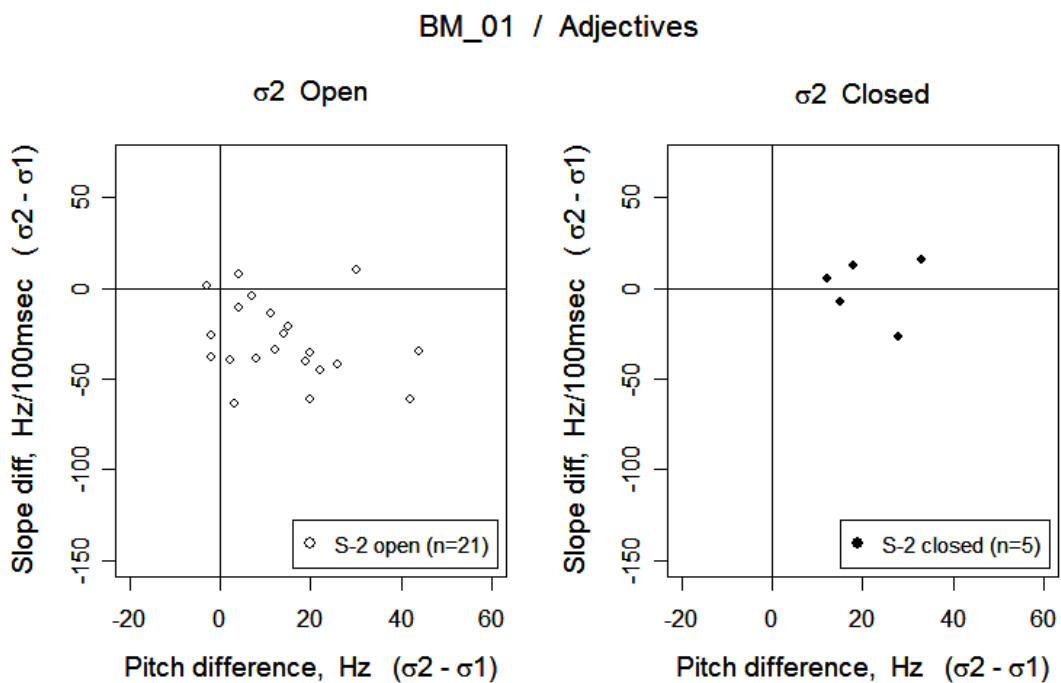
Paired-sample t-tests, summarized in Table 4.16 below, corroborate these observations. When $\sigma 2$ is open, the difference in pitch slopes across syllables is highly significant ($p \ll 0.05$), with an average difference of -30 Hz/100msec. When $\sigma 2$ is closed, the difference in slopes is not significant: $p > 0.05$, and the 95% confidence interval includes zero.

Table 4.16 BM_01 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BM_01	σ2 open	-6.0879	20	5.976e-06	-30	-40	-20
	σ2 closed	0.0263	4	0.9803	0.20	-21	22

When pitch differences are compared to pitch slope differences, as in Figure 4.70 below, the pattern is again quite similar to that observed in BM_01 nouns. The pitch difference across syllables is sometimes exaggerated, when much of the downward slope is borne by a sonorant coda.

Figure 4.70 BM_01 / Adjectives: Pitch difference vs. pitch slope difference

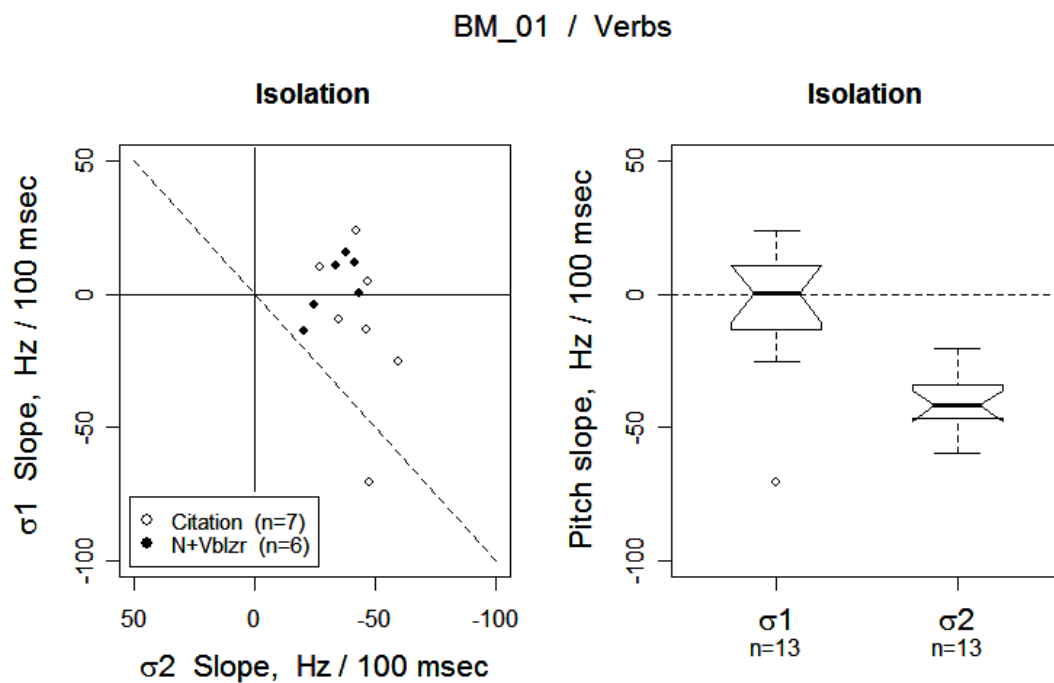


4.3.2.3 BM_01 Verbs

BM_01 verbs exhibit the same pitch slope patterns as the BM_01 non-verbs. For nouns and adjectives, as discussed above, I attribute the pitch slope pattern to the utterance-level intonation pattern. In the case of verbs, the effects of this factor cannot be distinguished from the effects of a second factor: the contrast in pitch across syllables, which was also observed in the case of BSh_03 verbs.

As illustrated in Figure 4.71 below, in $\sigma 1$ of verbs, pitch may slope upward or downward, but in $\sigma 2$ always slopes downward. With the exception of [ʔes.pa] *shes.pa* ‘to know’, all points fall to the right of the dashed line. There is no distinction between the behavior of citation forms and N+Vblzr forms.

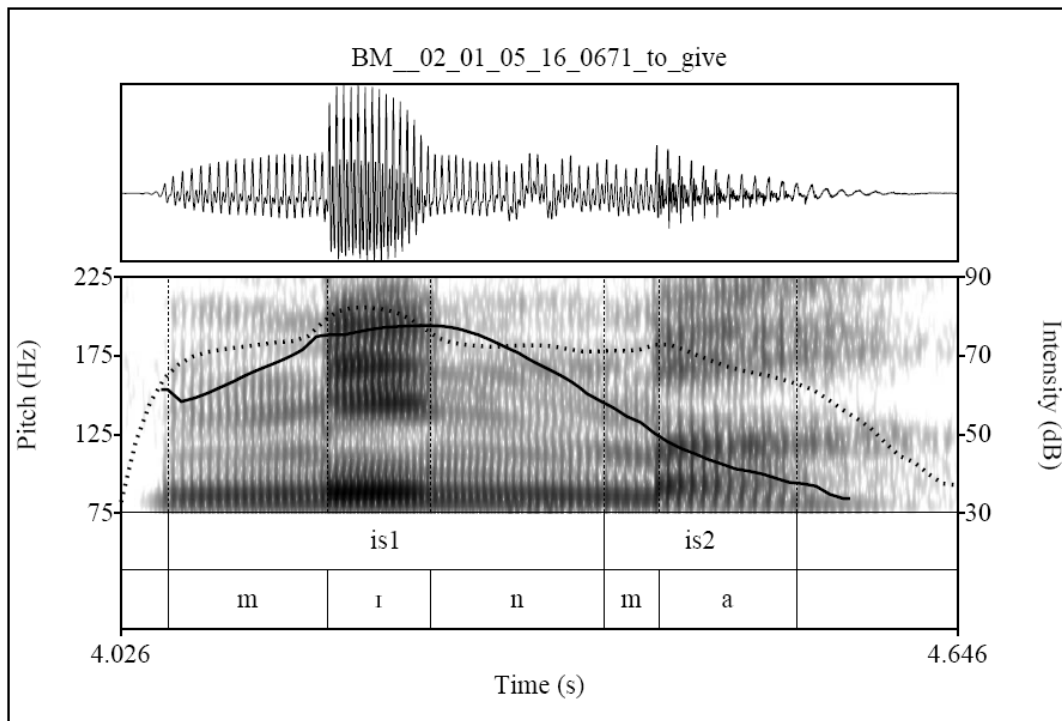
Figure 4.71 BM_01 / Verbs: Pitch slope



The pitch trace for the verb [ʔmin.ma] *sbyin.ma?* ‘to give’ in Figure 4.72 below is an example of a citation form in which pitch slopes upward in $\sigma 1$, and downward in $\sigma 2$. The mean pitch over the medial 50% of the vowel was measured as 192 Hz in $\sigma 1$, and 105 Hz in $\sigma 2$. This dramatic -87 Hz pitch contrast across syllables is the primary correlate of the perceived $\sigma 1$ stress pattern, and it plays a role in shaping the pitch slope pattern across syllables. As shown below, the falling pitch

slope on σ_2 is only an extension of the large-scale fall in pitch from σ_1 to σ_2 . In σ_1 , pitch slopes gently upward at 10 Hz/100msec; in σ_2 , pitch slopes downward at -27 Hz/100msec. The difference in slope is calculated as $(-27) - (10) = -37$ Hz/100msec. Of course, a second factor which may influence pitch slope here is that the word was produced in isolation, where a falling terminal contour is often observed.

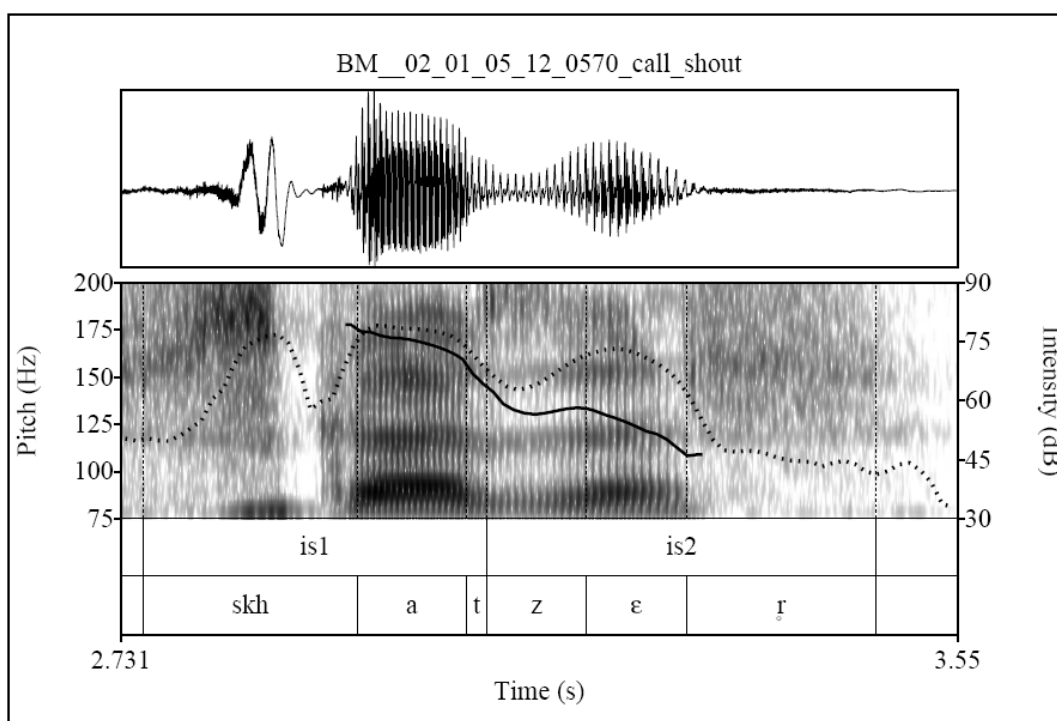
Figure 4.72 BM_01 / Verbs / Pitch trace: [ʼmɪn.ma] *sbyin.ma* ? ‘to give’



In contrast, the pitch trace for the N+Vblzr form [ʼskʰat.zer] *skad.zer* ‘to call, to shout’ is a case where pitch slopes downward in σ_1 , and downward more steeply in σ_2 , as illustrated by the pitch trace in Figure 4.73 below. The average pitch is 169 Hz in σ_1 , and 124 Hz in σ_2 , for a pitch difference across syllables of -45 Hz. Pitch slope

was measured as -14 Hz/100msec in σ_1 , and -20 Hz/100msec in σ_2 , for a slope difference of -6 Hz/100msec.

Figure 4.73 BM_01 / Verbs / Pitch trace: [ʰskʰat.zer] *skad.zer* ‘to call, to shout’



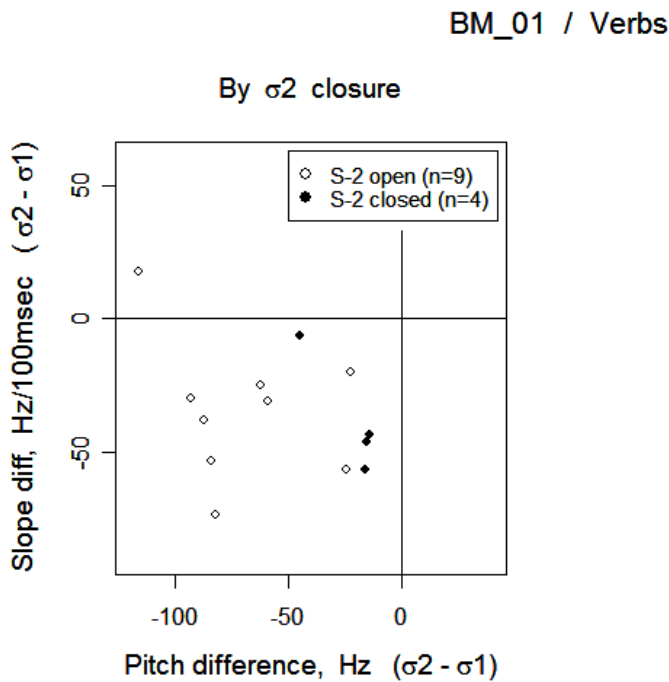
A paired-sample t-test comparing pitch slope across syllables is summarized in Table 4.17 below. The difference in slope across syllables is statistically significant ($p < 0.05$), and averages -34 Hz/100msec.

Table 4.17 BM_01 / Verbs / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
BM_01	Isolation	-5.2878	12	0.0001921	-34	-49	-20

The correspondence between pitch and pitch slope for BM_01 verbs is evident in Figure 4.74 below. Because there is such a dramatic drop in pitch across syllables, the slope difference is also always negative (again, except for ['çes.pa] *shes.pa* 'to know'). Virtually all points thus fall in the lower left quadrant, reflecting this relationship.

Figure 4.74 BM_01 / Verbs: Pitch difference vs. pitch slope difference

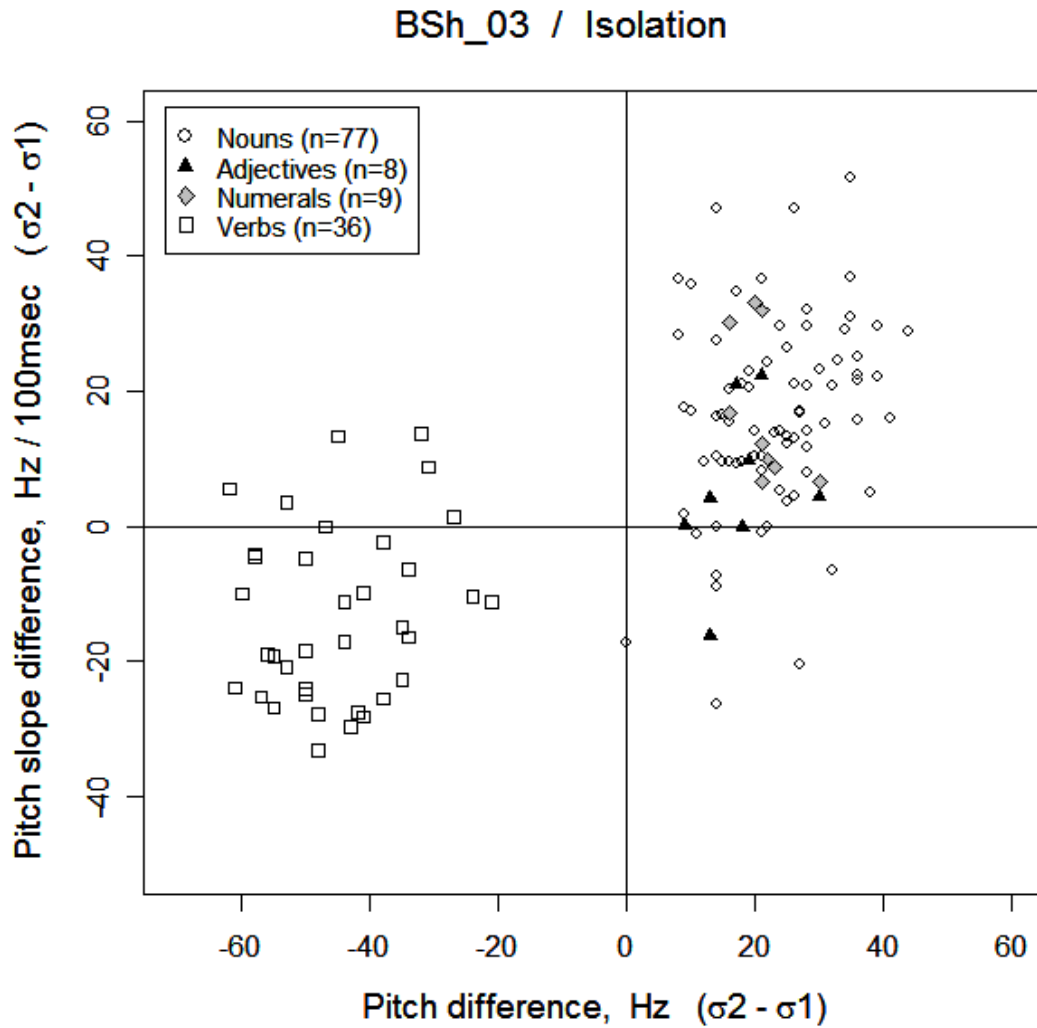


4.3.3 The interaction of pitch and pitch slope

In Figure 4.75 below, BSh_03 isolation forms from all lexical categories are plotted in terms of pitch differences and pitch slope differences. As shown, verbs vs. non-verbs are distinguished by both of these factors, with verbs plotting in the lower

left quadrant and non-verbs plotting in the upper right quadrant. In both cases, this reflects the fact that pitch slope is an epiphenomenon of pitch. For verbs, pitch is higher in σ_1 than in σ_2 ; i.e., pitch decreases across syllables. The “more downward” slope in σ_2 is a reflection and extension of that decrease. For non-verbs, pitch is higher in σ_2 than in σ_1 ; i.e., pitch increases across syllables. The “more upward” slope in σ_2 is a reflection and extension of that increase. There is no overlap of the two groups; prosodically, verbs are thoroughly distinct from non-verbs.

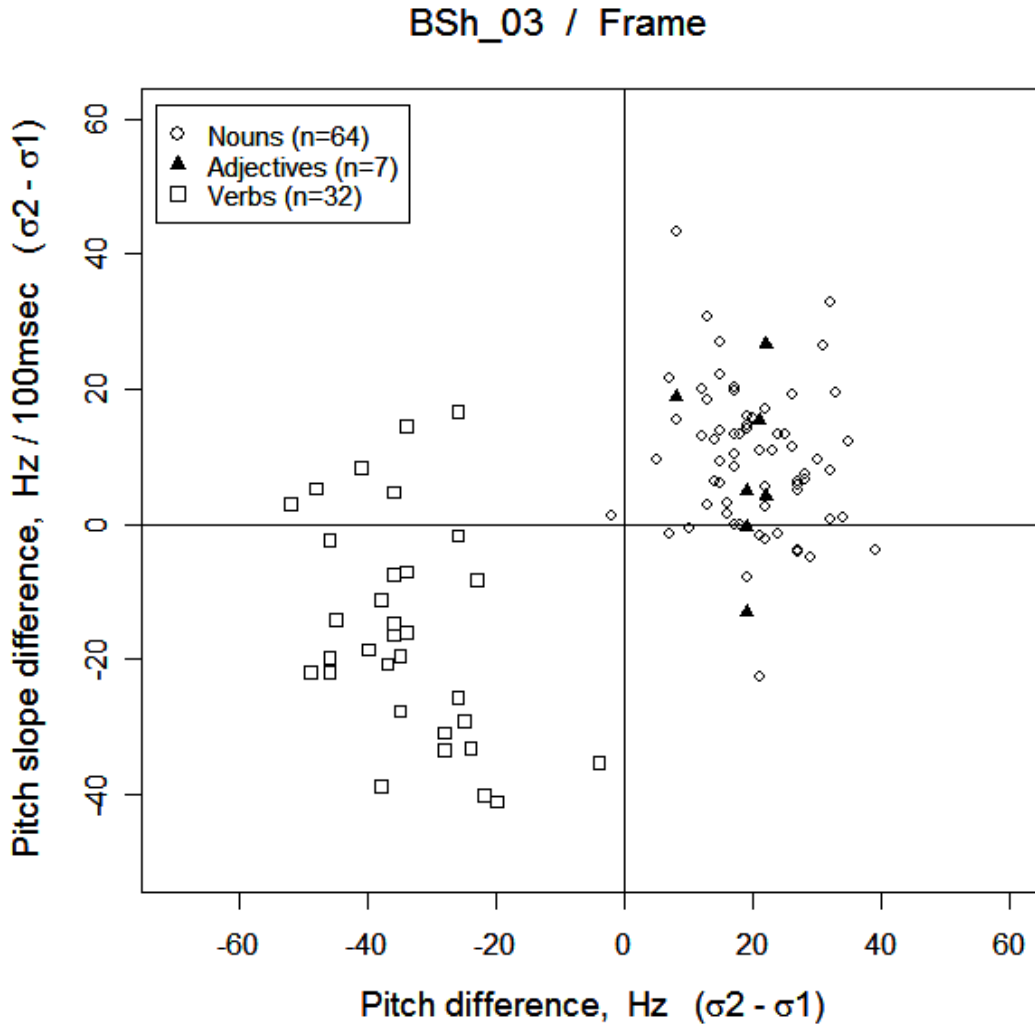
Figure 4.75 BSh_03 / Isolation: Pitch differences vs. pitch slope differences



Frame forms of BSh_03 verbs and non-verbs are plotted in Figure 4.76 below.

The same general patterns and distinctions pertain here.

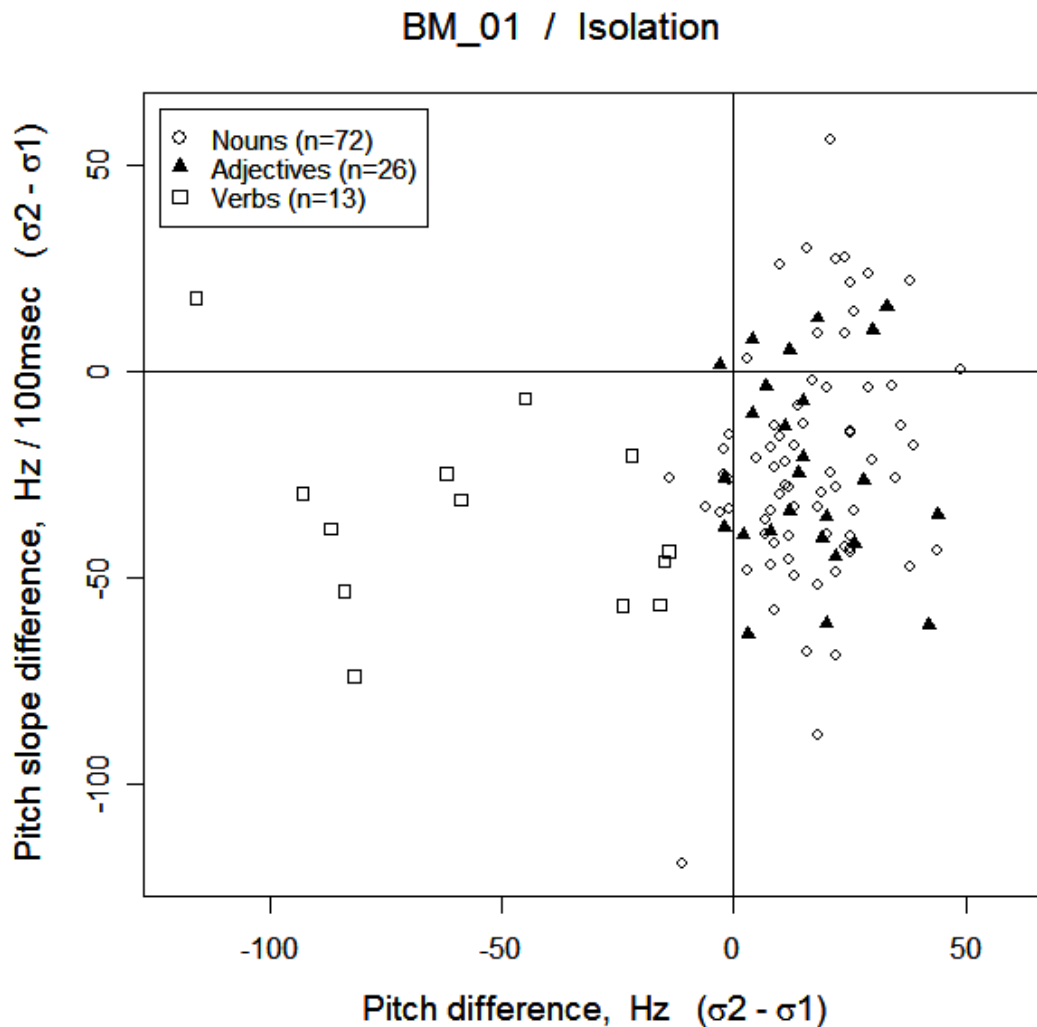
Figure 4.76 BSh_03 / Frame: Pitch differences vs. pitch slope differences



For speaker BM_01, the relationship is considerably different: verbs and non-verbs are distinguished by pitch differences, but not by pitch slope differences. As was the case for speaker BSh_03, verbs plot in the lower left quadrant of the graph: pitch is higher in σ_1 than in σ_2 , and this pitch decrease across syllables is extended as a downward pitch slope in σ_2 . Non-verbs, in this case, also fall below the x-axis, plotting in the lower right quadrant of the graph. (Those which plot above the x-axis

have a closed σ_2 , as discussed in sections 4.3.2.1 and 4.3.2.2.) Pitch increases from σ_1 to σ_2 , to convey stress, but then slopes downward in σ_2 as part of a natural utterance-final falling intonation.

Figure 4.77 BM_01 / Isolation: Pitch differences vs. pitch slope differences



In the case of speaker BSh_03, pitch slope patterns cannot be distinguished from pitch patterns, for both verbs and non-verbs. In the case of speaker BM_01,

pitch slope patterns cannot be distinguished from pitch patterns, for verbs, or from intonation patterns, for non-verbs. Because pitch slope thus lacks contextual significance, it is never an acoustic correlate of stress for either Balti speaker, for any lexical category.

4.4 Intensity

For both Balti speakers, intensity is a strong cue for the σ_1 stress perceived on verbs, but is not a correlate of the σ_2 stress perceived on nouns, adjectives, or numerals. In limited and controlled subsets of the data, where contrasts in vowel height are neutralized, intensity is sometimes higher in σ_2 . But even here, the contrast is not consistent or strong enough to be considered an indicator of stress.

The analyses for speakers BSh_03 and BM_01 are discussed in detail in sections 4.4.1 and 4.4.2, respectively.

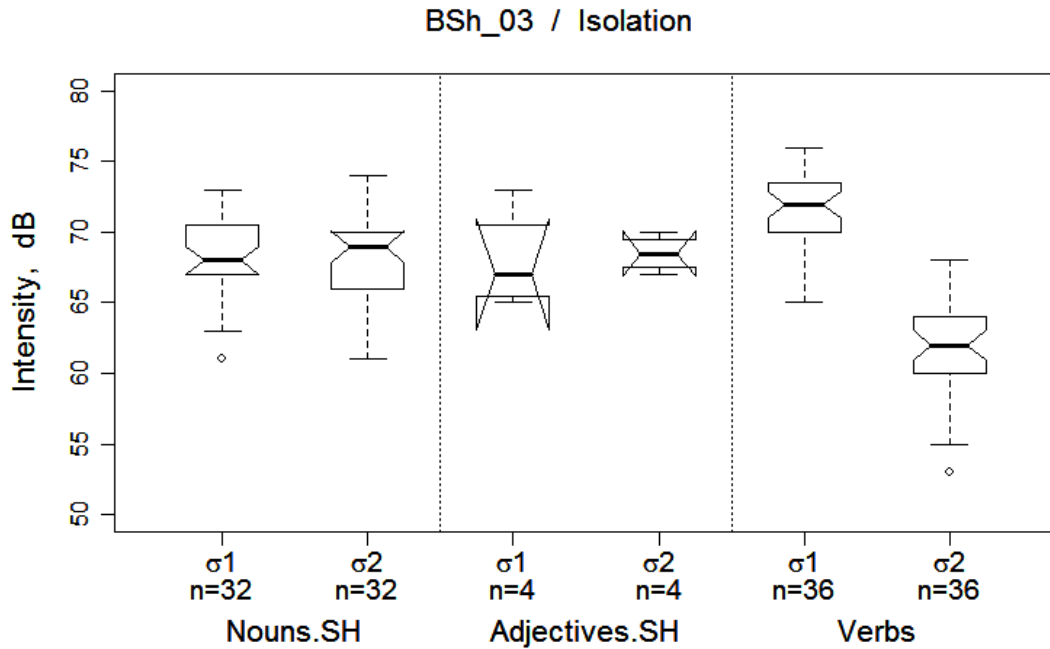
4.4.1 Intensity for speaker BSh_03

For both isolation and frame forms produced by speaker BSh_03, intensity is a strong acoustic correlate of first-syllable stress for verbs, but is not an acoustic correlate for the second-syllable stress perceived on nouns and adjectives.

The general pattern of intensity contrasts across syllables in isolation forms is illustrated in Figure 4.78 below. For each lexical category, the box-and-whisker plot is based on the data set which was crucial to the assessment of the role played by intensity in conveying stress. For nouns and adjectives, this means that only the control groups consisting of words with vowels of the same height in both syllables

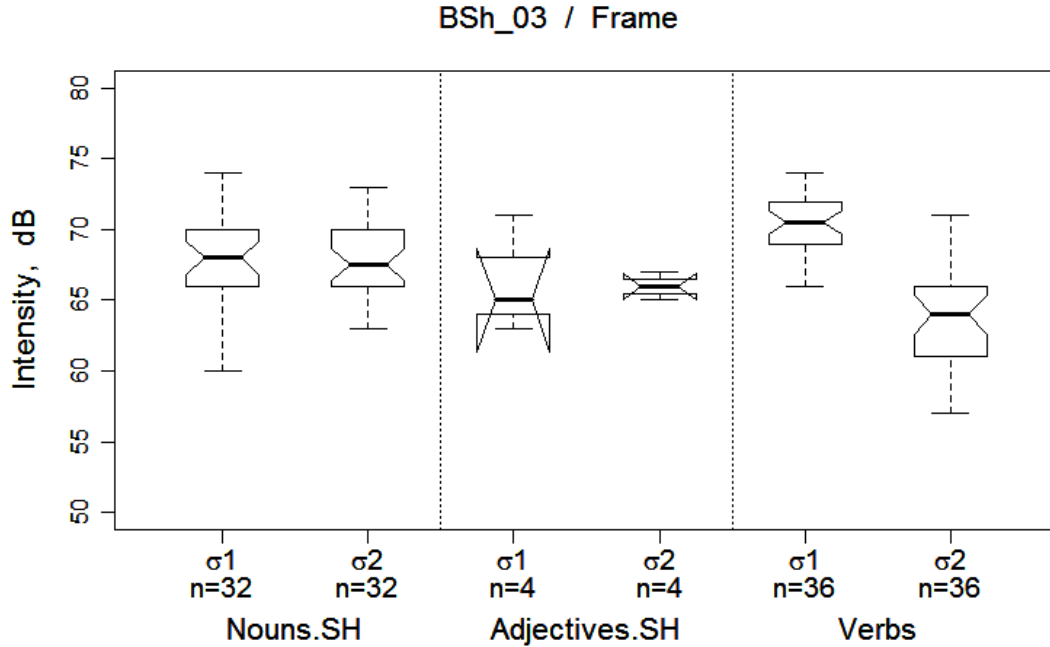
are shown. For verbs, the pattern was so robust and consistent that any variation as a function of vowel height is moot, so all tokens are plotted.

Figure 4.78 BSh_03 / Isolation: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)



Intensity contrasts for BSh_03 frame forms are similar, as illustrated in Figure 4.79 below. Again, only the subsets on which the final analysis was based are plotted.

Figure 4.79 BSh_03 / Frame: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)

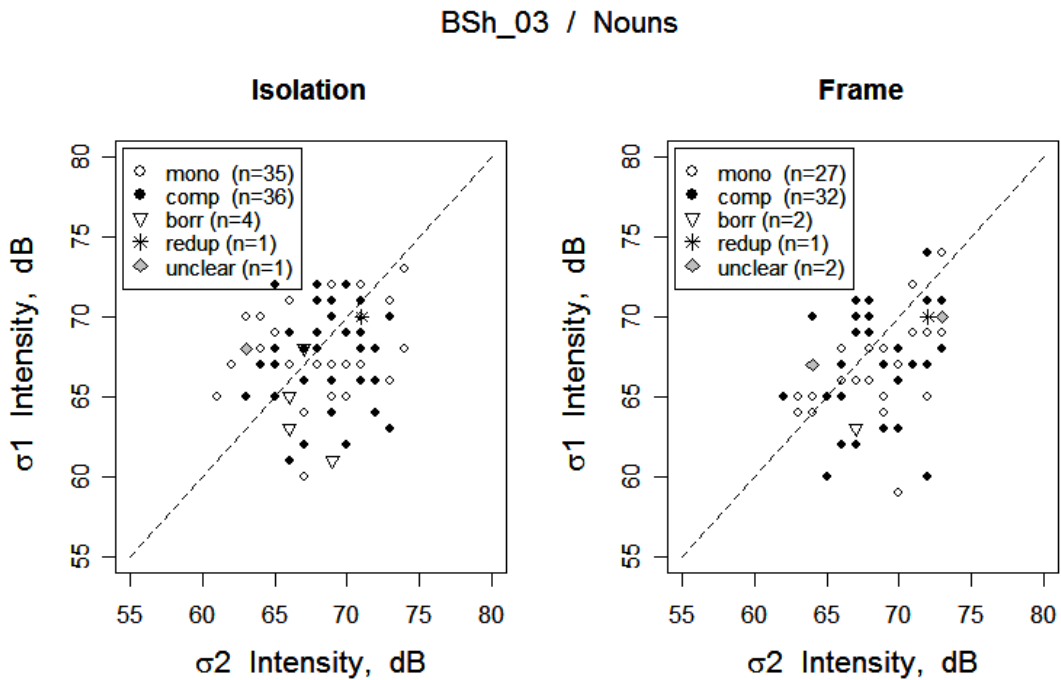


Intensity data for BSh_03 nouns, adjectives, numerals, and verbs are discussed in greater detail in sections 4.4.1.1 through 4.4.1.4 below.

4.4.1.1 BSh_03 Nouns

If intensity served as an acoustic cue for stress in BSh_03 nouns, one would expect to see a consistently higher intensity on σ_2 . However, this is not the case for either isolation or frame forms. As shown in Figure 4.80 below, points fall to either side of the dashed line, indicating that intensity is sometimes higher on σ_1 and sometimes higher on σ_2 . Morphological structure does not appear to influence the distribution.

Figure 4.80 BSh_03 / Nouns: Intensity

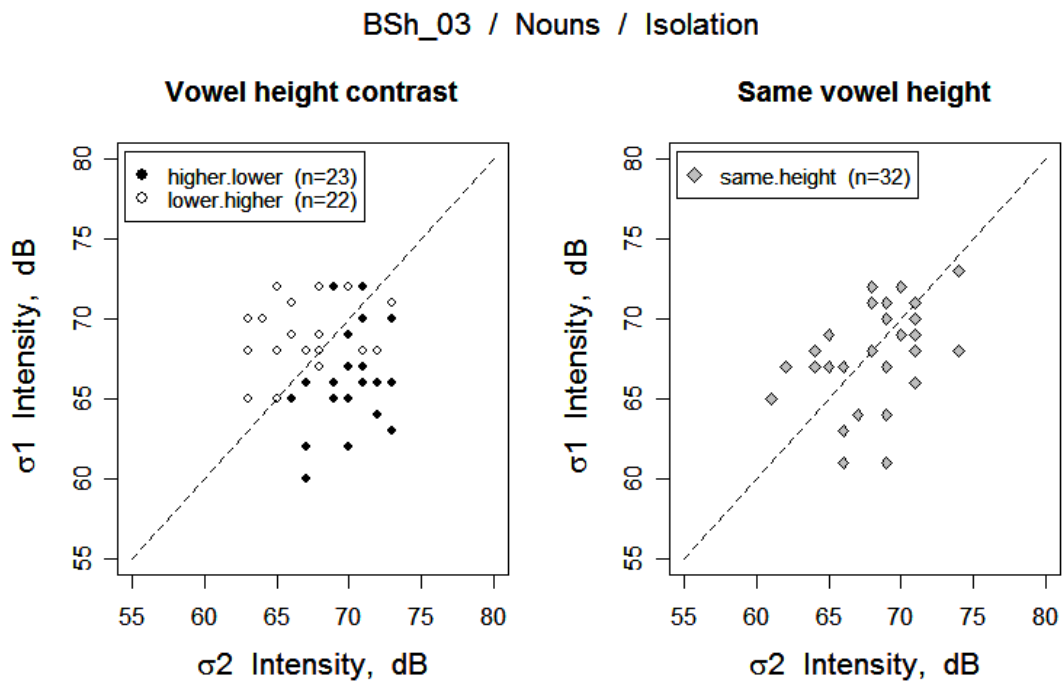


Except for a few anomalous points, this distribution is entirely attributable to the variation in intensity which is intrinsic to vowel height differences. While F0 tends to be higher on high vowels and lower on low vowels, the opposite is the case for intensity: intensity is inherently higher on low vowels and lower on high vowels (Lehiste 1970: 68, 120).

BSh_03 isolation forms are plotted in terms of vowel height contrasts in Figure 4.81 below. The graph on the left shows nouns with a contrast in vowel height across syllables. Nouns in the [lower.higher] group – such as [ndaɤ.'zu] *mda'.gzhu* ‘arrow and bow’ and [baɣ.'mo] *bag.mo* ‘bride’ – have a higher intensity on $\sigma 1$ and fall to the left of the dashed line. Nouns in the [higher.lower] group – such as

[mɛ.'ðɑ]⁴¹ *me.mda* ‘gun, rifle’ and [tut.'pa] *dud.ba* ‘smoke’ – have a higher intensity on σ_2 , and fall to the right of the dashed line. That is, intensity is higher on whichever syllable has the lower vowel, as one would predict following Lehiste.

Figure 4.81 BSh_03 / Nouns / Isolation: Intensity vs. vowel height contrast

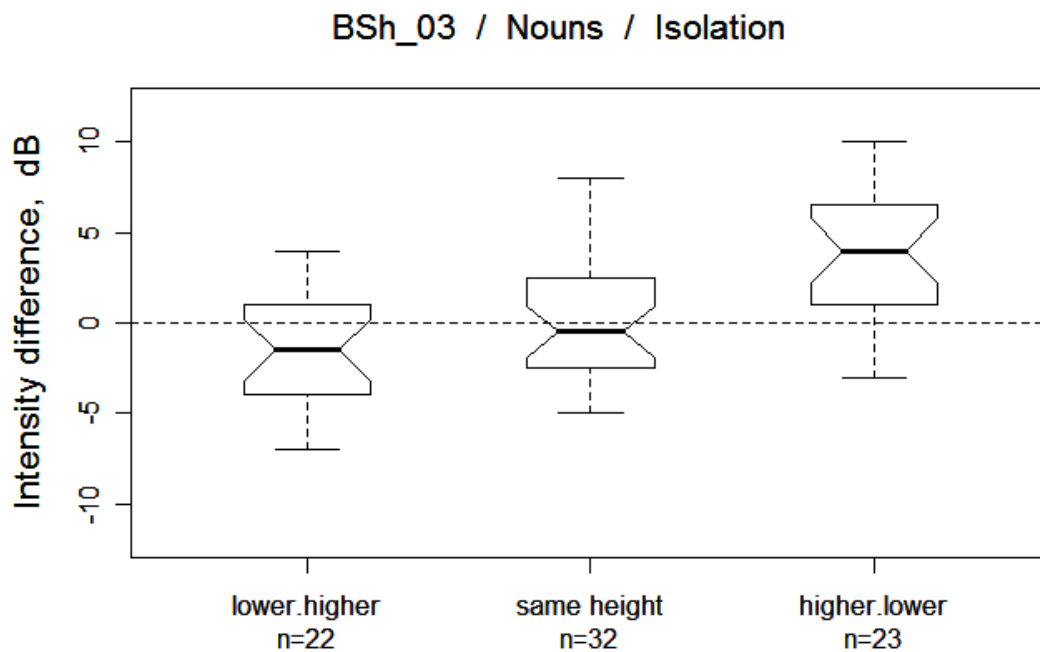


The graph on the right above shows that nouns with vowels of the same height in both syllables – such as [spal.'ba] *dpral.ba* ‘forehead’ and [bu.'striŋ] *bu.sriŋ* ‘woman’ – are about evenly distributed to either side of the dashed line. This is a random distribution; there is no evidence that the speaker is manipulating intensity to convey stress.

⁴¹ Bielmeier (p.c. 2008) points out that in Balti the word for gun is more commonly [tuaq].

As illustrated in the box-and-whisker plot in Figure 4.82 below, the differences in intensity across syllables for the three height groups define a steady trend, corresponding to vowel height contrasts. For the control group – those with vowels of the same height in both syllables – the notch representing the 95% confidence interval about the median difference spans the x-axis. This means that the median difference in pitch across syllables is not significantly different than zero.

Figure 4.82 BSh_03 / Nouns / Isolation: Intensity difference vs. vowel height contrast



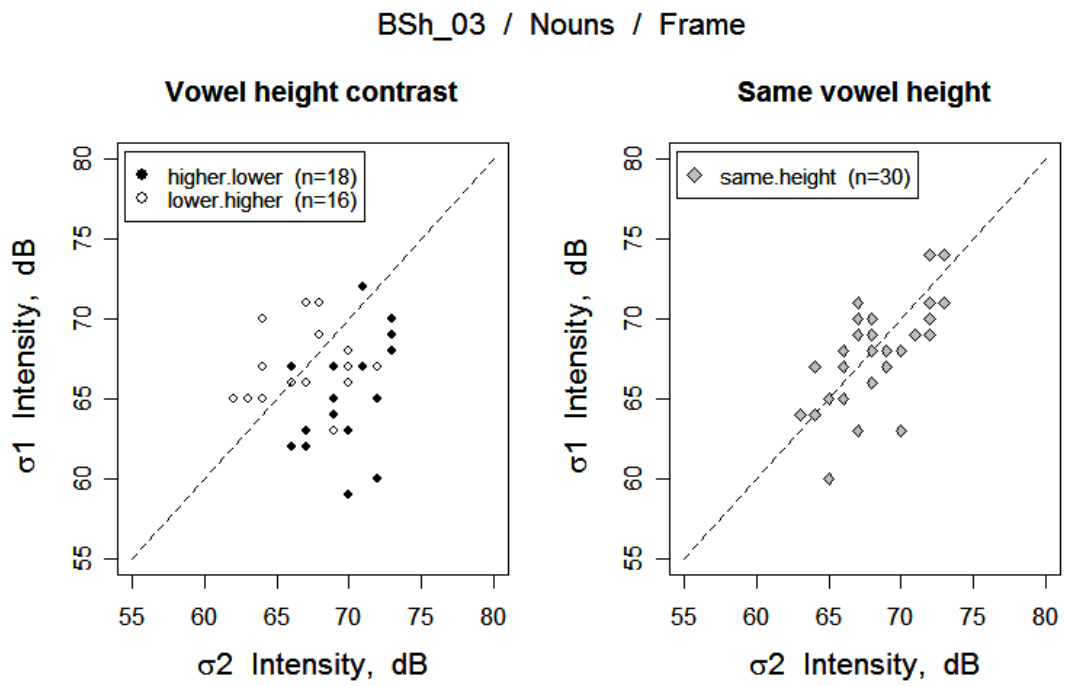
This is confirmed by the paired-sample t-tests summarized in Table 4.18 below. For the control group, $p > 0.05$ and the 95% confidence interval includes zero, indicating that there is no significant difference in intensity across syllables.

Table 4.18 BSh_03 / Nouns / Isolation: Intensity by height: paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	-2.5386	21	0.01911	-1.7	-3.1	-0.30
same height	0.2088	31	0.836	0.12	-1.1	1.3
higher.lower	5.7811	22	8.14e-06	3.9	2.5	5.3

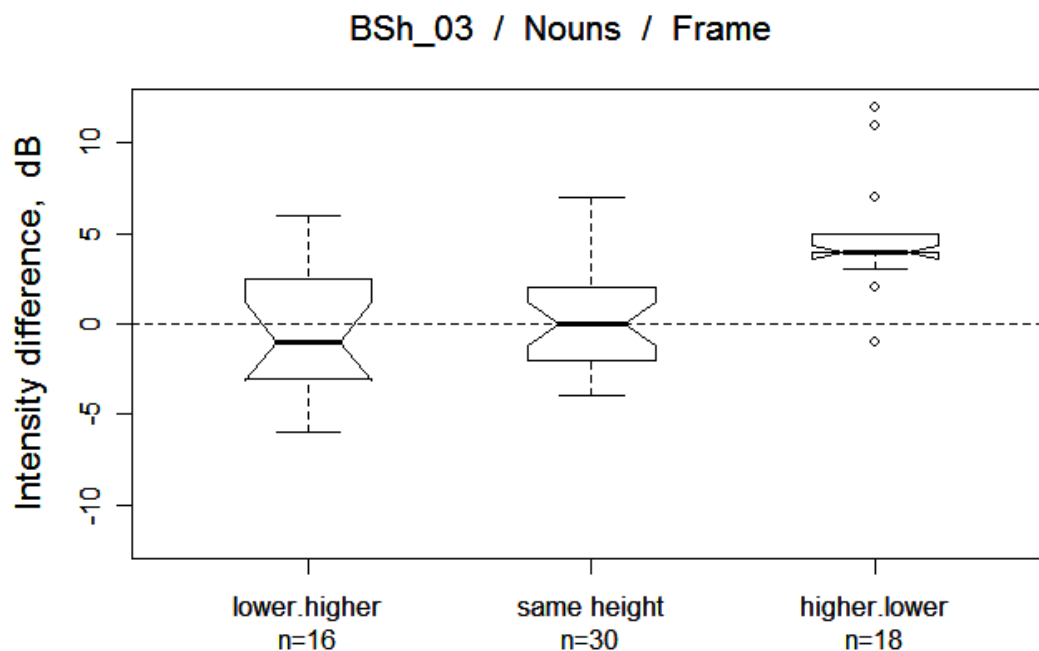
Intensity values for nouns produced in the sentence frame “In the Balti language, we *X* say” / [ˈbal.ti ˈskɑt.dɪŋ *X* zɑ.ret] are plotted in terms of vowel height in Figure 4.83 below. The pattern here is similar to that exhibited by the isolation forms. When there is a contrast in vowel height across syllables, intensity tends to be higher on whichever syllable has the lower vowel, as one would predict following Lehiste (1970). Nouns in the control group are more or less evenly distributed to either side of the dashed line.

Figure 4.83 BSh_03 / Nouns / Frame: Intensity vs. vowel height contrast



In the box-and-whisker plot in Figure 4.84 below, the median intensity difference for the control group is exactly zero.

Figure 4.84 BSh_03 / Nouns / Frame: Intensity difference vs. vowel height contrast



Paired-sample t-tests, summarized in Table 4.19 below, confirm that intensity differences across paired syllables do not define a significant contrast: $p > 0.05$ and the 95% confidence interval includes zero.

Table 4.19 BSh_03 / Nouns / Frame / Intensity by height: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	-0.2187	15	0.8298	-0.19	-2.0	1.6
same height	0.5801	29	0.5664	0.27	-0.67	1.2
higher.lower	6.0511	17	1.297e-05	4.7	3.0	6.3

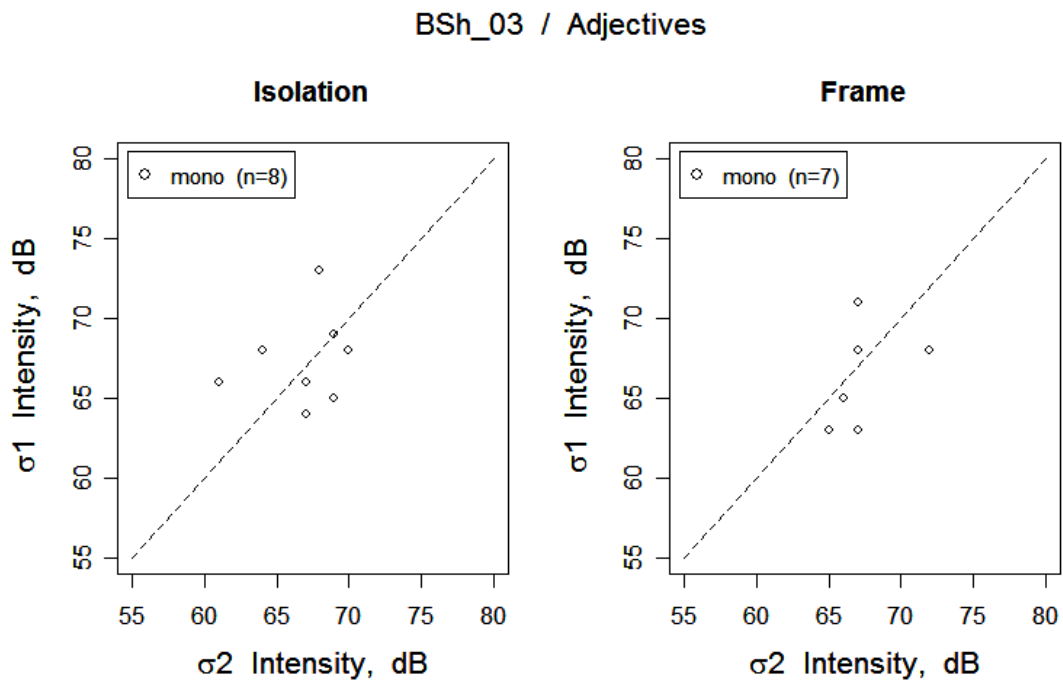
In conclusion, intensity is definitively *not* a correlate of stress for BSh_03 nouns, either in isolation or in the sentence frame.

4.4.1.2 BSh_03 Adjectives

For BSh_03 adjectives, the sample size is too small to determine whether or not there is a correlation between intensity and $\sigma 2$ stress. Furthermore, it seems that segmental properties may play an interfering role.

As shown in Figure 4.85 below, for both isolation and frame forms, intensity is sometimes higher on $\sigma 1$ and sometimes higher on $\sigma 2$, with points falling to either side of the dashed line.

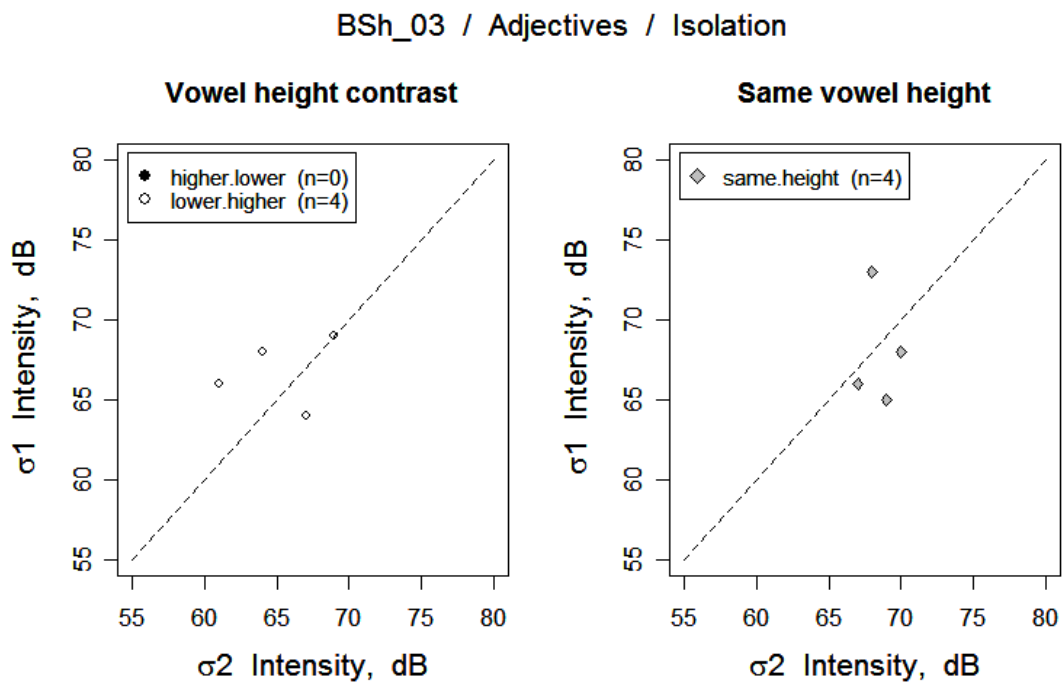
Figure 4.85 BSh_03 / Adjectives: Intensity



Isolation forms of BSh_03 adjectives are plotted in terms of vowel height contrast in Figure 4.86 below. The graphs are rather sparse and unconvincing. In the plot on the left, two of the [lower.higher] nouns – the color terms terms [kar.'p^ho]

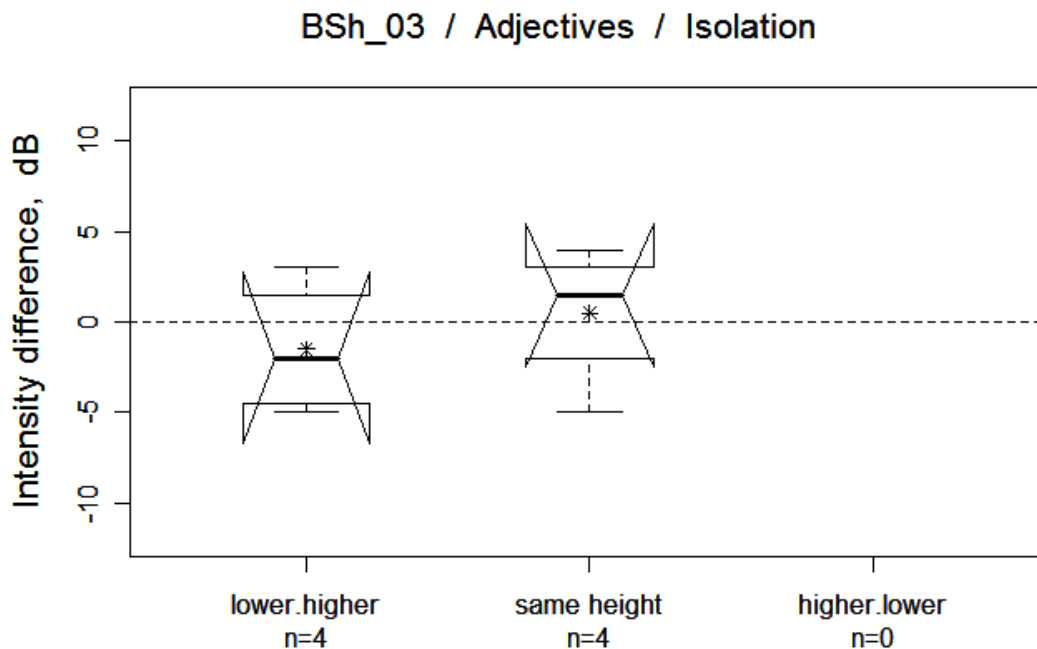
dkar.po ‘white’ and [mar.'p^ho] *dmar.po* ‘red’ – have a higher intensity on σ_1 , which one would predict following Lehiste (1970). In the plot on the right, most of the [same.height] nouns – including [bom.'bo] *sbom.po* ‘thick, fat (circumference)’ and [sar.'p^ha] *gsar.pa* ‘new’ – have a higher intensity on σ_2 , the stressed syllable.

Figure 4.86 BSh_03 / Adjectives / Isolation: Intensity vs. vowel height contrast



As illustrated in the box-and-whisker plot in Figure 4.87 below, for both subsets the notch representing the 95% confidence interval about the median difference is quite broad, which is not surprising given the small sample size. In both cases the notch spans the x-axis, indicating that the median difference in intensity across syllables is not significantly different than zero.

Figure 4.87 BSh_03 / Adjectives / Isolation: Intensity difference vs. vowel height contrast



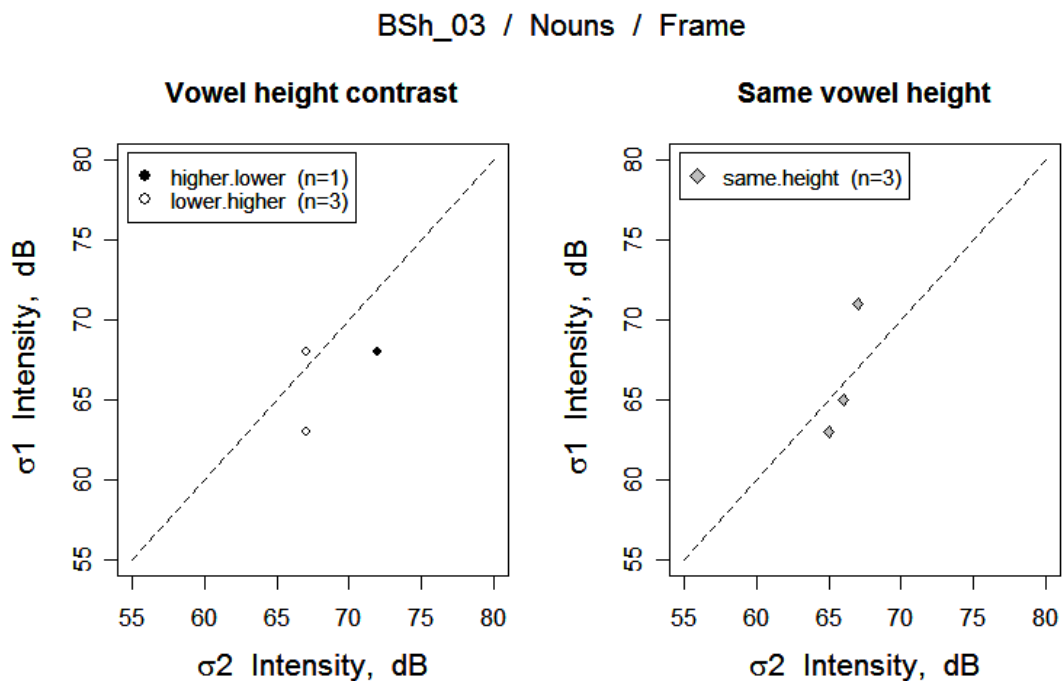
Paired-sample t-tests, summarized in Table 4.20 below, confirm that the difference in intensity across paired syllables is not statistically significant: $p > 0.05$, and the 95% confidence intervals about the mean include zero.

Table 4.20 BSh_03 / Adjectives / Isolation / Intensity by height: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	-0.8115	3	0.4765	-1.5	-7.4	4.4
same height	0.2582	3	0.813	0.5	-5.7	6.6
higher.lower	NA	NA	NA	NA	NA	NA

Intensity measurements for the frame forms are plotted in terms of vowel height contrast in Figure 4.88 below. (Note that [naq.'po] *nag.po* ‘black’ and [maĩ.'mo] *mang.mo* ‘many’ overlap at (σ_2 , σ_1) coordinates (67, 63), so only two [lower.higher] forms are visible.)

Figure 4.88 BSh_03 / Adjectives / Frame: Intensity vs. vowel height contrast



Here, it is quite clear that the sample size is too small to be really informative. As summarized in Table 4.21 below, paired-sample t-tests indicate that the contrast in intensity across syllables is not statistically significant, with $p > 0.05$ and 95% confidence intervals that include zero. (Since there was only one [higher.lower] form – [tʰr̩s.'pa] *mkhregs.pa* ‘difficult’ – the t-test could not be performed for this group.)

Table 4.21 BSh_03 / Adjectives / Frame / Intensity by height: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.4	2	0.2965	2.3	-4.8	9.5
same height	-0.1796	2	0.874	-0.33	-8.3	7.6
higher.lower	NA	0	NA	NA	NA	NA

In summary, there is no evidence that intensity is a correlate of σ_2 stress for either isolation or frame forms of BSh_03 adjectives. This situation is not quite the same as for nouns (discussed above) – where the sample was sufficiently large to demonstrate that intensity is definitively *not* a correlate of stress.

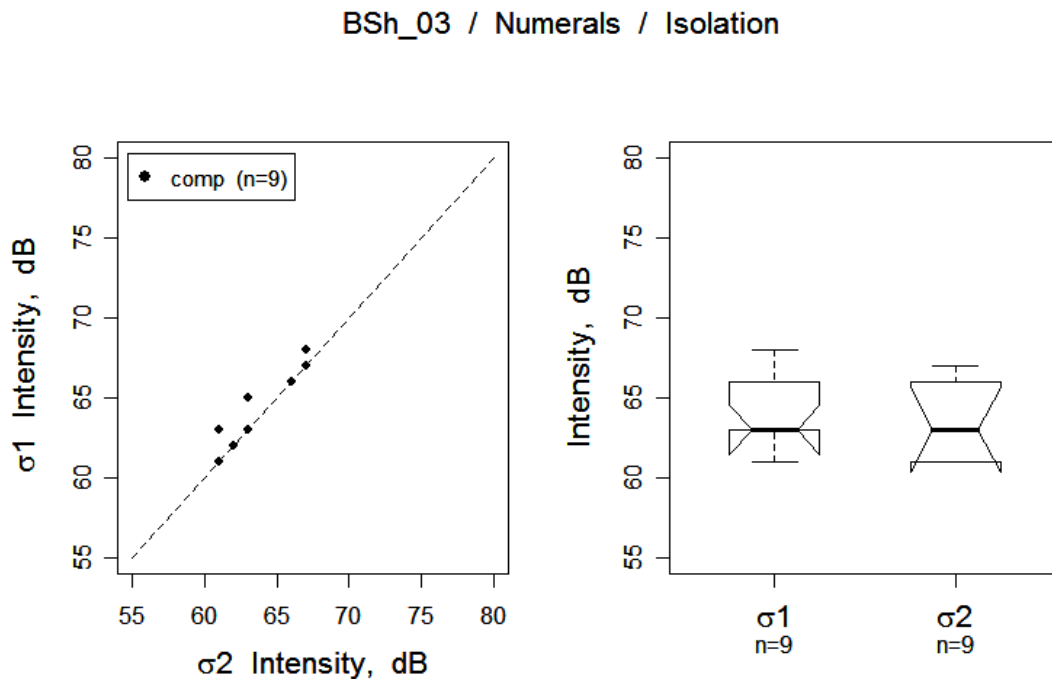
As a final note, it is interesting to observe that the three isolation forms and the two frame forms which have a higher intensity on σ_1 also happen to be the only adjectives in which σ_2 begins with an aspirated onset: [kar.'p^ho] *dkar.po* ‘white’, [sar.'pha] *gsar.pa* ‘new, and [mar.'p^ho] *dmar.po* ‘red’. As noted by Gordon and Ladefoged (2001), aspiration on a syllable onset may be associated with non-modal phonation in the succeeding vowel, which may in turn result in a reduced intensity. If intensity on σ_2 is thus depressed, intensity on σ_1 will then be relatively higher. This may play a confounding role in the distribution here, but a larger sample would be required to determine its significance.

4.4.1.3 BSh_03 Numerals

Intensity values for the nine numerals are plotted in Figure 4.89 below. As noted previously, the numerals show little acoustic variation, since they are part of a distinct set of related words and were recorded in sequence. Indeed, here only eight points are visible, because two of them overlap at (σ_2 , σ_1) coordinates (61, 63).

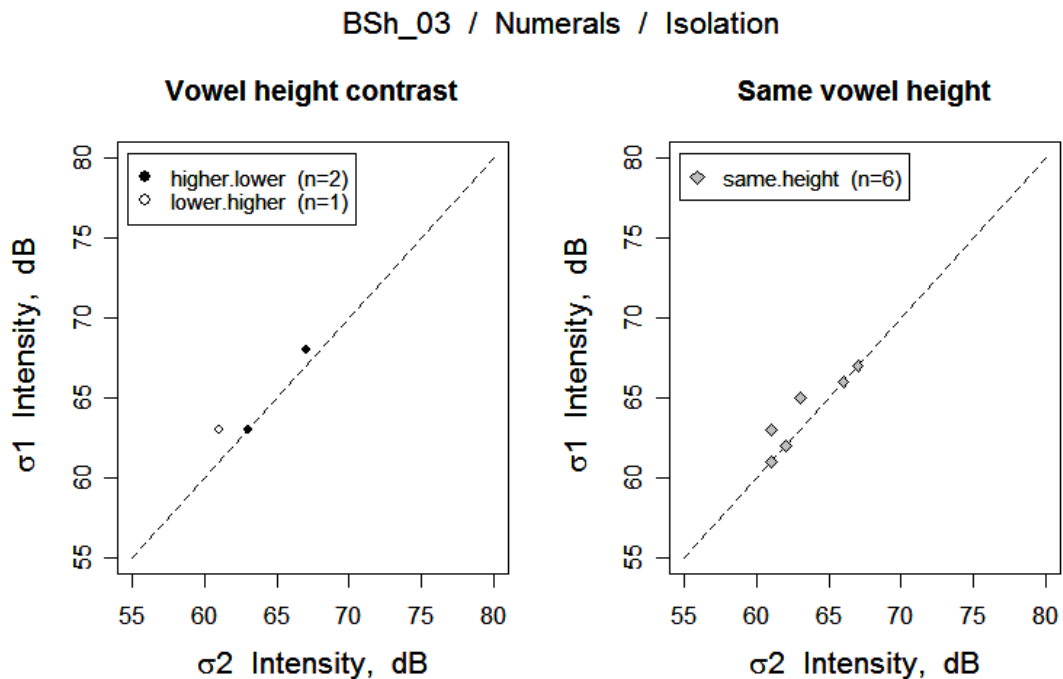
It is quite clear that intensity is not a correlate of σ_2 stress: five points fall exactly on the dashed line, and the remaining four have a higher intensity on σ_1 , the syllable which is *not* stressed. The largest contrast in intensity across syllables is -2 dB. In the box-and-whisker plot to the right below, σ_1 and σ_2 show almost complete overlap in range, and they have exactly the same median value, 63 dB.

Figure 4.89 BSh_03 / Numerals: Intensity



If all – or most of – the numerals had a [lower.higher] contrast in vowel height across syllables, then we could attribute this distribution to the intrinsic variation of intensity. But that is not the case: as shown in Figure 4.90 below, there is only one numeral – [ʔɤaf.tɕu] *lŋa.bcu* ‘fifty’ – with the [lower.higher] contrast in vowel height which is predicted to have – and does have – a higher intensity on σ_1 . The [higher.lower] and [same.height] numerals – such as [tɕo.'ŋas] *bcu.gnyis* ‘twelve’ and [tɕur.'gu] *bcu.dgu* ‘nineteen’ – are *not* predicted to have an inherently higher intensity on σ_1 , but they, too, plot on or slightly to the left of the dashed line.

Figure 4.90 BSh_03 / Numerals: Intensity vs. vowel height contrast



A paired-sample t-test shows that the contrast in intensity across syllables is marginally significant, with a p-value of 0.04. The mean difference across syllables

is -0.78 dB (i.e., higher on σ_1). The upper 95% confidence limit of -0.03 dB means that an average decrease in intensity across syllables which is considered *statistically* significant might be as small as only $^3/_{100}$ of a decibel, which is not even perceptible. Thus while intensity shows no increase from σ_1 to σ_2 , the decrease across syllables is not meaningful either, and it is clear that intensity shows no correlation with stress for this sample of BSh_03 numerals.

Table 4.22 BSh_03 / Numerals / Intensity: results of paired-sample t-tests (two-tailed)

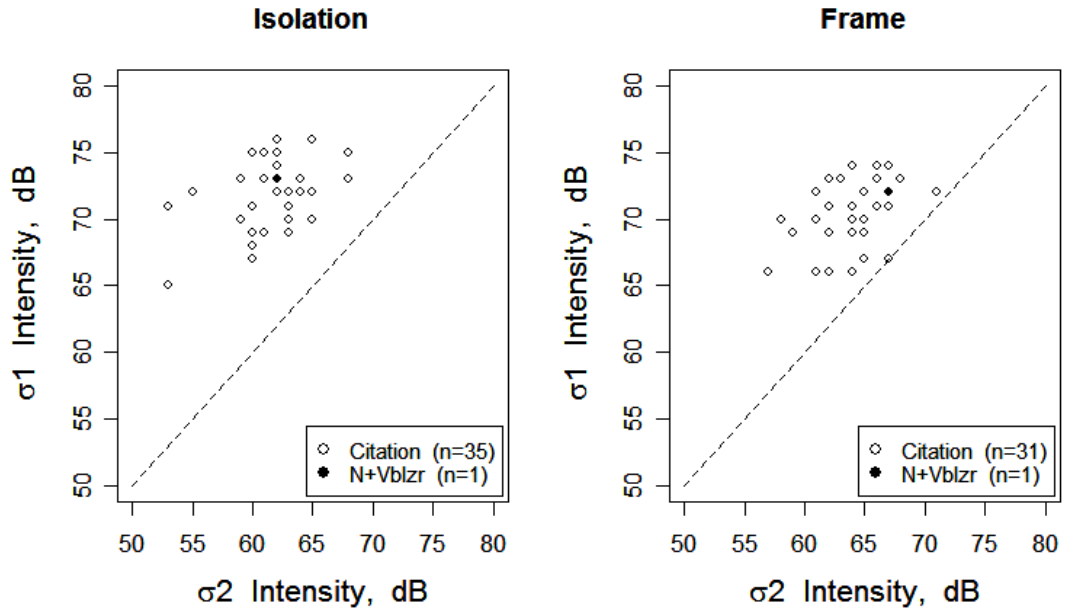
Speaker	Setting	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
						lower	upper
BSh_03	isolation	-2.401	8	0.04311	-0.78	-1.5	-0.03

4.4.1.4 BSh_03 Verbs

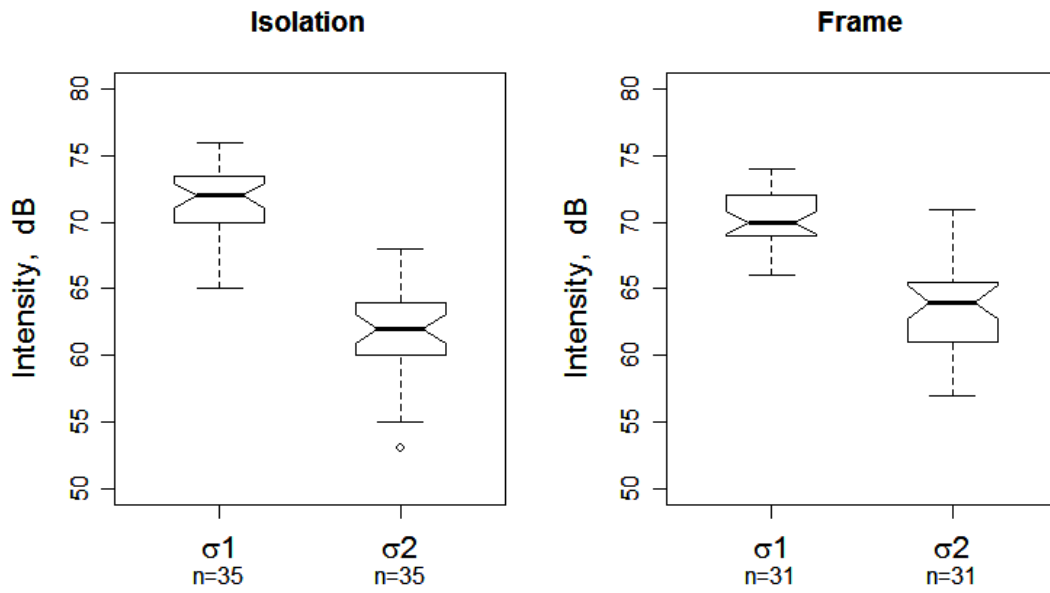
Intensity values for verbs produced by speaker BSh_03 are plotted in Figure 4.91 below. Except for the frame form ['smin.ma] *smin.ma* ‘to be ripe’ which falls exactly on the dashed line, all other verbs have a higher intensity on σ_1 , the stressed syllable. The contrast is more pronounced for the isolation forms than for the frame forms. The lone N+Vblzr form ['za:n.za] *zan.za* ‘to eat food’ plots among the citation forms in terms of intensity, but I have excluded it from further consideration here, as I have done when considering the other acoustic features. In the box-and-whisker plots in the bottom part of the figure below, there is no overlap of either the notches – representing the 95% confidence interval about the median – or of the boxes – representing the interquartile range.

Figure 4.91 BSh_03 / Verbs: Intensity

BSh_03 / Verbs



BSh_03 / Verbs / Citation



Paired-sample t-tests, summarized below, confirm that intensity differences across syllables are significant, with p-values $\ll 0.05$. Even the smallest probable contrast of -5.5 dB – represented by the lower 95% confidence limit for the frame forms – would almost certainly be highly perceptible.

Table 4.23 BSh_03 / Verbs / Citation / Intensity: results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
						lower	upper
BSh_03	isolation	-17.486	34	$< 2.2e-16$	-10	-11	-8.8
*	frame	-11.7417	30	$9.593e-13$	-6.7	-7.8	-5.5

* The lone BSh_03 N+Verbalizer form was excluded.

It is thus clear that intensity is a robust correlate of the σ_1 stress perceived on citation forms of verbs produced by speaker BSh_03, both in isolation and in the sentence frame. While we know that intensity varies as a function of vowel height – as discussed above during consideration of nouns – it is clear that, for BSh_03 verbs, this variation in no case yields a higher value on σ_2 . That is, the correlation between intensity and stress is exceptionless, and any intrinsic variation is moot.

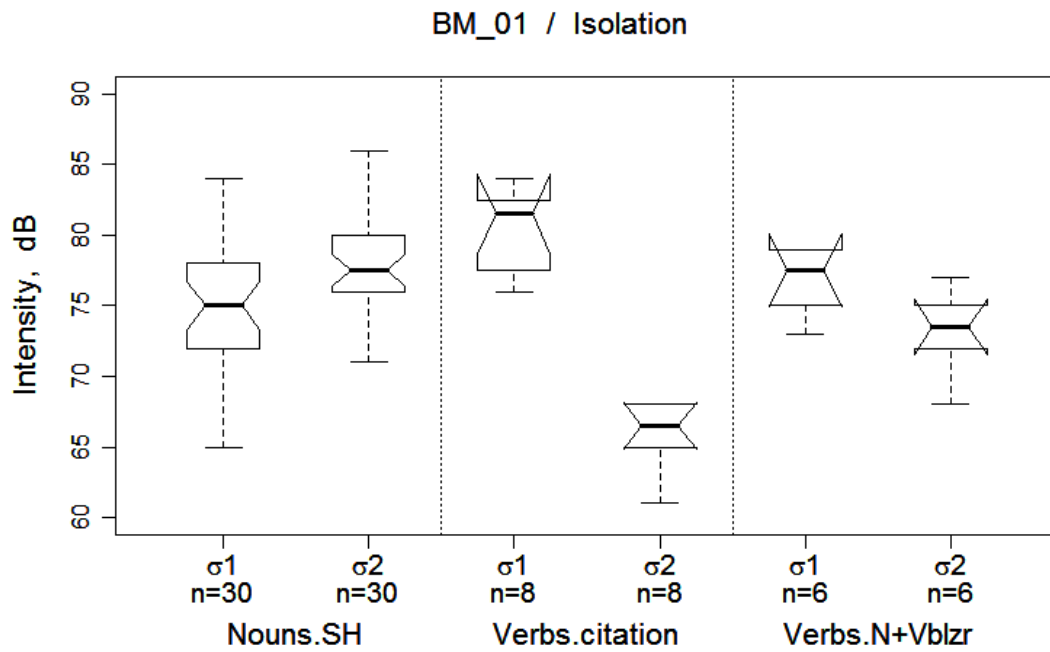
4.4.2 Intensity for speaker BM_01

The box-and-whisker plots in Figure 4.93 below provide a rough view of the contrast in intensity across syllables for BM_01 nouns and verbs. Adjectives are considered separately in section 4.4.2.2.

Nouns tend to have a slightly higher intensity on σ_2 , the stressed syllable. (Note that only nouns with vowels of the same height are plotted here.) For verbs,

regardless of morphological type, there is a much stronger correlation between intensity and σ_1 stress. There is a clear distinction between the patterns on non-verbs vs. verbs.

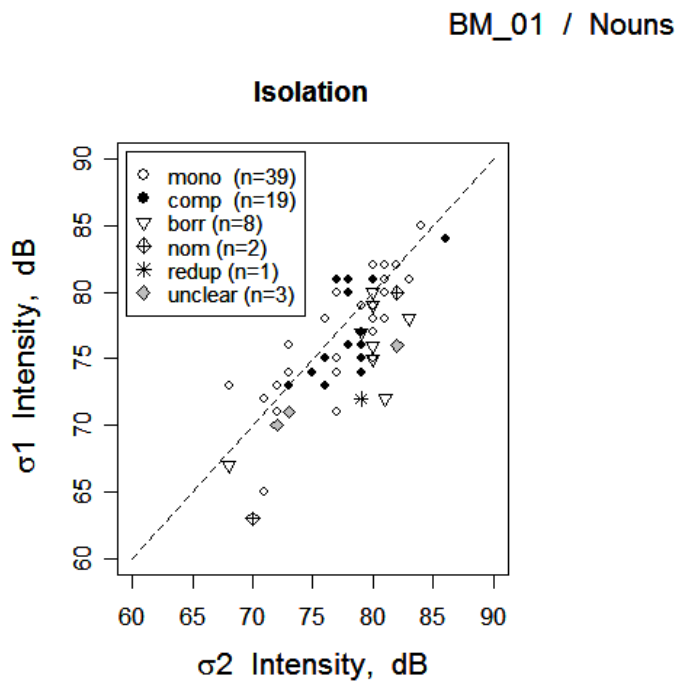
Figure 4.92 BM_01 / Isolation: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)



4.4.2.1 BM_01 Nouns

For BM_01 nouns, intensity is sometimes higher on σ_1 and sometimes higher on σ_2 ; points fall to either side of the dashed line in Figure 4.93 below. This distribution is not quite even, though: two thirds of the points plot to the right.

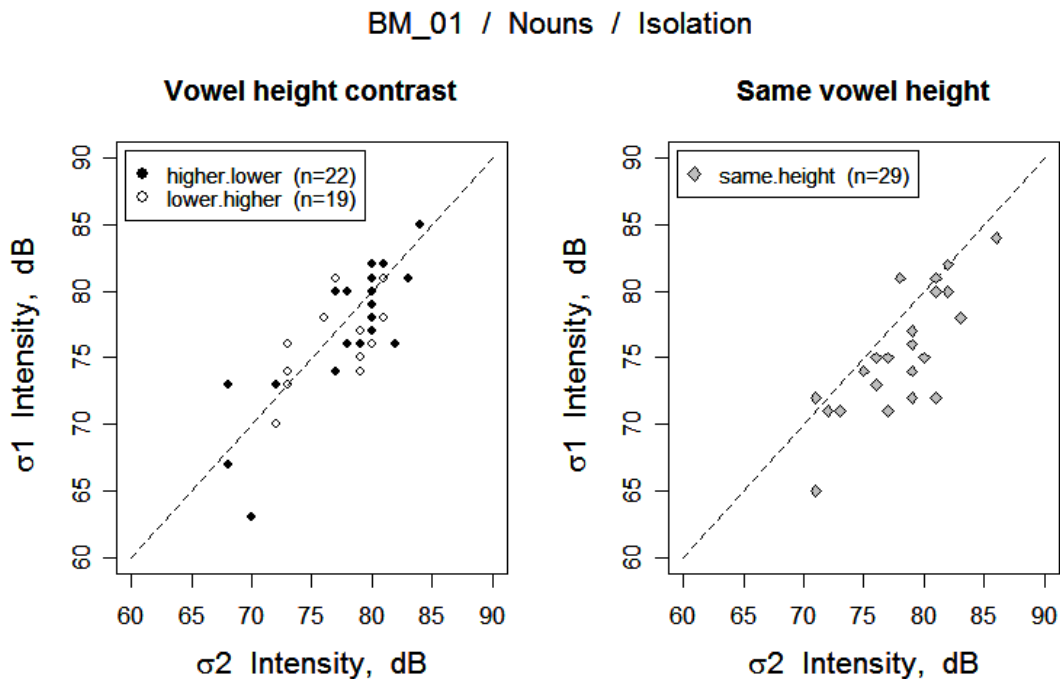
Figure 4.93 BM_01 / Nouns: Intensity



In Figure 4.94 below, intensity values for these same nouns (excluding the two with diphthongs) are plotted again, but now coded in terms of vowel height. The graph on the left shows nouns with a contrast in vowel height across syllables. These do not define the tidy distinctions that we observed in BSh_03 nouns (Figure 4.81 and Figure 4.83). If intensity were governed by contrasts in vowel height, we would expect to see nouns in the [lower.higher] group – such as [naŋ.'mi] *nang.mi* ‘family members’ and [ʃla.'ŋo] *lha.mo* ‘goddess’ – falling consistently to the left of the dashed line. Likewise, we would expect to see nouns in the [higher.lower] subset – such as [ʃi.'la] *byi.la* ‘cat’ and [broq.'pa] *'brog pa* ‘nomad, pastoralist’ – falling consistently to the right of the dashed line. Instead, the distribution seems random,

perhaps reflecting the inconsistency of production by this speaker. There is no evidence that vowel height is a controlling factor.

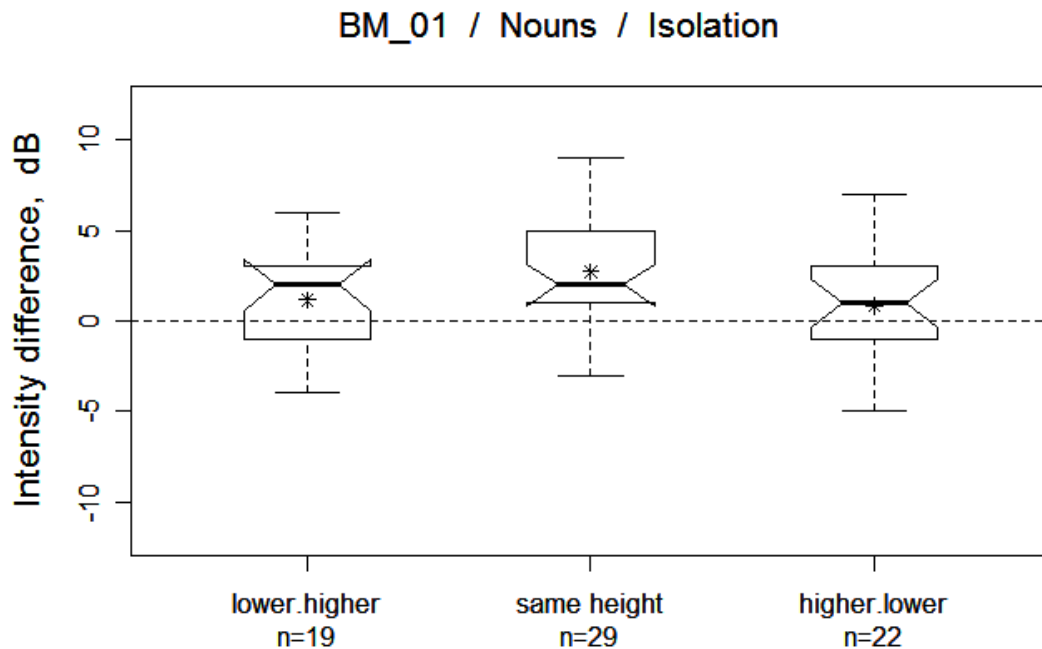
Figure 4.94 BM_01 / Nouns: Intensity vs. vowel height contrast



The graph on the right above shows nouns in which the vowel is the same height in both syllables – such as [la.'ma] *bla ma* ‘lama’ or [pɸi.'lim] ‘film, movie’. Almost all of the points (several of which overlap) fall to the right of the dashed line, indicating a higher intensity on σ_2 , the stressed syllable. This control group thus suggests that intensity is, indeed, a correlate of stress for this speaker. All different morphological structures are included in this group: monomorphemic, compound, borrowed, and, of course, the one reduplicated form (borrowed from Burushaski) [t□u.'t□u] ‘breast’, plotting at (σ_2, σ_1) coordinates (79,72).

The differences in intensity across syllables for the three height groups are illustrated in the box-and-whisker plot in Figure 4.95 below.

Figure 4.95 BM_01 / Nouns / Isolation: Intensity difference vs. vowel height contrast



Of greatest interest, as usual, are paired-sample t-tests. Results for the three vowel height groups are summarized in Table 4.24 below, along with test results for the full set of nouns.

Table 4.24 BM_01 / Nouns / Isolation / Intensity by height: paired t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
all	5.1692	71	2.074e-06	1.7	1.0	2.3
lower.higher	1.7936	18	0.08969	1.2	-0.20	2.5
same height	5.8217	28	2.96e-06	2.7	1.8	3.7
higher.lower	1.3438	21	0.1934	0.82	-0.45	2.1

When the full set of nouns is considered, the difference in intensity is statistically highly significant ($p \ll 0.05$), but the smallest probable mean contrast of 1.0 dB (the lower 95% confidence limit) is just at the difference limen identified by Lehiste (1970). For the subsets with a contrast in vowel height, the p-value is *not* < 0.05 , so the null hypothesis – that there is no significant difference across syllables – cannot be rejected. It is thus only for the control group that the increase in intensity can be considered both statistically significant ($p \ll 0.05$), and perceptible (lower 95% confidence limit = 1.8 dB).

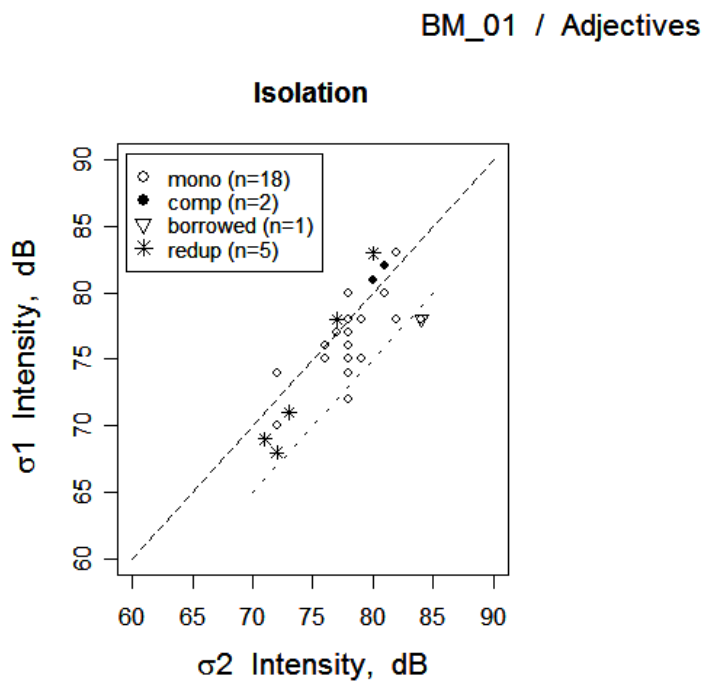
In conclusion, for BM_01 nouns, intensity shows a correlation with the σ_2 stress pattern only under limited, controlled conditions. It cannot be considered a robust and consistent cue for stress.

4.4.2.2 BM_01 Adjectives

Intensity measurements for adjectives recorded from speaker BM_01 are plotted in terms of morphological structure in Figure 4.96 below. Most points fall to the right of the dashed line, with a higher intensity on σ_2 . (The increase is generally < 5 dB, as indicated by the dotted reference line.) Note that the five reduplicated

forms fall to either side of the dashed line. This suggests that, when segmental content is completely controlled, intensity may randomly be either slightly higher on $\sigma 1$ or slightly higher on $\sigma 2$. If intensity were a correlate of stress, we would expect it to be consistently higher on $\sigma 2$ in such control groups.

Figure 4.96 BM_01 / Adjectives: Intensity

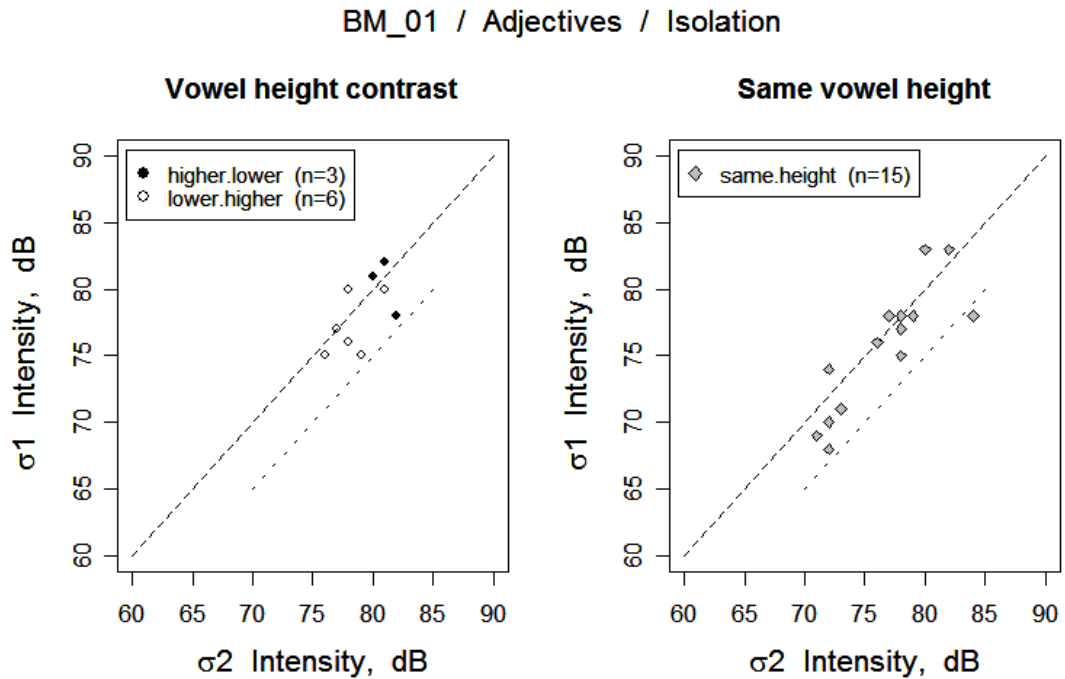


BM_01 adjectives are plotted in terms of vowel height in Figure 4.97 below.

Adjectives with a contrast in vowel height include [tɕ^ha.'ru] ‘dirty’ (borrowed from Burushaski) and [ʔzum.'bo] *zhim.po* ‘tasty, delicious’. For the control group – the set of adjectives with vowels of the same height in both syllables, such as [tɕsax.'ma]

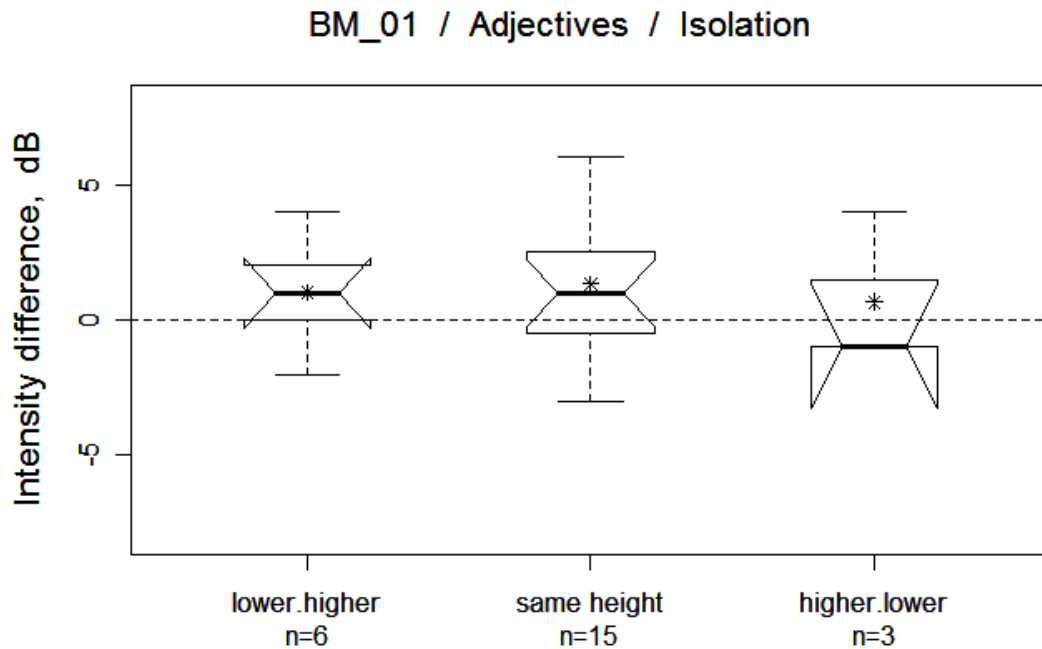
gtsang.ma ‘clean’ and [ʈuk.ʈuk] *mthug.mthug* ‘thick’ – points fall close to and to either side of the dashed line.

Figure 4.97 BM_01 / Adjectives / Isolation: Intensity vs. vowel height contrast



Boxplots illustrating the intensity difference for each of the three height groups are shown in Figure 4.98 below. In each case, the notch representing the 95% confidence interval about the median touches or spans the dashed line. This indicates that the median difference in intensity across syllables is not significantly different than zero.

Figure 4.98 BM_01 / Adjectives / Isolation: Intensity difference vs. vowel height contrast



Indeed, this is confirmed by the paired-sample t-tests summarized in Table 4.25. In all three cases, the p-value is *not* < 0.05 , so the null hypothesis – that intensity is the same on both syllables – cannot be rejected. The 95% confidence intervals span zero, meaning that, if groups of adjectives were repeatedly measured and analyzed, there is a possibility each time that there might be no significant contrast in intensity across syllables.

Table 4.25 BM_01 / Adjectives / Intensity by height: paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.2247	5	0.2752	1.0	-1.1	3.1
same height	1.9387	14	0.07297	1.3	-0.14	2.8
higher.lower	0.4	2	0.7278	0.67	-6.5	7.8

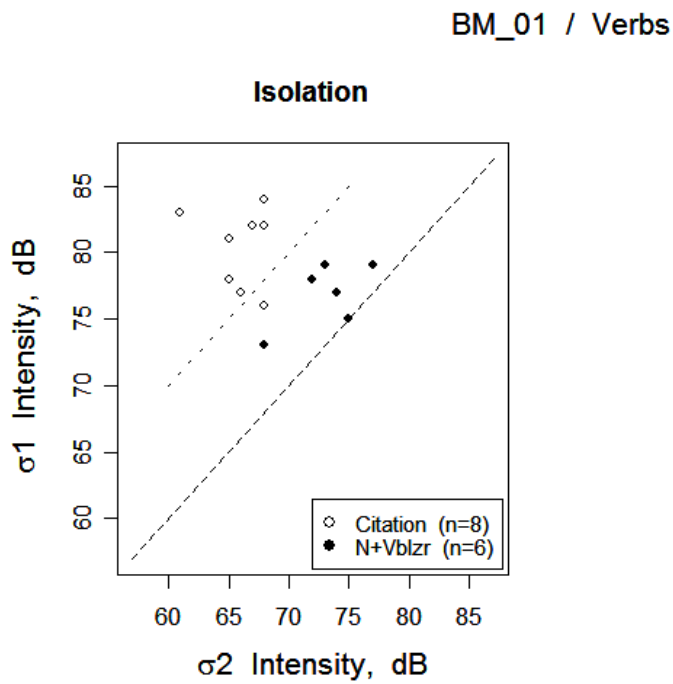
In conclusion, even though intensity is often higher on σ_2 in BM_01 adjectives, the contrast is not consistent or robust enough to be identified as a correlate of stress.

4.4.2.3 BM_01 Verbs

We have already seen that intensity corresponds with σ_1 stress for speaker BM_01 (Figure 4.18, page 181), and that N+Vblzr and citation forms exhibit different patterns (Table 4.8, page 182). The data is considered in greater detail here.

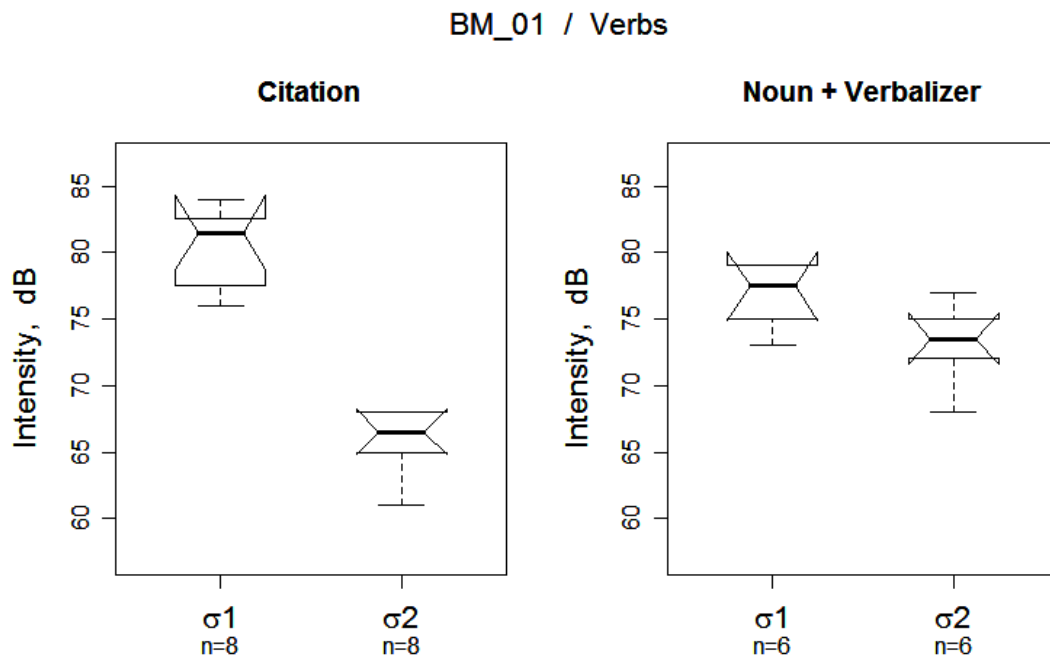
Measurements for both morphological types are plotted in Figure 4.99 below. Here, the N+Vblzr form [ˈzan.ɸtɕø̃s] *zan.bcos* ‘make food’ falls exactly on the dashed line, but all other verbs have a higher intensity on σ_1 , the stressed syllable. For the citation forms, with only one exception the intensity difference is > 10 dB.

Figure 4.99 BM_01 / Verbs: Intensity



The contrasts across syllables for both BM_01 verb types are illustrated by the box-and-whisker plots in Figure 4.100 below. As with pitch (Figure 4.19), the citation forms show a dramatic decrease from σ_1 to σ_2 , while for the N+Vblzr forms the contrast is less pronounced.

Figure 4.100 BM_01 / Verbs: Intensity contrasts vs. morphological category



Paired-sample t-tests for the two morphological types are summarized in Table 4.26 below. For the citation forms, $p \ll 0.05$ and, on average, intensity is 14 dB higher on $\sigma 1$ than on $\sigma 2$. Even the lower 95% confidence limit of -11 dB is quite dramatic. The contrast in intensity across syllables is also statistically significant for the N+Vblzr forms, with $p < 0.05$ and a mean difference across syllables of -3.7 dB. For this group, though, there is a probability that repeated sampling could yield a mean intensity contrast as small as -1.1 dB, which is barely more than the just-noticeable-difference limen of 1 dB (Lehiste, 1970)

Table 4.26 Intensity on Balti verbs: results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
						lower	upper
BM_01	citation	-9.9081	7	2.274e-05	-14	-18	-11
	N+Vblzr	-3.7081	5	0.01388	-3.7	-6.2	-1.1

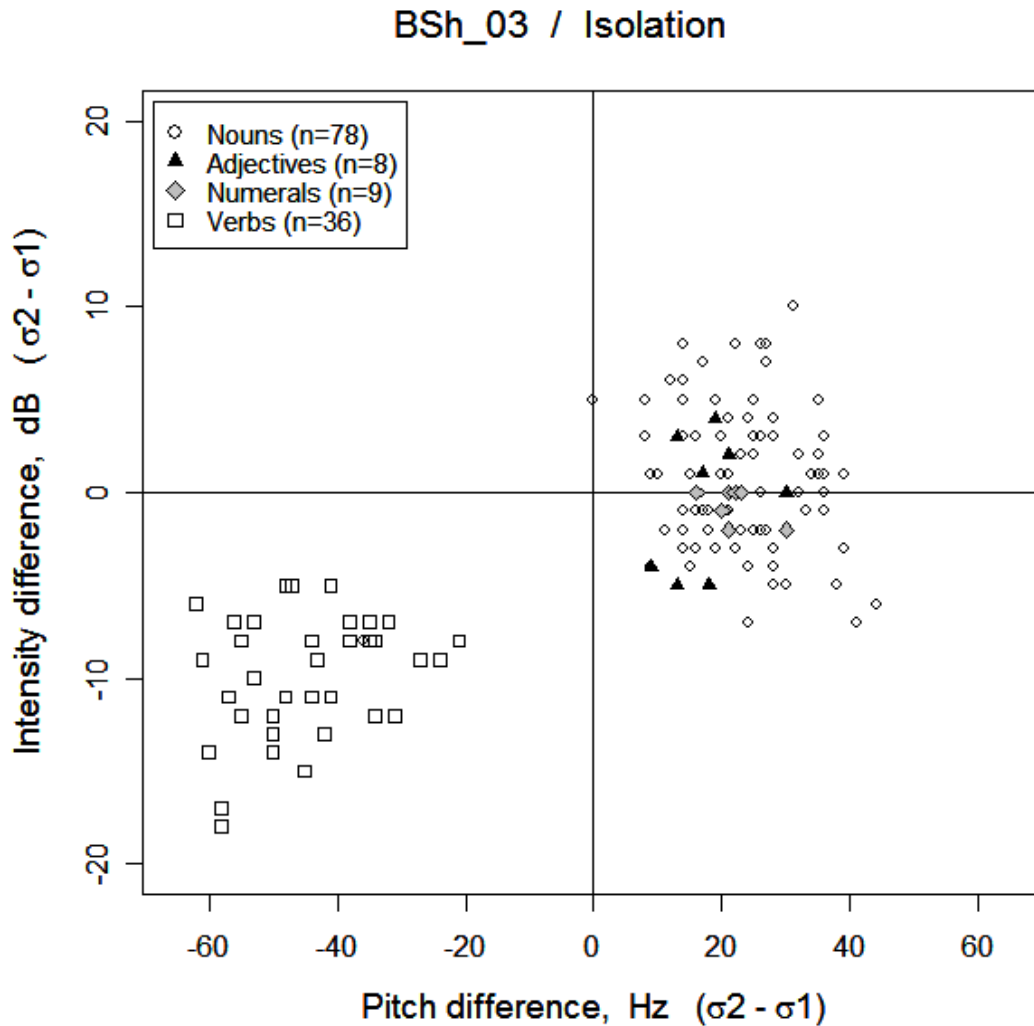
In summary, intensity is a robust cue for σ_1 stress for the BM_01 citation forms. For the N+Vblzr forms, intensity shows a consistent but sometimes weak correlation with stress.

4.4.3 The interaction of pitch and intensity

When pitch and intensity are considered together, the different acoustic signatures exhibited by verbs *vs.* non-verbs is quite striking.

For speaker BSh_03, isolation forms are plotted in terms of pitch difference and intensity difference in Figure 4.101 above. For verbs, both pitch and intensity are markedly higher in σ_1 , the stressed syllable. For non-verbs, pitch is clearly higher on σ_2 , while intensity is sometimes higher on one syllable, and sometimes higher on the other. The numerals fall in a tight cluster. Since they comprise a cohesive semantic category and were recorded consecutively within the span of a few minutes, it is not surprising that they show little variation in pitch or intensity.

Figure 4.101 BSh_03 / Isolation: Pitch difference vs. intensity difference

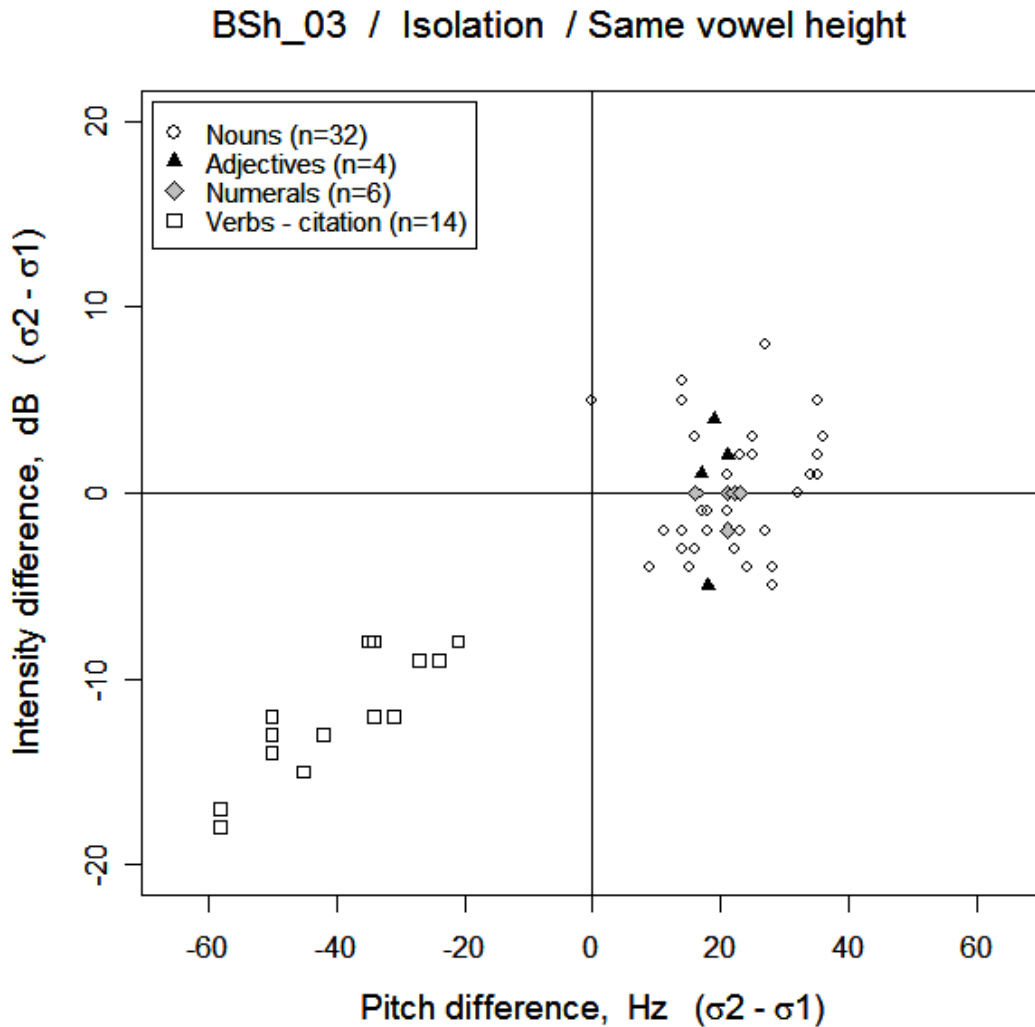


Plotting pitch and intensity together was also helpful in identifying anomalies. In the plot above, there is one noun which falls among the verbs at coordinates (-36,-8). This point – almost completely obscured by verbs – corresponds to the word ‘inner shirt’ (some kind of undergarment, as described by my consultant), and is pronounced [ʰŋɔn.mo] with a strong σ_1 stress. I have no idea why this one noun has

the stress pattern typical of verbs; it was excluded from consideration in all other elements of this analysis.

As demonstrated in section 4.4.1, intensity is highly sensitive to contrasts in vowel height. When the sample is controlled for vowel height – as in Figure 4.102 below – the distribution of the non-verbs about the x-axis confirms that intensity is not manipulated by the speaker in order to convey stress. The verbs define a linear trend, indicating that pitch and intensity vary together for this lexical category. (The lone N+Vblzr has been omitted here.)

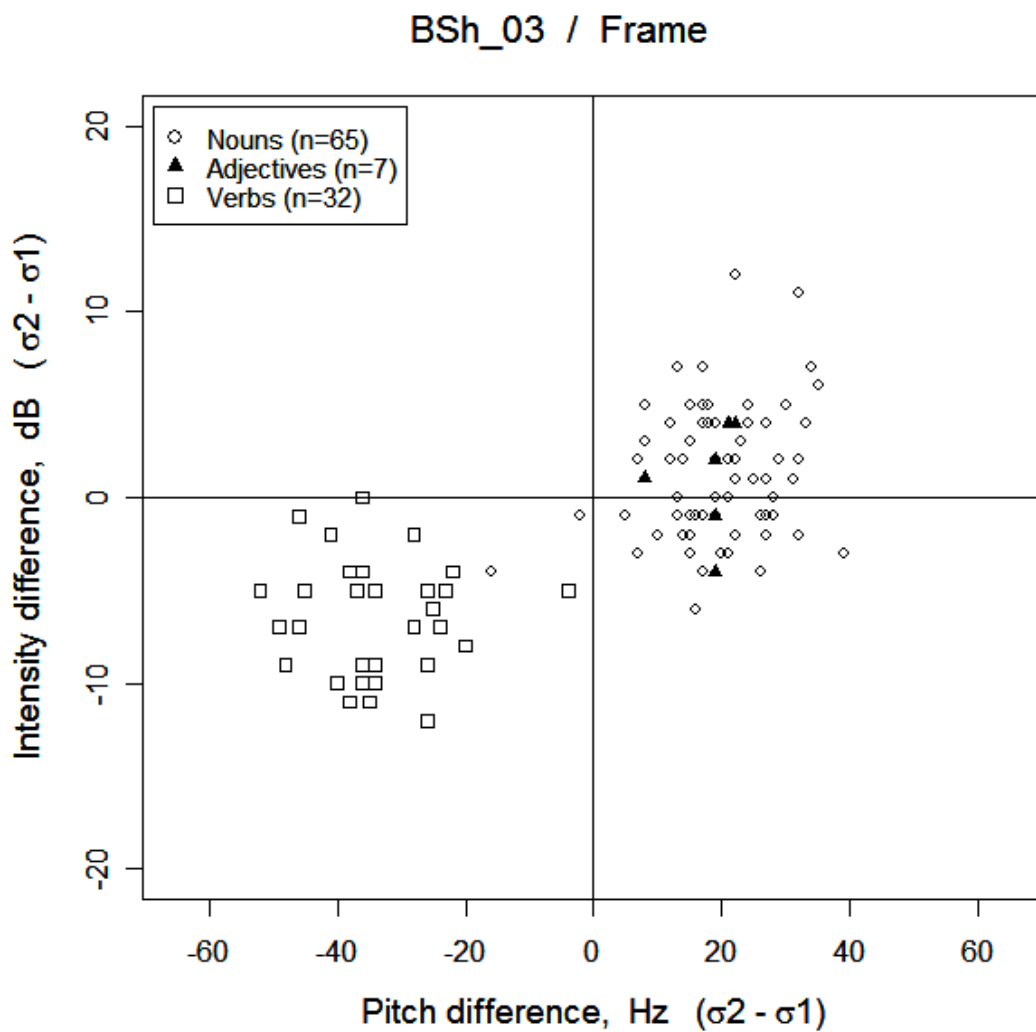
Figure 4.102 BSh_03 / Isolation / Same vowel height: Pitch difference vs. intensity difference



BSh_03 frame forms are plotted in terms of pitch and intensity differences in Figure 4.103 below. Compared to Figure 4.101, both clusters are closer to the graph's origin, reflecting the constraints imposed in the sentence frame. The clusters are also shifted slightly upward, indicating that the intensity on σ_2 is a bit higher relative to the intensity on σ_1 . This is likely a consequence of the word's position at the focal

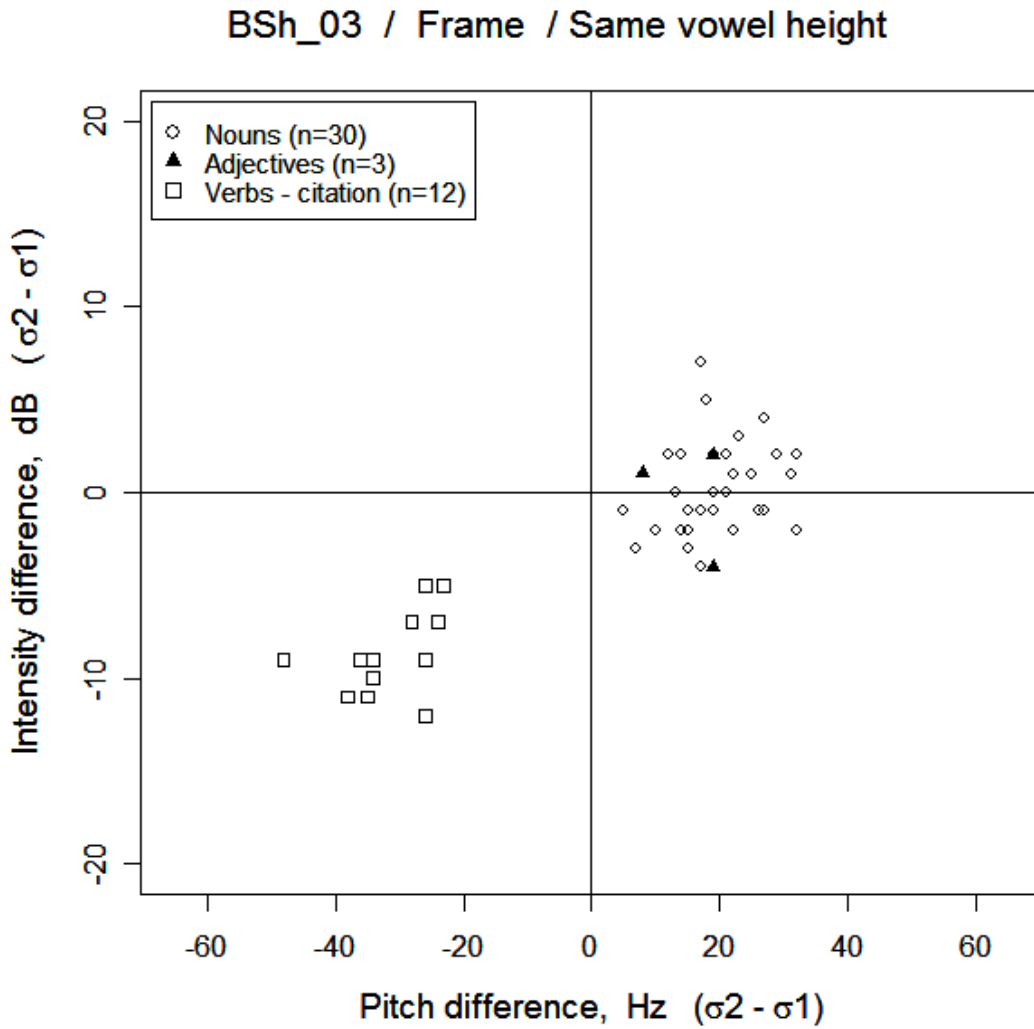
penultimate position in the carrier sentence “We X say” or “In the Balti language we X say”. In the case of frame forms, too, the noun ‘inner shirt’ [ʈŋɔn.mo] plots with the verbs, at coordinates (-16, -4); again, this anomaly was excluded from the analyses above.

Figure 4.103 BSh_03 / Frame: Pitch difference vs. intensity difference



When controlled for vowel height, the patterns are sharpened, as shown in Figure 4.104 below.

Figure 4.104 BSh_03 / Frame / Same vowel height: Pitch difference vs. intensity difference

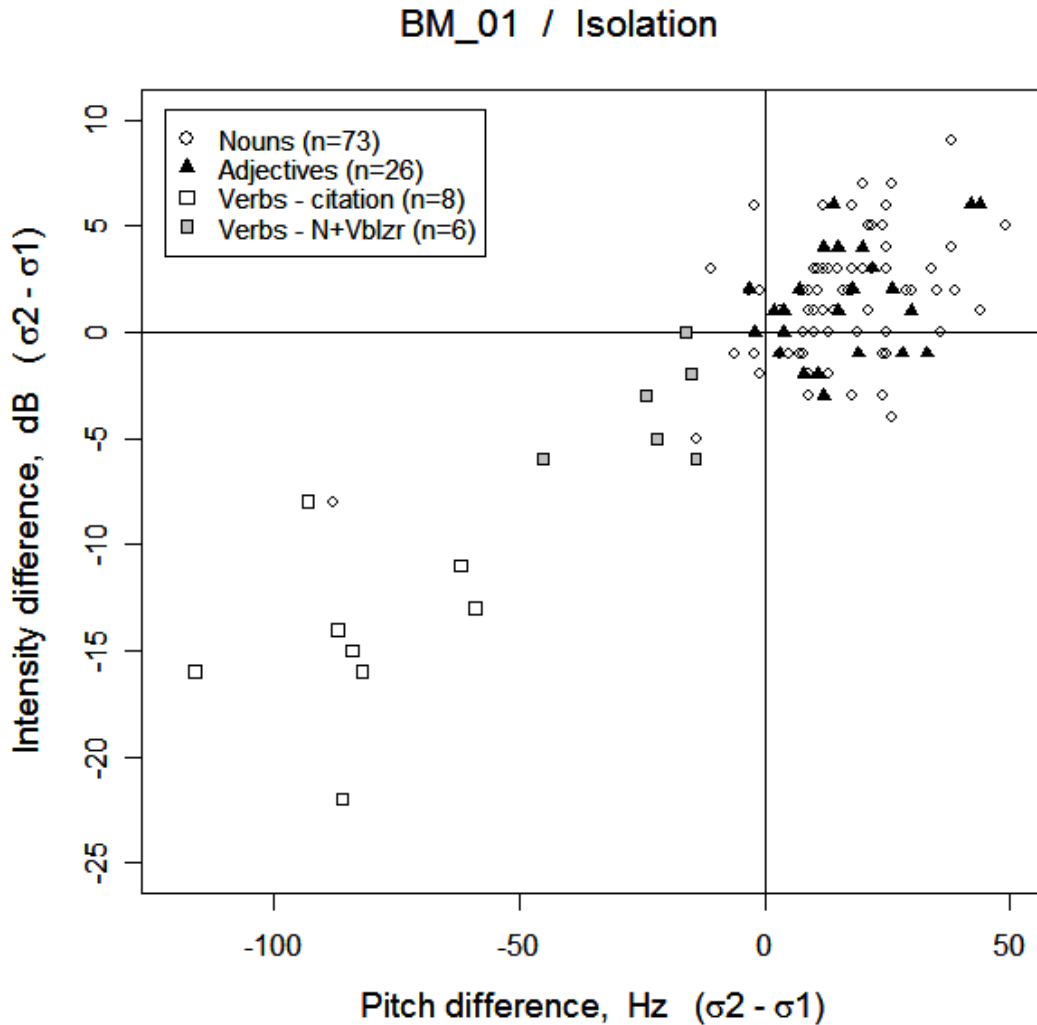


One very important point which becomes clear from these graphs is that, even when an acoustic contrast is at its very weakest, it would still be very difficult for a listener to confuse lexical categories. That is, for both isolation and frame forms, if

we look at the non-verb and the verb which plot closest to one another, they are still quite far apart. Verbs and non-verbs occupy unmistakably distinct acoustic spaces.

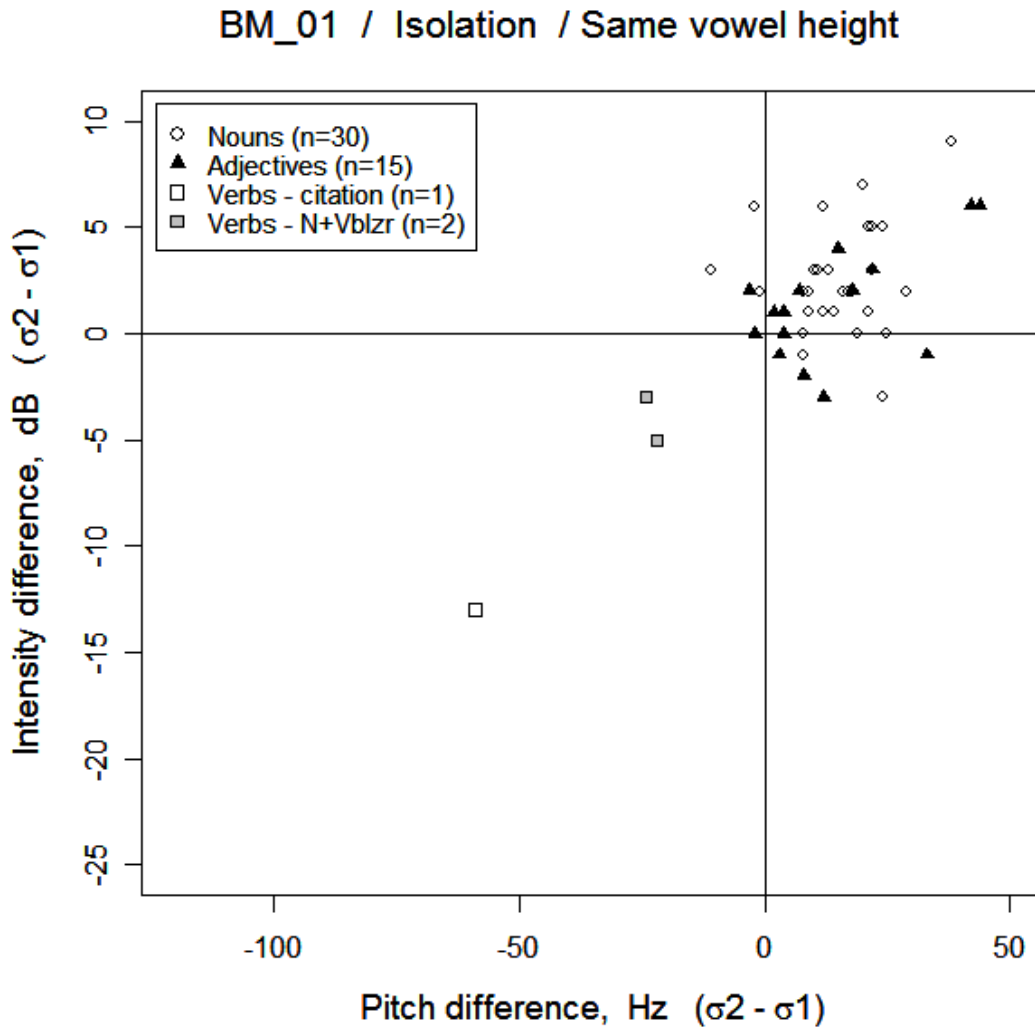
Pitch differences and intensity differences across syllables are plotted for speaker BM_01 in Figure 4.105 below. Again, nouns and adjectives pattern together and are distinct from verbs, with the exception of one point at coordinates (-88,-8). This corresponds to the noun [ʰɕuk.po] *gzhug.po?* ‘buttocks’, which is pronounced with a strong stress on the first syllable, just like a verb. This anomalous form was excluded from all analyses in this study.

Figure 4.105 BM_01 / Isolation: Pitch difference vs. intensity difference



This plot shows that there is less of a distinction between verbs and non-verbs than there was for speaker BSh_03. This may reflect the more casual variability observed throughout the recording session with this speaker. But it also reflects the intrinsic variation of both intensity and pitch with vowel height. In Figure 4.106 below, when this factor is eliminated by considering only forms with vowels of the same height in both syllables, the contrast in lexical categories is sharpened.

Figure 4.106 BM_01 / Isolation / Same vowel height: Pitch difference vs. intensity difference



4.5 Vowel duration

Evaluation of vowel duration as an acoustic correlate of stress must take into consideration several factors which may have an incidental effect on duration. As discussed in section 2.3.2.1, (a) vowels are often longer in open syllables than in

closed syllables; (b) low vowels are often longer than high vowels; and (c) vowels are often lengthened in utterance-final position. The first two factors mean that syllable closure and vowel quality must be controlled for. The third factor means that, if a longer vowel occurs in σ_2 of a noun, adjective, or numeral produced in isolation, it will be impossible to determine whether this lengthening reflects a correlation with σ_2 stress, or is simply related to the word's position in the utterance. That is, there is essentially no point in examining vowel duration in nouns, adjectives, or numerals produced in isolation, because we can never really know what the duration measurements mean.

This third factor is not pertinent to the verbs produced in isolation, since the syllable of interest is not the final syllable of the utterance – verbs are stressed on σ_1 , and this is where we must look for a correlation with stress. It is also not pertinent to words of any lexical category produced in the sentence frame; here, even if we are interested in properties of σ_2 of a noun or adjective, this is not the final syllable of the utterance.

Analyses of vowel duration measurements for speakers BSh_03 and BM_01 are presented in sections 4.5.1 and 4.5.2 below. It is only in verbs that we see some evidence that the speaker manipulates vowel duration in accord with the stress pattern.

4.5.1 Vowel duration for speaker BSh_03

If vowel duration were a meaningful correlate of stress for non-verbs, we would expect the vowel in σ_2 to be longer than the vowel in σ_1 . Though many

individual tokens do exhibit this pattern, once incidental factors are controlled for, the remaining sample is too small to be very informative. For the frame forms of nouns, only the one reduplicated form shows a possible perceptually significant duration contrast. The control group of adjectives produced in the sentence frame is too small to assess. The numerals were only produced in isolation, and so cannot be evaluated at all.

For verbs, when we consider subsets which are controlled for vowel quality, both isolation and frame forms do, indeed, exhibit a tendency for vowels to be longer in σ_1 , the stressed syllable. In most cases the duration contrast is probably not perceptually significant. But it is notable nonetheless, because all of these forms have a [closed.open] syllable structure, which favors a longer vowel in σ_2 . Thus the evidence suggests that the speaker manipulates vowel duration in accord with the stress pattern, but the resultant contrast is not highly perceptible, and is only observed in a limited and controlled subset.

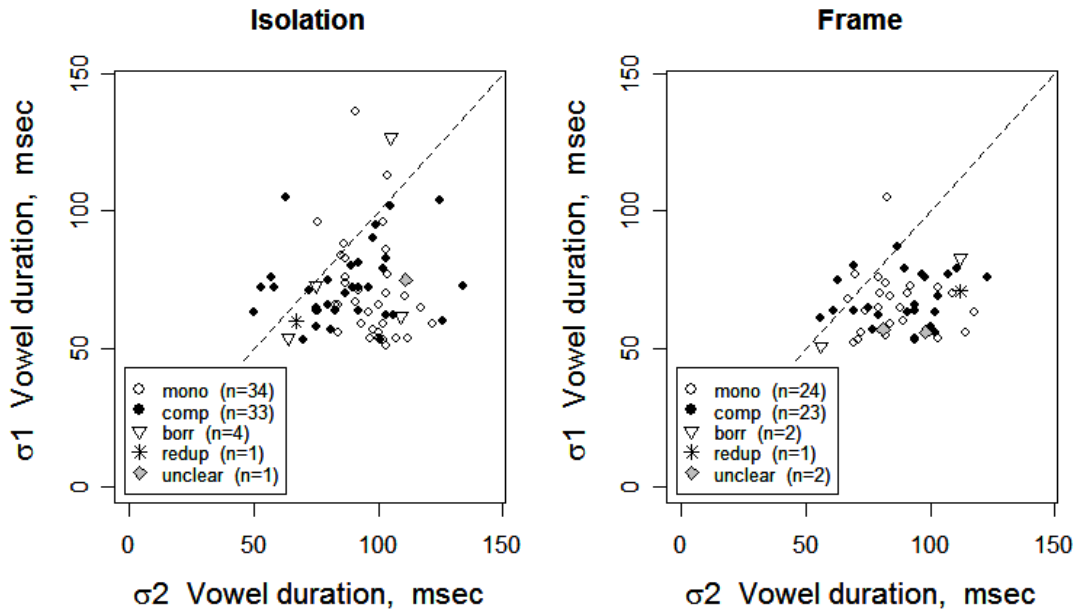
4.5.1.1 BSh_03 Nouns

Vowel duration measurements for BSh_03 nouns are plotted in Figure 4.107 below. As shown, most nouns produced in isolation and in the sentence frame have a longer vowel in σ_2 ; most points plot to the right of the dashed line, and the box-and-whisker plots show no overlap of either the notches or the boxes. At first glance, this gives the impression that vowel duration must be a strong correlate of σ_2 stress. In fact, though, this is not the case: when the intrinsic effects of position in the utterance,

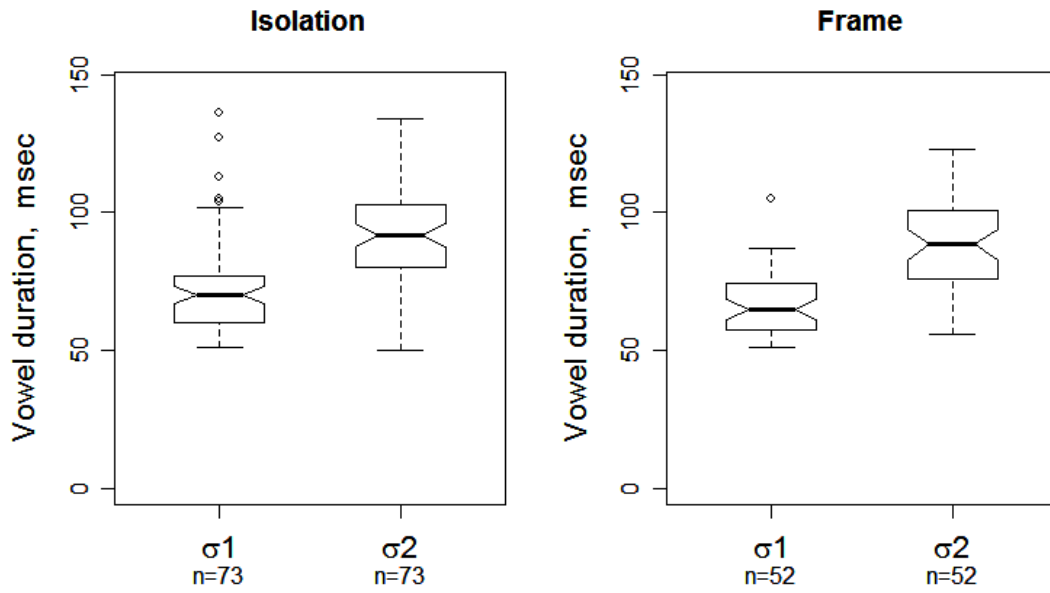
vowel quality, and syllable structure are taken into consideration, there is no evidence of a meaningful correlation between vowel duration and stress.

Figure 4.107 BSh_03 / Nouns: Gross vowel duration (words with diphthongs and compensatory lengthening excluded)

BSh_03 / Nouns



BSh_03 / Nouns



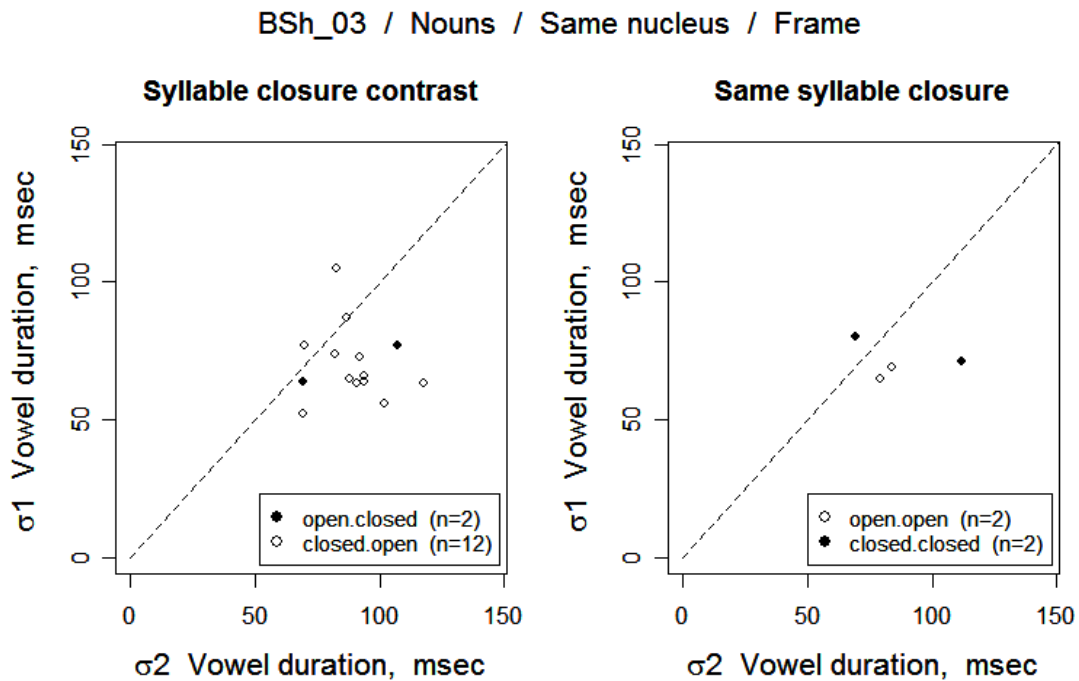
To begin, the isolation forms can be immediately excluded from further consideration. As noted above, it is impossible to distinguish between σ_2 lengthening associated with stress, and σ_2 lengthening associated with a syllable's utterance-final position. Even though, in the graphs above, there is little apparent difference between isolation and frame forms, there is still no point in pursuing the analysis; we can never definitively know whether or not the speaker manipulates vowel duration to convey stress.

For the frame forms – where the stressed syllable is *not* in utterance-final position – the intrinsic variation in vowel duration as a function of vowel quality can be controlled for by considering only those nouns which have the same vowel in both syllables. This subset can then be evaluated in terms of syllable closure, as in Figure 4.108 below.

The graph on the left below shows the distribution of words with different syllable closure patterns. In nearly all cases, the vowel is longer in σ_2 , the stressed syllable. For the [closed.open] nouns – such as [gar.'ba] *mgar.ba* 'blacksmith' and [lom.'go] *lo.mgo* 'first year' – this correlation may be attributable to the contrast in syllable closure. Of greater interest is the fact that the two [open.closed] nouns – [da.'p^haŋ] *mda'. 'phangs / mda'. 'phang* 'archery festival' and [lo.'skor] *lo.skor* 'twelve-year calendar cycle' – also have a longer vowel in σ_2 . On the basis of syllable closure, they might be predicted to instead have a longer vowel in σ_1 .

However, with only two tokens, this is not a convincing indication of a correlation between stress and vowel duration.

Figure 4.108 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration vs. syllable closure contrast duration (words with diphthongs and compensatory lengthening excluded)

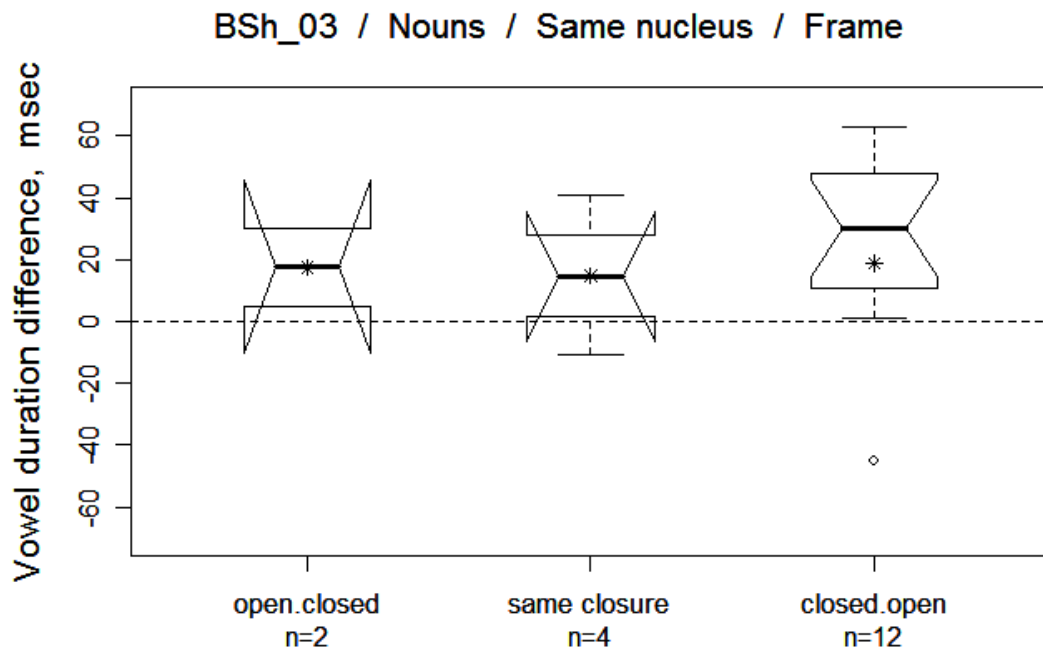


The real test case for a correlation with stress is the subset of nouns which not only have the same vowel in both syllables, but which also have the same closure type on both syllables; this control group is shown on the right above. If vowel duration were a correlate of stress, we would expect words in this group to consistently have a longer vowel in σ_2 . This is the case for both of the [open.open] nouns – [la.'ma] *bla.ma* ‘lama’ and [ra.'ma] *ra.ma* ‘goat’. But the [closed.closed]

nouns – [bar.'tʃat] *bar.chad* ‘obstruction, obstacle’ and [çox.'çox] *shog.shog* ‘paper’ – are split, one falling to either side of the dashed line.

Indeed, in the box-and-whisker plot in Figure 4.109 below, for both the [open.closed] and the [same.closure] nouns the notches representing the 95% confidence interval about the median are quite broad – since the sample size is so small – and span the dashed “Vowel duration difference = 0” line. This indicates that the median vowel duration difference is not significantly different than zero.

Figure 4.109 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration differences vs. syllable closure contrast



Paired-sample t-tests, summarized in Table 4.27 below, show the same result for these two groups, with $p > 0.05$. It is only the nouns with a [closed.open] structure

that have a significantly longer vowel in σ_2 . For this group, the mean difference in duration across syllables is 19 msec. Here, of course, the contrast may be attributable to syllable structure.

Table 4.27 BSh_03 / Nouns / Same nucleus / Frame: Vowel duration by syllable closure type: paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
					lower	upper
open.closed	1.4	1	0.3949	18	-141	176
same closure	1.3892	3	0.2589	15	-19	48
closed.open	2.9971	11	0.01214	19	5.0	32

Finally, vowel duration ratios (short vowel : long vowel) for these four words are shown in Table 4.28 below. As discussed in section 2.5.2.3, this ratio should be ≤ 0.5 for the duration contrast to be perceptually significant. This is not the case for any of the nouns in the control group.

Table 4.28 BSh_03 / Nouns / Same nucleus / Same closure / Frame: Vowel duration ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_1/σ_2)
lama	la. 'ma	bla.ma	14	0.82
goat	ra.'ma	ra.ma	15	0.82
obstruction	bar.'tʃat	bar.chad	-11	1.16
paper	ʃoʔ.'ʃoʔ	shog.shog	41	0.63

In conclusion, though gross measurements for the full sample show longer vowels in σ_2 than in σ_1 , once we control for all the factors which may have an incidental effect on duration, there is little evidence to suggest that vowel duration may be an acoustic correlate of stress for BSh_03 nouns. The only token with a

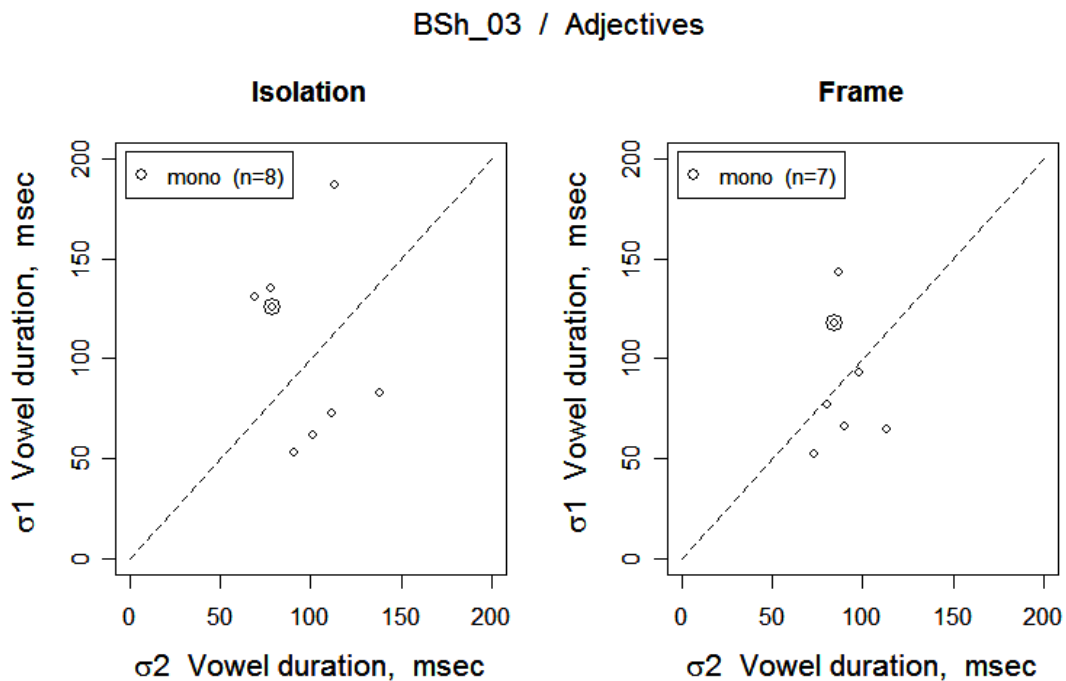
notably longer vowel in σ_2 is the reduplicated form, [çok.'çok] *shog.shog* ‘paper’. It might be helpful, in future, to evaluate a larger set of reduplicated nouns.

4.5.1.2 BSh_03 Adjectives

The sample of adjectives recorded from speaker BSh_03 is simply too small to determine whether or not vowel duration is an acoustic correlate of stress. There are not enough tokens to control for the effects of position in the utterance, vowel quality, and syllable closure.

Vowel duration measurements are plotted in Figure 4.110 below. For both isolation and frame forms, points fall to either side of the dashed line. The one word with compensatory lengthening in σ_1 is [mã.'mo] *mang.po* ‘many’. This is the only word with an [open.open] syllable closure type.

Figure 4.110 BSh_03 / Adjectives: Gross vowel duration (words with compensatory lengthening circled)



All of the other BSh_03 adjectives have a [closed.open] structure. On this basis, they are all predicted to have an intrinsically longer vowel in σ_2 . Furthermore, in the isolation forms, σ_2 is in utterance-final position, a second factor favoring a longer vowel here. Given these confounding influences, it is impossible to determine if vowels are lengthened in σ_2 in order to convey stress.

As an aside: among both isolation and frame forms, there are several adjectives in which the vowel is longer in σ_1 , plotting close to or left of the dashed line in the figure above. These three words – [mar.'p^ho] *dmar.po* ‘red’, [kar.'p^ho] *dkar.po* ‘white’, and [sar.'p^ha] *gsar.pa* ‘new’ – all have an [-r] coda in σ_1 , which may have a phonological lengthening effect on the preceding vowel. They also all have an

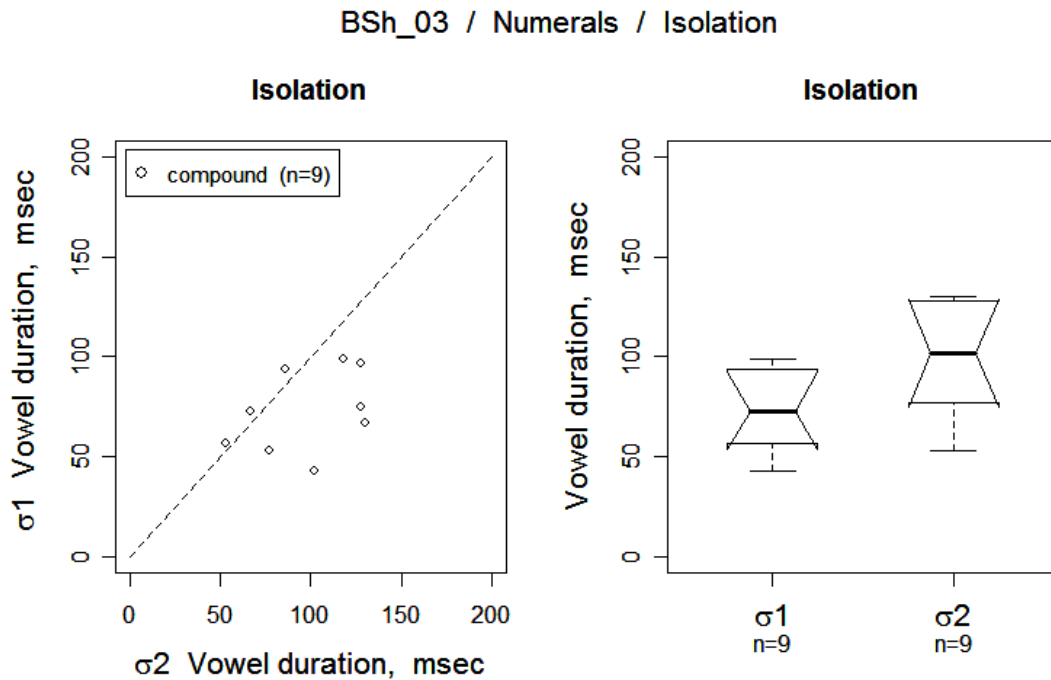
aspirated onset in σ_2 . When segment boundaries are identified in the textgrid, the period of aspiration is assigned to the onset consonant rather than to the nuclear vowel. This may result in a relatively short vowel, compared to words in which the onset consonant is unaspirated. These two hypotheses could be investigated in future with a larger data set.

4.5.1.3 BSh_03 Numerals

No conclusions can be formed regarding vowel duration in BSh_03 numerals, either. Since the numerals were recorded only in isolation, the syllable of interest as the locus of stress – σ_2 – is subject to lengthening due to its utterance-final position. Contrasts in vowel quality and syllable closure also play a confounding role here.

As shown in Figure 4.111 below, vowel duration is indeed longer in σ_2 for the majority of numerals, with six of the nine points falling to the right of the dashed line, and a higher median value in σ_2 than in σ_1 . However, since none of the incidental factors which influence vowel length can be teased out, it is impossible to determine whether or not this represents a genuine correspondence between duration and stress.

Figure 4.111 BSh_03 / Numerals: Vowel duration

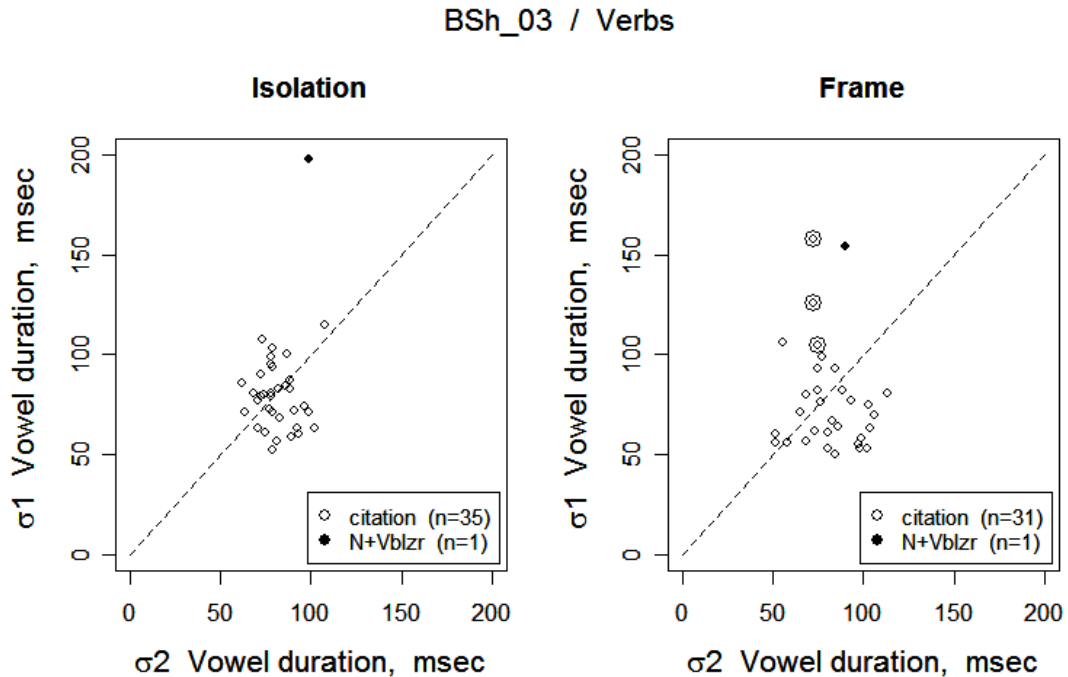


4.5.1.4 BSh_03 Verbs

Balti verbs present a different scenario than the non-verbs. Here, the stressed syllable is $\sigma 1$, which – in disyllabic words – is never subject to utterance-final lengthening, even when produced in isolation. Thus isolation forms as well as frame forms can be evaluated.

Vowel duration measurements for BSh_03 verbs are plotted in Figure 4.112 below. These graphs show the gross, overall pattern; at this stage of the analysis, the intrinsic effects of vowel quality and syllable structure have not been taken into consideration.

Figure 4.112 BSh_03 / Verbs: Gross vowel duration (circles indicate compensatory lengthening on $\sigma 1$)



Among the isolation forms, the only verb with a long vowel is the lone N+Vblzr form ['za:n.za] *zan.za* 'to eat food'. The remaining verbs – the citation forms – are more or less evenly distributed to either side of the dashed line, indicating that sometimes the vowel in $\sigma 1$ is longer, and sometimes the vowel in $\sigma 2$ is longer.

['za:n.za] *zan.za* 'to eat food' also has a long vowel when produced in the sentence frame. Several other frame forms have a long vowel in $\sigma 1$ due to compensatory lengthening: ['strã:ma] *srang.ma* 'to straighten', and ['lã:ma] *lang.ma* 'to get up, to rise'. The remaining frame forms fall to either side of the dashed line, but the majority have a longer vowel in $\sigma 2$.

Since the objective here is to determine whether the speaker manipulates vowel duration to convey stress, the N+Vblzr form and the cases of compensatory lengthening are excluded from further consideration.

The two other factors which need to be controlled for are vowel quality contrast and syllable closure contrast. Since all of the forms which remain under consideration are citation forms – with a variant of [-pa] or [-ma] as the second syllable – vowel quality can be controlled for by considering only those verbs which have [a] in the first syllable, such as ['blaq.pa] *'breg.pa* 'to shave' and [ʔ'ɛzar.pa] *'dzar.ba / gzar.ba / bzar.ba* 'to drip'.

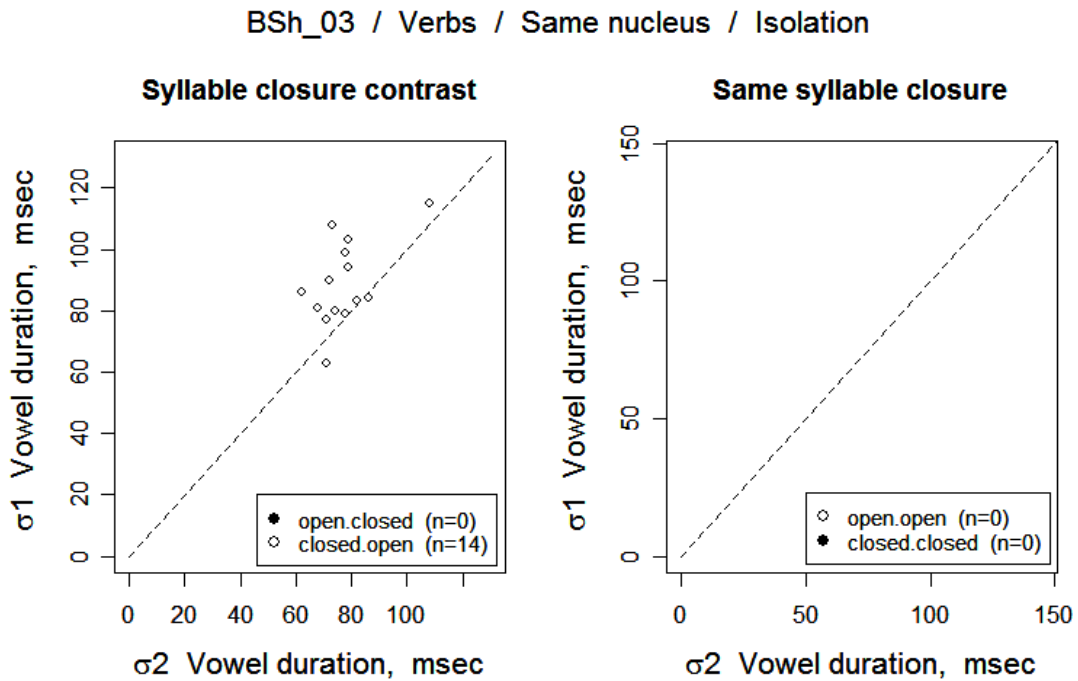
Isolation forms of the [a.a] verbs are plotted in terms of syllable closure contrast in Figure 2.11 above. Of course, among the citation forms – ending with a variant of [-pa] – the [open.closed] and [closed. closed] subsets have no members. And since there are no [open.open] [a.a] verbs⁴², there is nothing to represent in the graph on the right. But there are fourteen [a.a] verbs with a [closed.open] syllable closure pattern. For this group, there are two factors which favor a longer vowel duration in σ_2 , and which would lead us to predict that these verbs would plot to the right of the dashed line. First, the syllable template is [closed.open], and vowels are generally longer in open syllables than in closed syllables. Second, in isolation, the

⁴² The only [open.open] verbs are [ʔtɛ.ua] ?? 'to destroy, to raze' and [tʰri.bya] *dri.bya* 'to smell (transitive)'. In these cases, the vowel is longer in σ_2 , the syllable which is *not* stressed. This may be attributable to the [higher.lower] vowel height pattern, since high vowels are intrinsically shorter than low vowels (Lehiste, 1970:18).

vowel in σ_2 is in utterance-final position and likely to be lengthened on that basis, as well.

Instead, though, the [a.a] isolation forms plot to the left of the y-axis, indicating that the vowel is longer on σ_1 . The most plausible explanation for this is that the vowel is preferentially lengthened in correspondence with the stress pattern.

Figure 4.113 BSh_03 / Verbs / Same nucleus / Isolation: Vowel duration vs. syllable closure contrast (N+Vblzr forms omitted)



A paired-sample t-test, summarized in Table 4.29 below, indicates that the difference in vowel duration across syllables for this group is statistically significant, with $p < 0.05$. On average, though, the vowel in σ_1 is only 12 msec longer than the vowel in σ_2 , a contrast which is only trivially longer than the difference limen of 10 msec identified by Lehiste (1970). On repeated sampling, a mean difference of a mere

4.6 msec is probable, as indicated by the upper 95% confidence limit. Such a contrast would not be perceptible, much less perceptually significant.

Table 4.29 BSh_03 / Verbs / [closed.open] / Same nucleus / Isolation / Vowel duration: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
						lower	upper
BSh_03	isolation	-3.5824	13	0.003342	-12	-18	-4.6

Vowel duration ratios (short vowel : long vowel) for the verbs in the control group are summarized in Table 4.30 below. The ratio is never ≤ 0.5 ; Lehiste (1970) identified this as a threshold for perceptual significance, though she noted considerable cross-linguistic variation. It is possible that the contrast in vowel duration across syllables is indeed perceptually significant for the top few words in the table below. For the remainder, though, it is probably not perceptually significant.

Table 4.30 BSh_03 / Verbs / Isolation / [closed.open] / Same nucleus: Vowel duration ratios

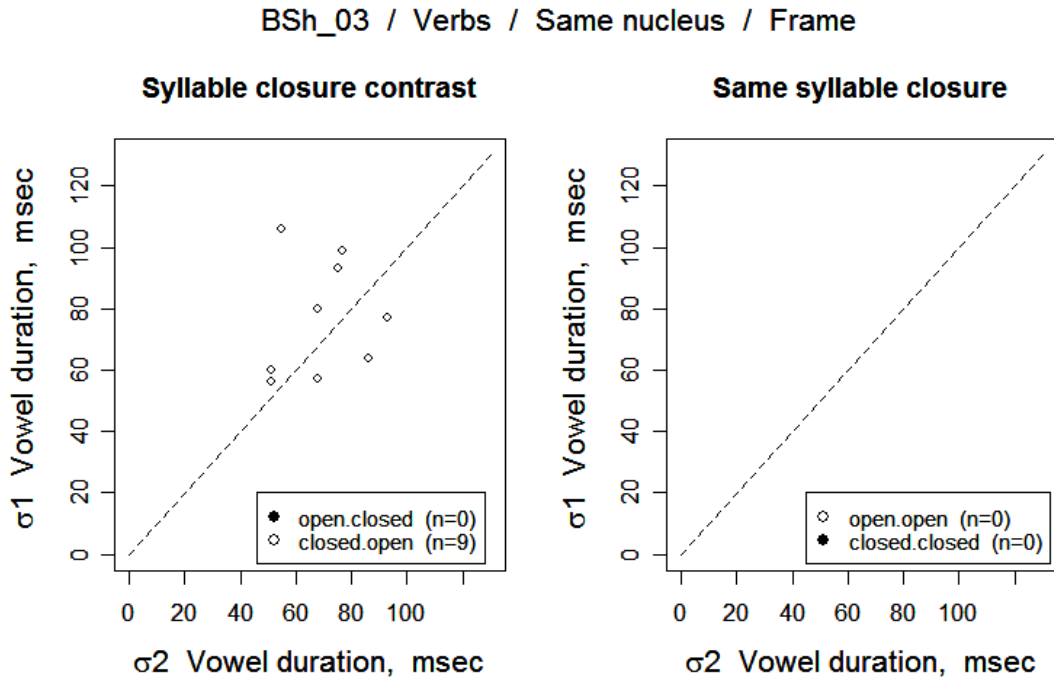
Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_2/σ_1)
to deprive	ʰzgak.pa	<i>sgag.pa / bsgags.pa</i>	-35	0.68
to light ignite	'spar.ba	<i>spar.ba</i>	-24	0.72
to rise (sun)	'ʃar.ba	<i>shar.ba</i>	-24	0.77
to drip	ʰɛzar.pa	<i>'dzar.ba / gzar.ba / bzar.ba</i>	-21	0.79
to assign a name	'taq.pa	<i>btag.pa</i>	-18	0.80
to fill	'skaŋ.ma	<i>bkang.ma ?</i>	-13	0.84
to be clean, pure	'daq.pa	<i>dag.pa</i>	-15	0.84
to be tired	'ɣlat.pa	<i>*glad.pa</i>	-6	0.92
to stretch	'škyaŋ.ma	<i>brkyang.ma ? rkyangs.ma ?</i>	-6	0.93
to straighten	'strāŋ.ma	<i>srang.ma</i>	-7	0.94
to grind	'thaq.pa	<i>'thag.pa</i>	-1	0.99
to shave	'blak.pa	<i>'breg.pa</i>	-1	0.99
to get up, rise	'laŋ.ma	<i>lang.ma</i>	2	1.02
to release	'škat.pa	??	8	1.13
Mean			-12	0.82

* reconstructed Proto-Tibetan form provided by Roland Bielmeier (p.c., 2008).

Nonetheless, this contrast is notable in light of the fact that duration is predicted – on the basis of syllable structure and utterance-final position – to be longer on σ_2 , instead. What this suggests is that the speaker does make some [unconscious] effort to manipulate vowel duration in accord with the stress pattern, but that the effect is not sufficient to override contradictory incidental factors, such as vowel quality and syllable closure.

Among the BSh_03 frame forms, there are only a few verbs that show a correspondence between vowel duration and stress. The subset of citation forms with the vowel [a] in both syllables is plotted in Figure 4.114 below in terms of syllable closure type. Here – in contrast to the isolation forms – the [closed.open] forms fall to either side of the dashed line. The verbs [ʃʁkyaŋ.ma] *brkyang.ma* ? *rkyangs.ma* ? ‘to stretch’, [xlɑt.pa] **glad.pa* ‘to be tired’, and [ʃpar.ba] *spar.ba* ‘to light, to ignite’ have a longer vowel in σ_2 , as one might predict based on the syllable closure pattern. But six of the nine verbs – including [ʃar.va] *shar.ba* ‘to rise (sun)’ and [ʔzgak.pa] *sgag.pa* / *bsgags.pa* ‘to deprive’ – have a longer vowel in σ_1 , which is counter to expectations. In these six cases, it is possible that the speaker has lengthened the vowel because it is stressed

Figure 4.114 BSh_03 / Verbs / Same nucleus / Frame: Vowel duration vs. syllable closure contrast (N+Vblzr forms and cases of compensatory lengthening omitted)



A paired-sample t-test comparing vowel duration across syllables for the nine [closed.open] frame forms is summarized in Table 4.31 below. As shown, the contrast in duration is not significant: $p > 0.05$, and the 95% confidence interval includes zero.

Table 4.31 BSh_03 / Verbs / [closed.open] / Same nucleus / Frame / Vowel duration: results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
						lower	upper
BSh_03	frame	-1.0129	8	0.3408	-7.6	-25	9.6

Vowel duration ratios (short vowel : long vowel) for this subset, shown in Table 4.32 below, show results consistent with these statistics. It is only in the case of

the verb [ʷzgak.pa] *sgag.pa* / *bsgags.pa* ‘to deprive’ that the vowel duration ratio is close to 0.50, the threshold for perceptual significance.

Table 4.32 BSh_03 / Verbs / Frame / [closed.open] / Same nucleus: Vowel duration ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_2/σ_1)
to deprive	ʷzgak.pa	<i>sgag.pa</i> / <i>bsgags.pa</i>	-51	0.52
to drip	'kzar.pa	'dzar.ba / <i>gzar.ba</i> / <i>bzar.ba</i>	-22	0.78
to rise (sun)	'ʃar.va	<i>shar.ba</i>	-18	0.81
to be clean, pure	'dak.pa	<i>dag.pa</i>	-12	0.85
to assign a name	'taq.pa	<i>btag.pa</i>	-9	0.85
to grind	'thaq.pa	' <i>thag.pa</i>	-5	0.91
to be tired	'xlat.pa	* <i>glad.pa</i>	11	1.19
to light, ignite	'spar.ba	<i>spar.ba</i>	16	1.21
to stretch	'ʃkyaŋ.ma	<i>brkyang.ma</i> ? <i>rkyangs.ma</i> ?	22	1.34
Mean			-7	0.84

For the frame forms, then, there are a few verbs – controlled for incidental factors – in which the vowel is longer in σ_1 , the stressed syllable. However, these individual cases do not constitute a consistent or significant pattern.

In summary, there are some cases among both isolation and frame forms of BSh_03 verbs in which the vowel is longer in σ_1 , the stressed syllable, even when it is predicted to be longer in σ_2 on the basis of syllable closure type. But this correspondence between vowel duration and stress is quite limited. Since there are so

many other verbs in which the vowel is longer in σ_2 – the *unstressed* syllable – a listener cannot depend on vowel duration as a reliable indicator of stress.

4.5.2 Vowel duration for speaker BM_01

For speaker BM_01, all words were recorded in isolation only. This means that, for non-verbs, the stressed syllable – σ_2 – is always in utterance-final position. There is no way to distinguish vowel length that might be attributable to stress from vowel length that might be attributable to this final position in the utterance. Thus nothing useful can be gained by analyzing nouns and adjectives produced by this speaker, and I do not consider them here at all.

This problem does not arise for verbs, which are stressed on σ_1 . As demonstrated in section 4.5.2.1 below, most verbs have a longer vowel in this syllable; the contrast across syllables is probably perceptually significant.

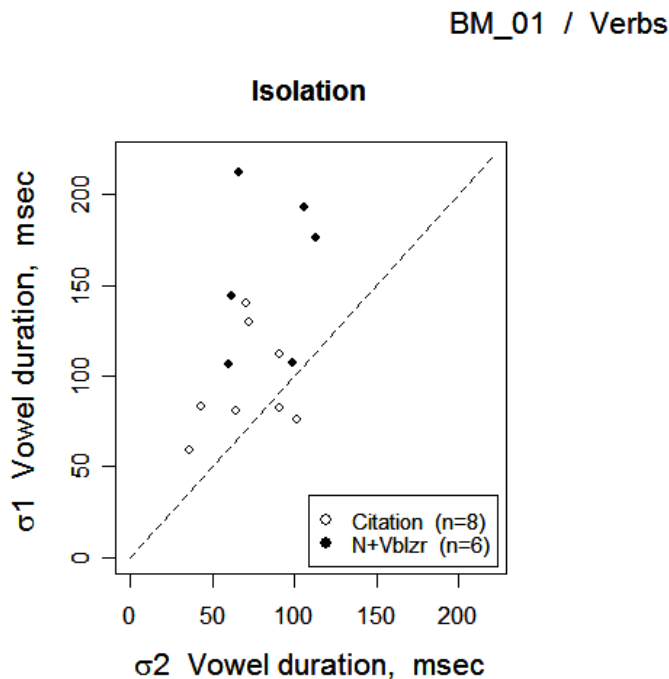
In general, as with the other acoustic parameters, the vowel duration measurements for speaker BM_01 show greater variation than those for speaker BSh_03.

4.5.2.1 BM_01 Verbs

Since Balti verbs are stressed on σ_1 , any potential correlation between vowel duration and stress will not be obscured by the potential for utterance-final lengthening – which was the case for non-verbs.

Duration measurements for both citation and N+Vblzr forms are plotted in Figure 4.115 below. As shown, with only two exceptions all verbs fall to the left of the dashed line, indicating a longer vowel in σ_1 , the stressed syllable.

Figure 4.115 BM_01 / Verbs: Gross vowel duration contrasts



This pattern is particularly noteworthy, given that several factors instead favor a longer vowel in σ_2 : For both citation forms and N+Vblzr forms, σ_2 is in utterance-final position, and thus might be expected to have the longer vowel on that basis. Furthermore, all of the citation forms have a [closed.open] syllable structure, and all but one of them have a [higher.lower] vowel height contrast – two other factors which favor a longer vowel in σ_2 . However, the speaker has evidently manipulated vowel duration to override these intrinsic effects, and to instead correspond with σ_1 stress.

For the N+Vblzr form ['sk^hat.zɛŋ] *skad.zer* 'to call, to shout', the difference in duration across syllables is only -8 msec, which is below Lehiste's minimal difference limen of 10 msec. For all of the other N+Vblzr forms, the difference ranges from -46 msec to -146 msec.

The two citation forms which fall to the right of the dashed line are ['dʒik.p^ha] *'jigs.pa* 'to fear' and ['mɪn.ma] *sbyin.ma* ? 'to give', where the vowels in σ_2 are longer by 9 msec and 26 msec, respectively. In these cases, the intrinsic effects of vowel quality probably play a role: vowels are shorter in high vowels than in low vowels, and these are the only two verbs with [i] or [ɪ] in σ_1 , contrasting with [a] in σ_2 .

With all of the variation in syllable closure type and vowel height, it is not possible to identify a control group for either morphological type of verb. It would certainly be useful to evaluate a larger sample of verbs in future, but even with this limited sample, it is clear that vowels are indeed mostly longer in σ_1 , the stressed syllable.

For the N+Vblzr forms, a paired-sample t-test shows that this correlation is statistically significant: $p < 0.05$, as shown in Table 4.33 below. For the citation forms, the contrast in duration across syllables is *not* statistically significant: $p > 0.05$.

Table 4.33 BM_01 / Verbs / Vowel duration: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
						lower	upper
BM_01	N+Vblzr	-3.8206	5	0.01236	-72	-120	-24
	citation	-2.1448	7	0.06914	-24	-51	2.5

Vowel duration ratios for the N+Vblzr forms are summarized in Table 4.34 below. The short vowel : long vowel ratio averages 0.57 – or 0.50 if the anomalous case of ['sk^hat.zɛɽ] *skad.zer* ‘to call, to shout’ is omitted. This suggests that the duration difference is probably perceptually significant. Lehiste (1970) reported a considerable degree of cross-linguistic variation in perceptually significant duration ratios; we cannot really be certain about perceptual significance in Balti without experimental evidence.

Table 4.34 BM_01 / Verbs / N+Vblzr: Vowel duration ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ2/σ1)
to listen	'sna.bya	<i>rna.bya</i>	-146	0.31
to close a door	'zgo.tɕ ^h uk ^h	<i>sgo.bcug</i>	-82	0.43
to cook food	'zan.ɸtɕøɽ	<i>zan.bcos</i>	-87	0.55
to spread manure	'lut.taŋ	<i>lud.btang</i>	-46	0.57
to eat food	'zan.za	<i>zan.za</i>	-63	0.64
to call, to shout	'sk ^h at.zɛɽ	<i>skad.zer</i>	-8	0.93
Mean			-72	0.57
Mean w/o ‘to shout’			-85	0.50

As shown in Table 4.35 below, vowel duration ratios for the citation forms average 0.63, when [ˈdzɪk.pʰa] *'jigs.pa* ‘to fear’ and [ˈmɪn.ma] *sbyin.ma ?* ‘to give’ – the two forms where vowel height contrast plays a role – are excluded. This may also be a perceptually significant contrast for Balti.

Table 4.35 BM_01 / Verbs / Citation / [closed.open]: Vowel duration ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_2/σ_1)
to know	ˈçɛs.pa	<i>shes.pa</i>	-69	0.51
to meet	ˈthuk.pə	<i>thug.pa</i>	-40	0.52
to light a fire	ˈspar.va	<i>spar.ba</i>	-58	0.55
to keep, protect	ˈstun.ma	<i>srung.ma</i>	-23	0.61
to fly	ˈphur.ba	<i>'phur.ba</i>	-17	0.79
to be late	ˈŋgor.va	<i>'gor.ba</i>	-21	0.81
to fear	ˈdzɪk.pha	<i>'jigs.pa</i>	9	1.11
to give	ˈmɪn.ma	<i>sbyin.ma ?</i>	26	1.34
Mean			-24	0.78
Mean w/o [i], [ɪ]			-38	0.63

These results suggest that vowel duration is probably a perceptually significant correlate of σ_1 stress in both N+Vblzr and citation forms of BM_01 verbs, especially since incidental factors favor a longer vowel in σ_2 . It would be helpful, in future, to consider a sample of verbs which is controlled for vowel quality and syllable closure.

4.6 Recapitulation

Table 4.36 below summarizes the contributions made by pitch, pitch slope, intensity, and vowel duration to the stress patterns perceived in Balti.

Pitch is the most consistent and most robust correlate of stress for words in all lexical categories – both the non-verbs, stressed on σ_2 , and the verbs, stressed on σ_1 . A listener would be able to identify the σ_2 stress pattern on BSh_03 nouns, adjectives, and numerals from the pitch contrast alone.

Intensity is also a robust correlate of stress for verbs. It is a limited and weak correlate of stress for BM_01 nouns, but otherwise is not important to the perception of stress in Balti non-verbs.

Pitch slope does not play any role at all in Balti stress. Vowel duration shows a limited and weak correlation with stress in BSh_03 verbs. For verbs produced by speaker BM_01, the duration contrast across syllables is probably perceptually significant for N+Vblzr and citation forms.

Table 4.36 Acoustic correlates of stress in Balti

	Speaker	Setting	Pitch	Pitch slope	Intensity	Vowel duration
Nouns (σ_2)	BSh_03	isol	robust	no	no	NA
		frame	robust	no	no	no – SS
	BM_01	isol	yes	no	limited	NA
Adjs (σ_2)	BSh_03	isol	robust	no	no – SS	NA
		frame	robust	no	no – SS	no – SS
	BM_01	isol	yes	no	no	NA
Nums (σ_2)	BSh_03	isol	robust	no	no	NA
Verbs (σ_1)	BSh_03	isol	robust	no	robust	limited
		frame	robust	no	robust	limited
	BM_01	citation	robust	no	robust	probable *
		N+Vblzr	robust	no	limited	probable *

NA This parameter could not be analyzed.

SS Small Sample. The conclusion presented in the table should be considered tentative, as it was based on a very limited data set.

* For both verb types, the vowel is almost always longer in σ_1 than in σ_2 , despite a number of factors which would favor the opposite pattern. This suggests that the speaker manipulates duration to accord with stress. However, without a controlled sample, a conservative conclusion is to say that this is a “probable” stress correlate.

In sum, pitch is an important cue for stress in Balti for all lexical categories.

Intensity is an important correlate of stress for verbs. Vowel duration also seems to be manipulated in accord with the stress pattern perceived on verbs, but the effect is limited, given the incidental effects of vowel quality and syllable structure.

5. Acoustic correlates of stress in Rebkong Amdo Tibetan

In this chapter I present my analysis of the acoustic correlates of stress in Rebkong Amdo, based on recordings from speakers AR_04 and AR_05.

Results are summarized in section 5.1, organized by lexical category. Details of the analysis can be found in sections 5.2 through 5.5, organized by acoustic parameter. Pitch data is considered in section 5.2 – first for speaker AR_04 (first nouns, then adjectives, then verbs), and then for speaker AR_05 (first nouns, then adjectives, then numerals, then verbs). Pitch slope is similarly considered in section 5.3, intensity data in section 5.4, and vowel duration data in section 5.5.

Finally, in section 5.6 I provide a recapitulation. Table 5.33 on page 511 provides a grand summary of the role played by each acoustic parameter in conveying stress.

5.1 Summary of acoustic correlates of stress in Rebkong Amdo

The acoustic correlates of stress are summarized for non-verbs in section 5.1.1, and for verbs in section 5.1.2. For both speakers, the majority of the words recorded and analyzed were nouns; there were about a dozen adjectives and a dozen numerals, and even fewer verbs. Since some of the verbs recorded could not be reliably segmented, only a handful remained for analysis.

5.1.1 Non-verbs

As in Balti, Rebkong Amdo non-verbs (nouns, adjectives, and numerals) are stressed on σ_2 , in contrast to verbs, which are stressed on σ_1 .

5.1.1.1 Nouns

Nouns in Rebkong Amdo are stressed on σ_2 . The fundamental frequency-related parameters pitch and pitch slope are the primary cues for this perceived stress pattern. Intensity shows a limited and weak correspondence with stress, with patterns influenced by the intrinsic effects of vowel height. There is no evidence to indicate that vowel duration correlates with stress.

The only morphological types of nouns recorded from speakers AR_04 and AR_05 were monomorphemic and compound. There were no borrowed, reduplicated, or “unclear” forms, as there were in the Balti data set. Rebkong Amdo monomorphemic nouns included [na.'ma] *mna'.ma* ‘bride’, [h^hlai.'pa] *klad.pa* / *glad.pa* ‘brain’, and [ndzɣ.'ɣɣ] *mdzug.gu* ‘digit; finger, toe’; compound nouns included [xtar.'ga] *rta.sga* (horse+saddle) ‘horse saddle’, [hɬɣ.'xa] *phag.sha* (pig+meat) ‘pork’, and [h^hna.'tɕɣɕ] *sna.khrag* (nose+blood) ‘bloody nose; blood from the nose’. For speaker AR_04 I was able to segment and analyze 93 isolation forms and 74 frame forms; for speaker AR_05 I was able to segment and analyze 96 isolation forms and 87 frame forms.

5.1.1.2 Adjectives

Like nouns, adjectives in Rebkong Amdo are stressed on the second syllable. This stress pattern is conveyed primarily by the fundamental frequency-related parameters pitch and pitch slope, in some cases functioning in complementarity.

Intensity exhibits a limited and weak correlation with stress only in the AR_05 frame forms; in adjectives produced by speaker AR_04, intensity is actually higher on σ_1 , in contrast to the perceived stress. There is no evidence of a correlation between vowel duration and stress, though the sample size is too small to be conclusive.

Only a small number of adjectives were recorded for this dialect. Twelve isolation forms and ten frame forms were recorded from speaker AR_04, all of which are monomorphemic: e.g., [ka.'ru] *dkar.po* ‘white’; [tʃʰɣ.'ko] *phyug.po* ‘rich’. The twelve isolation forms and twelve frame forms recorded from speaker AR_05 include several reduplicated forms – e.g., [tʃʰɣ.'tʃʰɣ] *chung.chung* ‘small’, and [lev.'lev] *leb.leb* ‘flat’ – as well as monomorphemic forms, like [sɔ.'ma] *so.ma* ‘new’.

5.1.1.3 Numerals

Like nouns and adjectives, Rebkong Amdo numerals are stressed on the second syllable. This pattern is conveyed primarily by pitch slope, and also by intensity. Pitch measurements suggest a weak correlation with stress. Because the numerals were recorded only in isolation, it is not possible to determine whether there is any correlation at all between vowel duration and stress.

Numerals were recorded only from speaker AR_05. The disyllabic numerals recorded include nominalized forms (e.g., [ksum.'ba] *gsum.pa* ‘third’; [ʔdun.'ba] *bdun.pa* ‘seventh’) as well as compound forms (e.g., [tʃʊb.'dun] *bcu.bdun* ‘seventeen’; [hʌdun.'tʃʊ] *bdun.cu* ‘seventy’). Not all of the disyllabic numerals could

be reliably segmented, leaving only two nominalized forms and nine compound forms to be analyzed. This included three pairs of compounds like ‘seventeen’ and ‘seventy’, in which the same elements are combined in reverse order. These are of particular interest because stress falls on σ_2 regardless of the order of morphemes – clear evidence that stress is governed by syllable position, rather than by segmental or semantic content.

5.1.2 Verbs

Verbs in Rebkong Amdo are stressed on σ_1 , in contrast to the σ_2 stress perceived on non-verbs.

Only a handful of verbs were recorded for each speaker, since the focus at the time I began my study was on nouns. For speaker AR_04, only the frame forms are considered here. Except as monosyllabic imperatives, it proved very difficult for this speaker to produce verbs in complete isolation, without the normal context of tense, aspect, and evidentiality. Given his discomfort with this element of the task, the isolation forms were generally stressed on both syllables, in a very unnatural way, regardless of the pattern on the prompt provided by my research assistant. The speaker was more comfortable producing the verbs in a sentence frame; these recordings sound more natural, and here a σ_1 stress pattern is easy to hear.

Despite the small sample size, several patterns emerge. Both pitch and intensity are always higher on σ_1 . For the AR_04 frame forms, the sample is too small ($n=2$) to determine whether or not these apparent correlations with stress are

statistically significant. For AR_05 verbs, both pitch and intensity are statistically significant correlates of stress, for isolation and frame forms.

Pitch slope is never an acoustic correlate of stress in verbs, because it lacks contextual significance: the downward slopes on each syllable are a direct result of the drop in pitch across syllables. For both speakers, the sample is too small to determine whether vowel duration shows a correlation with stress.

All of the verbs recorded from both speakers have a Noun+Verbalizer structure – e.g., [ˈnta.hɛn] *mda* ‘to shoot an arrow’; [ˈnʌm.biʊ] *gnam.babs* ‘to rain’; [ʒ̥ɛ.tɕʌv] *lan.brgyab* ‘to give an answer’. For speaker AR_05, the verb [ˈnʌm.ˈbau] *gnam.babs* ‘to rain’ was anomalously produced with stress on both syllables in the frame form. This was the only one of the five verbs that sounded like this, and it was only in the frame form – not in the two isolation tokens, and not in the two tokens in the short sentence the speaker made up. What I think is that the speaker produced this word, in this context, with an anomalous stress pattern. I include it on the graphs for reference, but I do not include it in the statistics.

5.2 Pitch

Pitch can be regarded as one of two reflexes of fundamental frequency, the other being pitch slope, which is discussed in section 5.3.

In Balti, pitch was identified as a robust correlate of stress for both speakers, for words in all lexical categories (see section 4.2). This is not quite the situation in

Rebkong Amdo; here, pitch is often – but not always – a robust correlate of stress. In cases where the pitch contrast across syllables is not so strong or so consistent, another acoustic correlate – pitch slope or intensity – generally plays a role in conveying stress.

For speaker AR_04, analyses of nouns, adjectives, and verbs in isolation and in the sentence frame are presented in section 5.2.1. For this speaker, the magnitude of the pitch contrast across syllables is influenced by whether σ_2 is open or closed. The intrinsic variation of pitch as a function of vowel height does not appear to play a significant role.

Pitch patterns in nouns, adjectives, and verbs produced by speakers AR_04 and AR_05 are analyzed in sections 5.2.1 and 5.2.2, respectively.

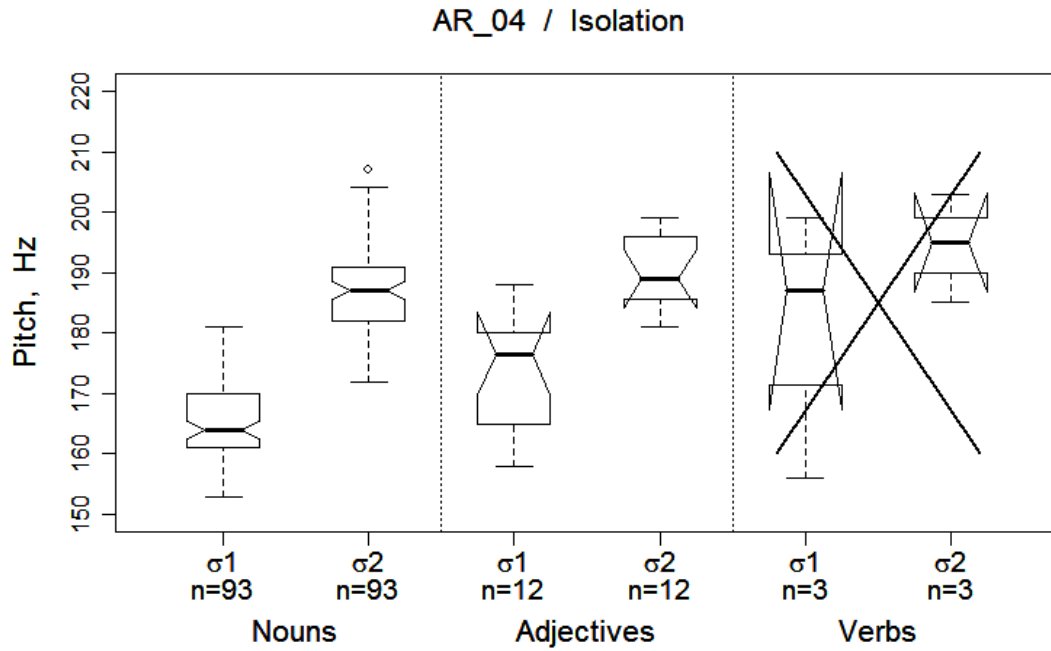
5.2.1 Pitch for speaker AR_04

For speaker AR_04, pitch is a robust and consistent cue for σ_2 stress in both nouns and adjectives, in isolation and in the sentence frame.

Because this speaker was not able to comfortably produce verbs in isolation – without the context of tense / aspect / evidentiality – the data collected for this group of samples is not useful or meaningful. For verbs produced in the sentence frame, though, there is some evidence of a potential correlation between pitch and σ_1 stress.

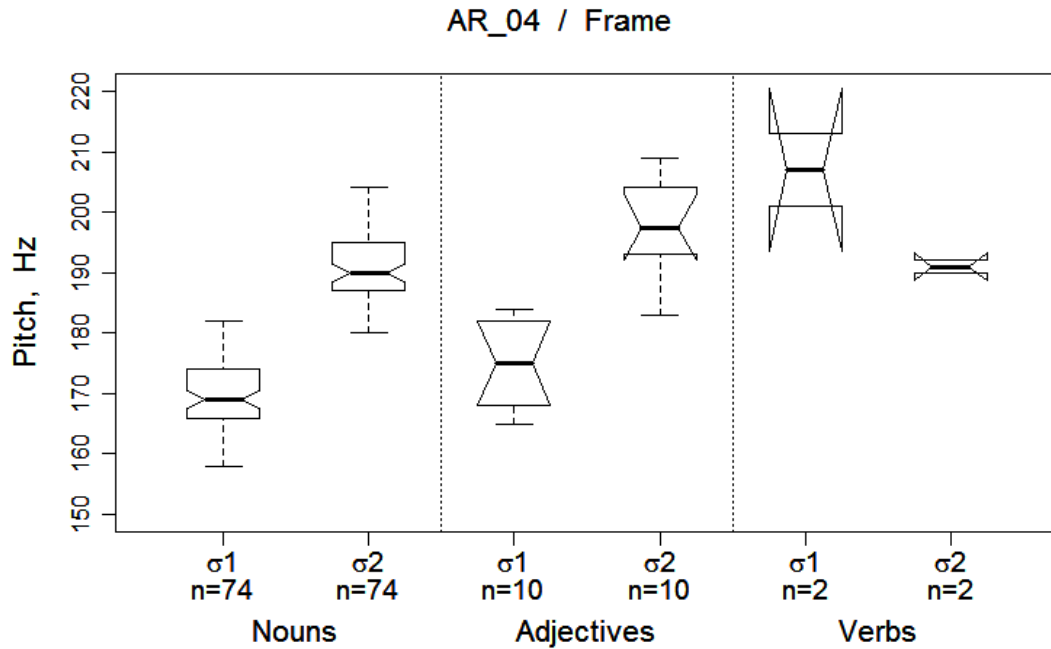
The overall patterns of pitch on σ_1 and σ_2 of isolation forms are illustrated in Figure 5.1 below. There is a great “X” through the verb portion of the plot, since this data is not meaningful.

Figure 5.1 AR_04 / Isolation: Pitch contrasts



Pitch contrasts for the frame forms are illustrated in Figure 5.2 below. Here, pitch clearly distinguishes the σ_2 stress perceived on non-verbs from the σ_1 stress perceived on verbs.

Figure 5.2 AR_04 / Frame: Pitch contrasts



Pitch data for AR_04 nouns, adjectives, and verbs are considered in greater detail in sections 4.2.1.1 through 5.2.1.3.

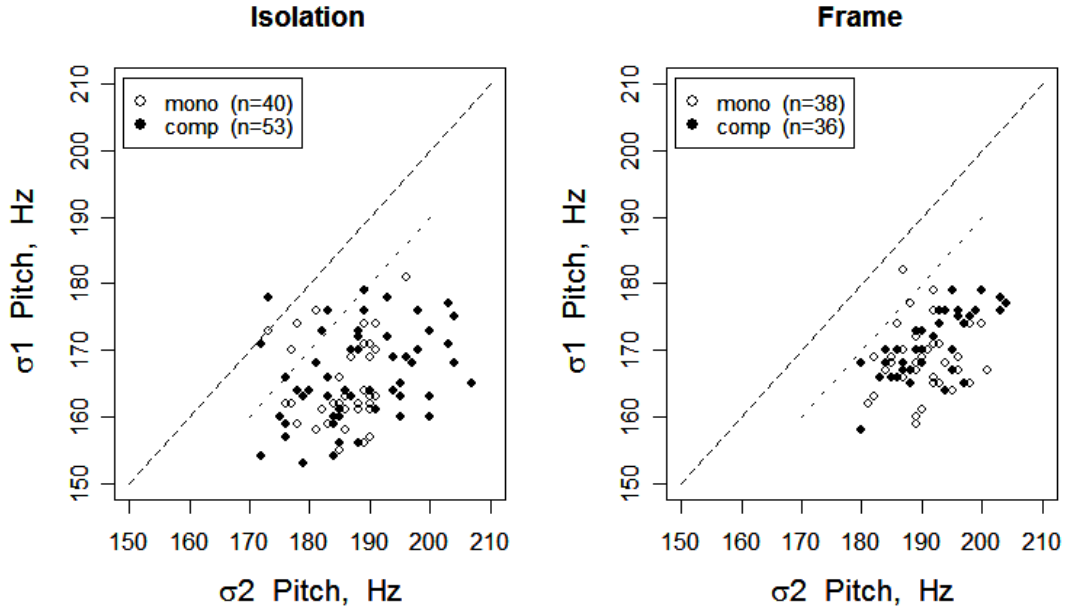
5.2.1.1 AR_04 Nouns

Pitch is a robust cue for stress for AR_04 nouns both in isolation and in the sentence frame. The top part of Figure 4.3 above shows that pitch is higher on σ_2 – the stressed syllable – in all but two or three words. In fact, in most cases, points fall to the right of the dotted “Pitch difference = 10 Hz” reference line. In the frame forms the points are more tightly clustered, reflecting the prosodic constraints imposed when the target word is produced in a fixed sentence. There is no evidence that either isolation or frame forms behave differently depending on morphological structure.

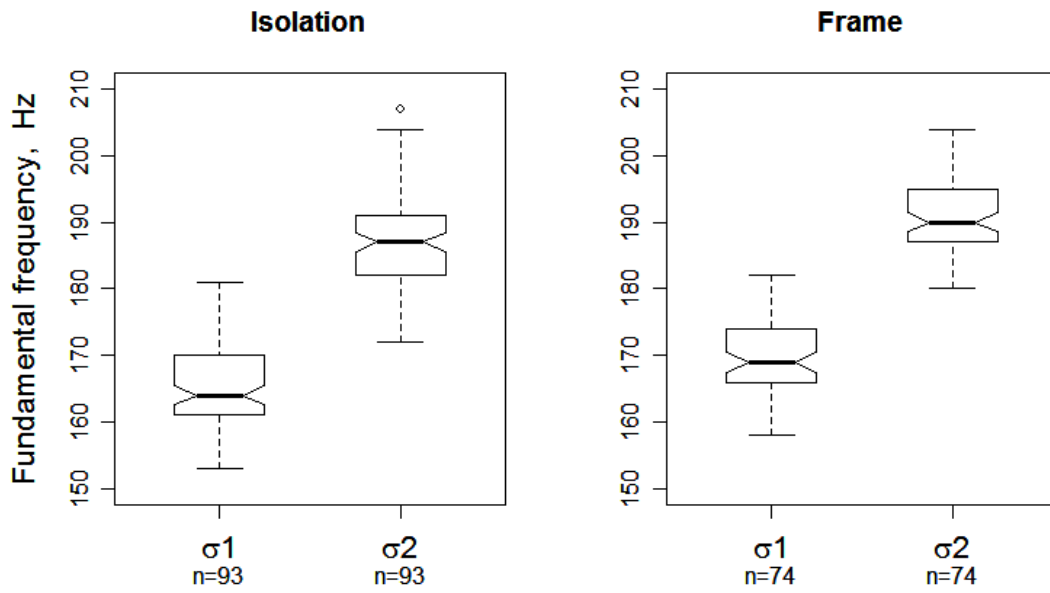
The box-and-whisker plots in the bottom part of the figure show no overlap of either the notches or the interquartile ranges. Thus we can reject the null hypothesis and conclude – with 95% confidence – that there is a significant difference between the median pitch of σ_1 and the median pitch of σ_2 , for nouns in both settings.

Figure 5.3 AR_04 / Nouns: Pitch

AR_04 / Nouns



AR_04 / Nouns



Paired-sample t-tests confirm the statistical significance of the contrast in pitch across syllables. As summarized in Table 5.1 below, p-values are $\ll 0.05$, and the mean difference is a considerable 21 or 22 Hz for both isolation and frame forms. The 95% confidence intervals are fairly tight: if we were to repeatedly record groups of nouns from speaker AR_04 and compare pitch across syllables for each group, there is a 95% probability that the mean difference each time would fall between 20 Hz and 23 Hz.

Table 5.1 AR_04 / Nouns / Pitch: Results of paired-sample t-tests (two-tailed)

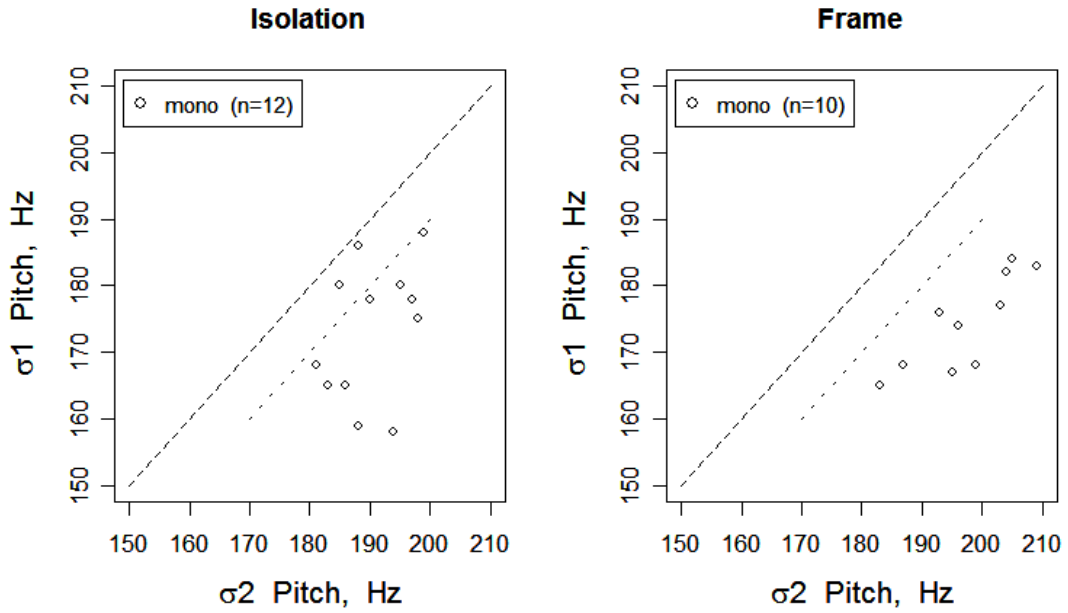
Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
AR_04	isolation	23.9072	92	$< 2.2e-05$	22	20	23
	frame	32.478	73	$< 2.2e-16$	21	20	22

5.2.1.2 AR_04 Adjectives

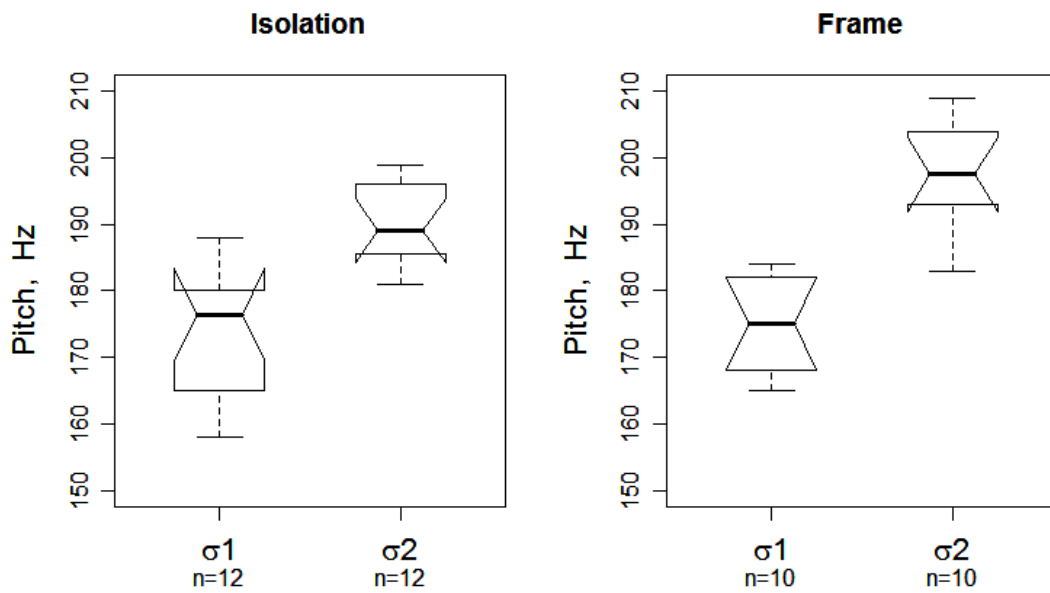
Pitch contrasts across syllables for AR_04 adjectives are illustrated in Figure 5.4 below. Here, too, pitch is a robust and consistent cue for σ_2 stress. In the top part of the figure, all of the adjectives fall to the right side of the dashed line, and most also fall to the right side of the dotted “+10 Hz” reference line. In the bottom part of the figure, the boxplots show no overlap of the notches, indicating that the median values of σ_1 and σ_2 are significantly different at the 95% confidence level. The boxes representing the interquartile ranges for each syllable show no overlap, either.

Figure 5.4 AR_04 / Adjectives: Pitch

AR_04 / Adjectives



AR_04 / Adjectives



Paired-sample t-tests for AR_04 adjectives are summarized in Table 5.2. With p-values $\ll 0.05$, the null hypothesis must be rejected, confirming that the difference in pitch is highly significant. The mean increase across syllables is 17 Hz for isolation forms, and 23 Hz for frame forms.

Table 5.2 AR_04 / Adjectives / Pitch: Results of paired-sample t-tests (two-tailed)

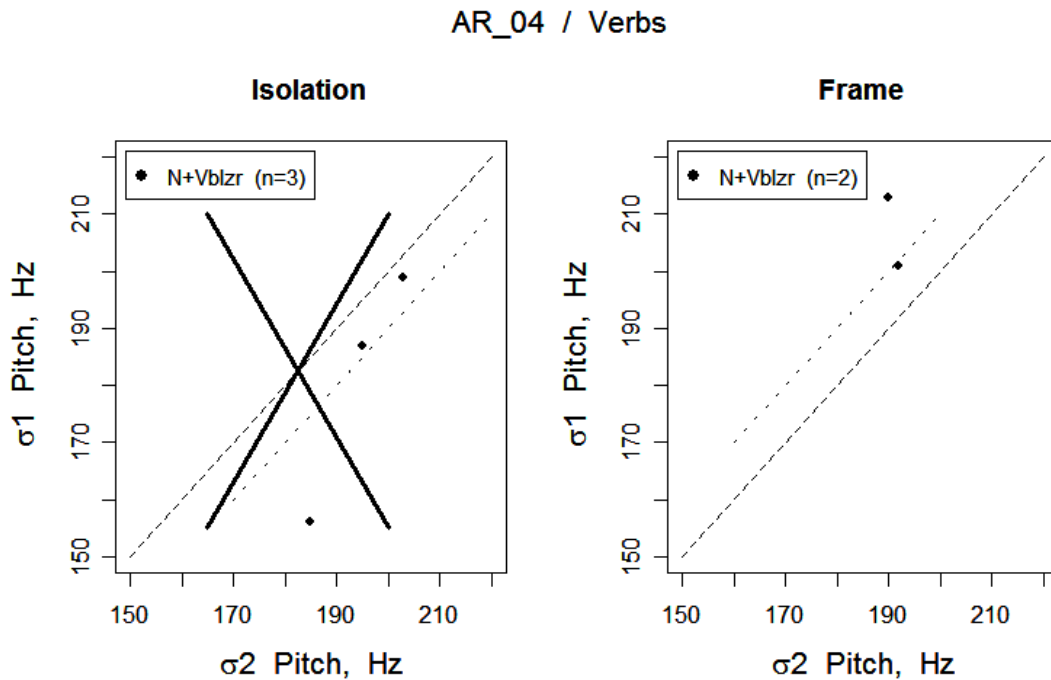
Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
AR_04	isolation	6.1397	11	7.314e-05	17	10.90	23.09
	frame	15.8297	9	7.062e-08	23	15.83	26.29

Thus even though the sample size is small, the pattern is consistent, and the correlation between pitch and σ_2 stress is robust.

5.2.1.3 AR_04 Verbs

Pitch measurements for AR_04 verbs are plotted in Figure 5.5 below. As noted above, this speaker was not at all comfortable producing the verbs in isolation, without the context of tense/aspect/evidentiality. Thus the graph on the left below does not faithfully represent pitch in AR_04 verbs. The two frame forms – which were produced more naturally – both have a higher pitch on σ_1 , the stressed syllable, and so plot to the left of the dashed line below.

Figure 5.5 AR_04 / Verbs: Pitch



With only two tokens, a paired-sample t-test would not be meaningful here.

Nonetheless, the two verbs considered clearly behave differently than non-verbs, and suggest a potential correlation between pitch and stress. It would be useful to consider a larger sample in future.

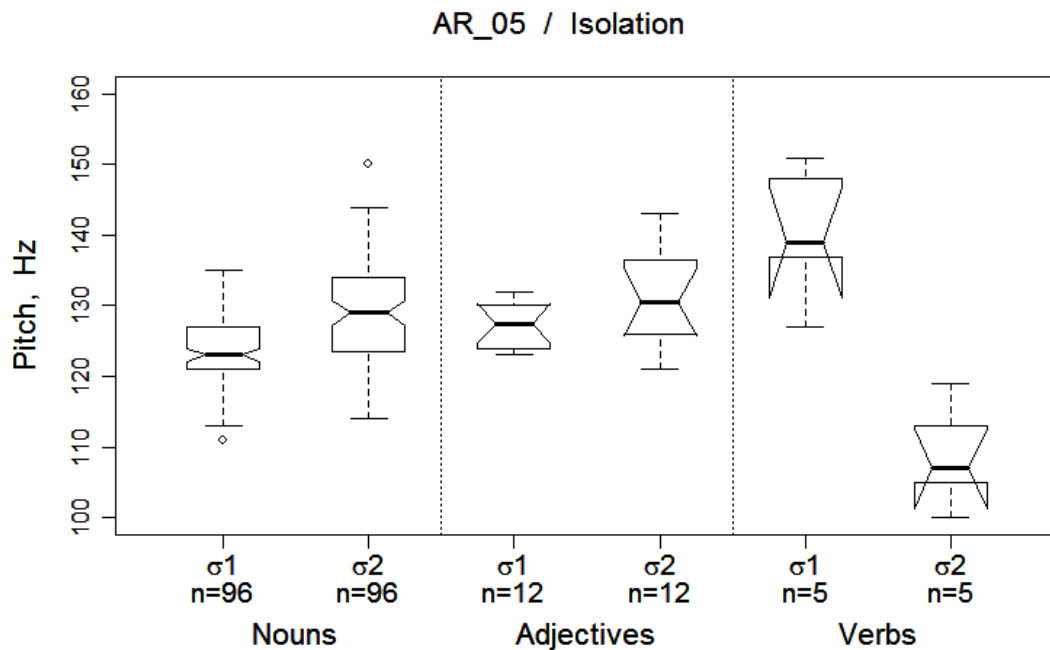
5.2.2 Pitch for speaker AR_05

Pitch has been identified as a correlate of $\sigma 2$ stress in non-verbs and of $\sigma 1$ stress in verbs for both Balti speakers (section 4.2) and for speaker AR_04 (section 5.2.1). For speaker AR_05, however, the correlation between pitch and stress is not so reliable. Specifically, for non-verbs produced in isolation, pitch is not consistently higher on $\sigma 2$. In fact, for nearly 20% of the nouns, pitch is actually higher on $\sigma 1$, the syllable which is *not* stressed. (However, as I demonstrate in sections 5.3.2 and 5.3.3

below, pitch slope is a robust cue for stress in the cases where pitch is not; pitch and pitch slope are complementary reflexes of fundamental frequency in Rebkong Amdo.)

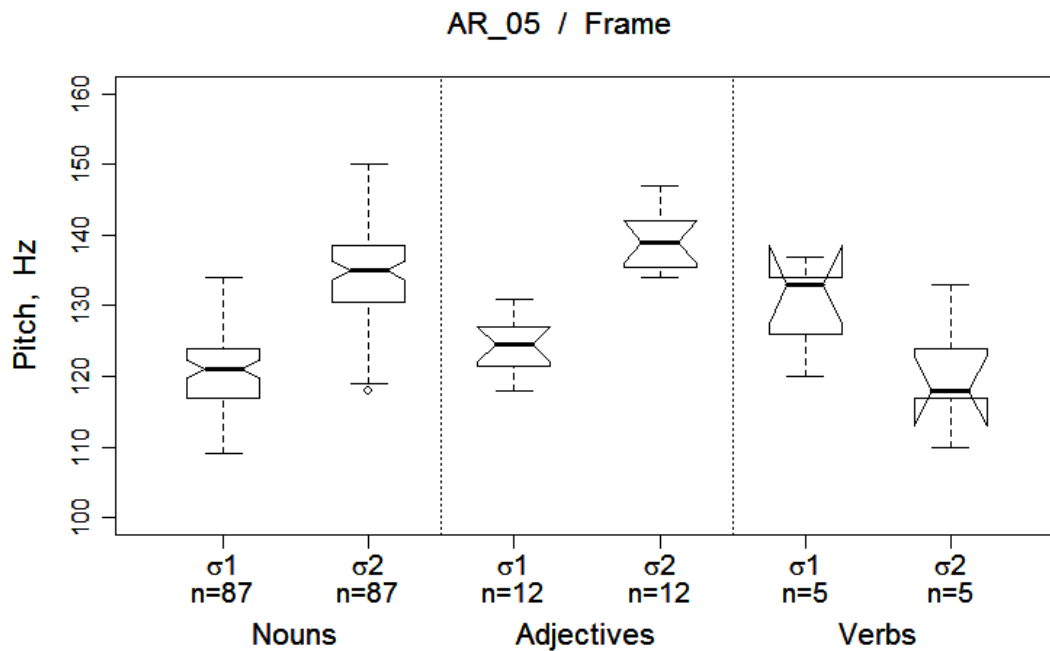
Pitch contrasts for AR_05 isolation forms of nouns, adjectives, and verbs are illustrated in the box-and-whisker plots in Figure 5.6 below. (Numerals are considered in section 5.2.2.3.) As shown, for nouns and adjectives there is considerable overlap in the range of pitch values measured on the two syllables; for adjectives, the notches overlap as well. (This is considerably different than the pattern observed for speaker AR_04, as illustrated in Figure 5.1 on page 330.) For verbs, pitch is clearly a strong cue for the perceived $\sigma 1$ stress.

Figure 5.6 AR_05 / Isolation: Pitch contrasts



For frame forms, the pitch contrasts are much sharper, as illustrated in Figure 5.7 below. The patterns here are reminiscent of those of the Balti speakers and of speaker AR_04, with pitch clearly distinguishing the σ_2 stress perceived on non-verbs from the σ_1 stress perceived on verbs.

Figure 5.7 AR_05 / Frame: Pitch contrasts



Pitch data for AR_05 nouns, adjectives, numerals, and verbs are considered in greater detail in sections 5.2.2.1 through 5.2.2.4.

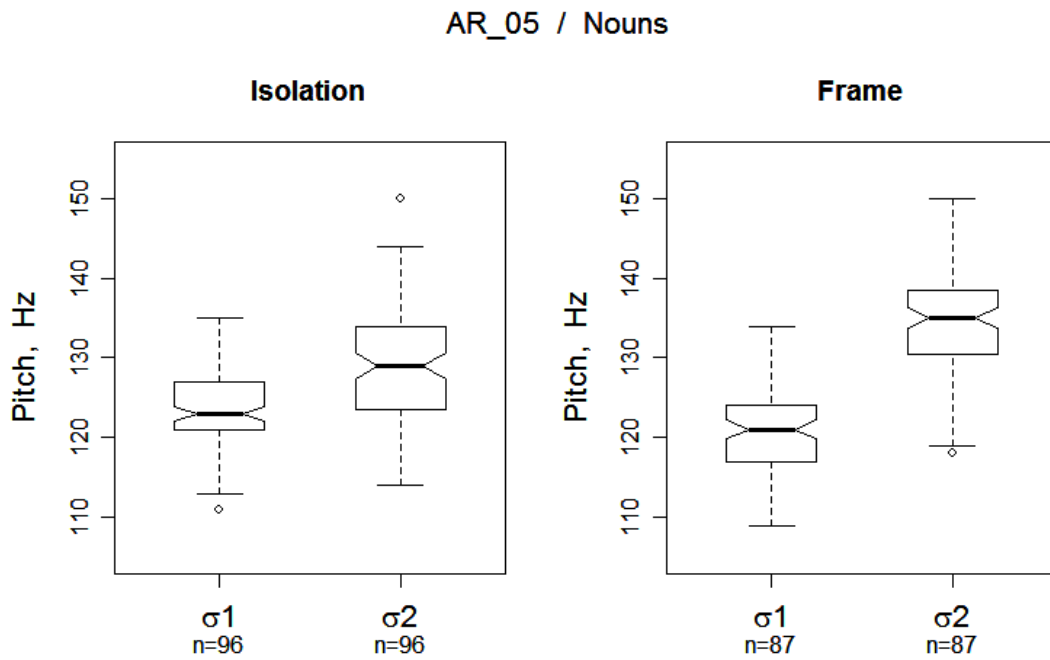
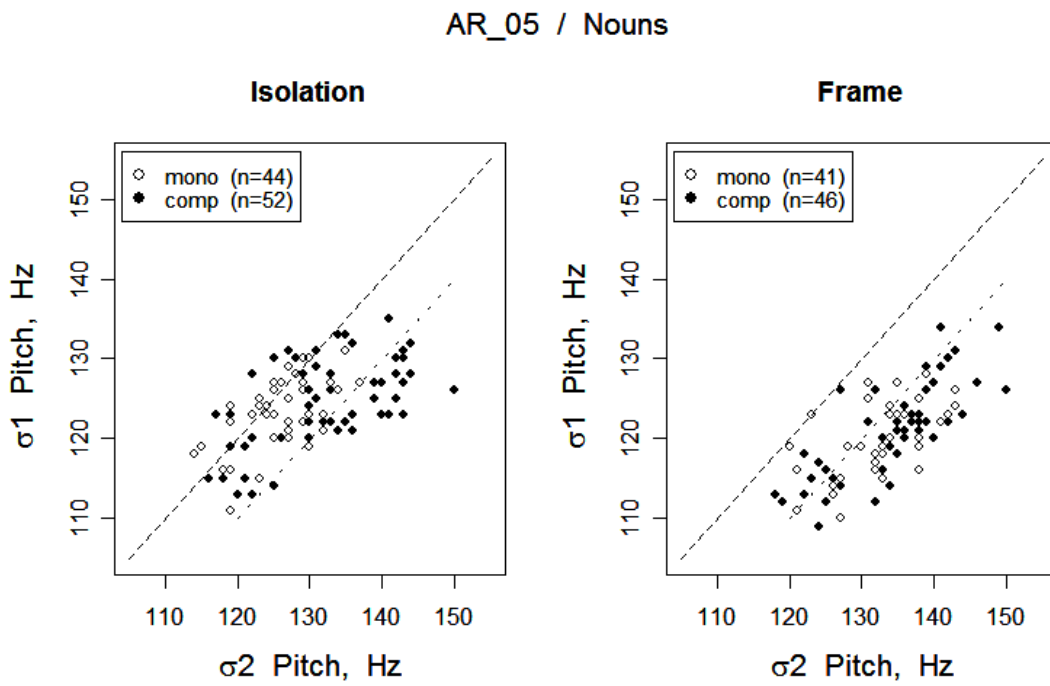
5.2.2.1 AR_05 Nouns

Pitch measurements for AR_05 nouns are plotted in Figure 5.8 below. If pitch were a robust cue for the perceived σ_2 stress pattern, then nearly all nouns would fall

well to the right of the dashed “0” line. For the isolation forms, this is clearly not the case. Many nouns plot to the left of the dashed line, indicating a higher pitch on σ_1 , the unstressed syllable. And even among those nouns which *do* have a higher pitch on σ_2 , the contrast is $< +10$ Hz for all but a handful of compound forms. In fact, there is a distinct rightward “bulge” of isolation forms in which σ_2 pitch is greater than ~ 135 Hz; all of these forms are compounds.

For the frame forms, pitch is indeed higher on σ_2 in almost all cases, though the contrast across syllables is often $< +10$ Hz. Here, too, many of the nouns with a high pitch on σ_2 are compounds.

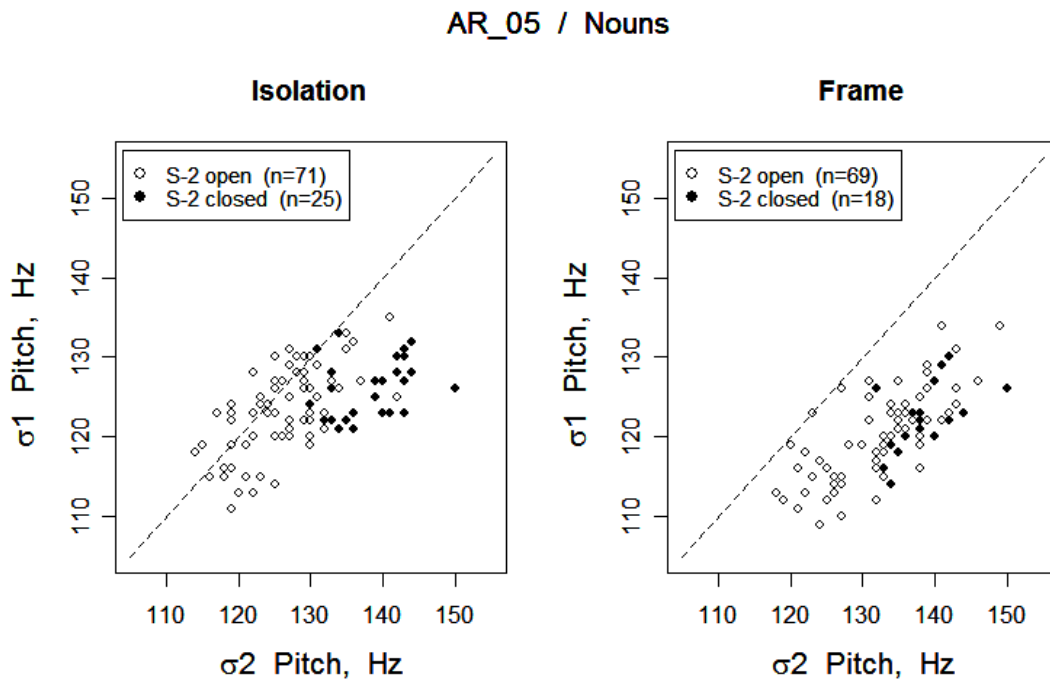
Figure 5.8 AR_05 / Nouns: Pitch



For both isolation and frame forms, the observed pitch pattern is influenced by two factors: (a) a tendency for pitch to be higher on σ_2 when σ_2 is closed; and (b) the intrinsic variation of pitch with vowel height.

The effect of σ_2 closure is illustrated by the plots in Figure 5.9 below.

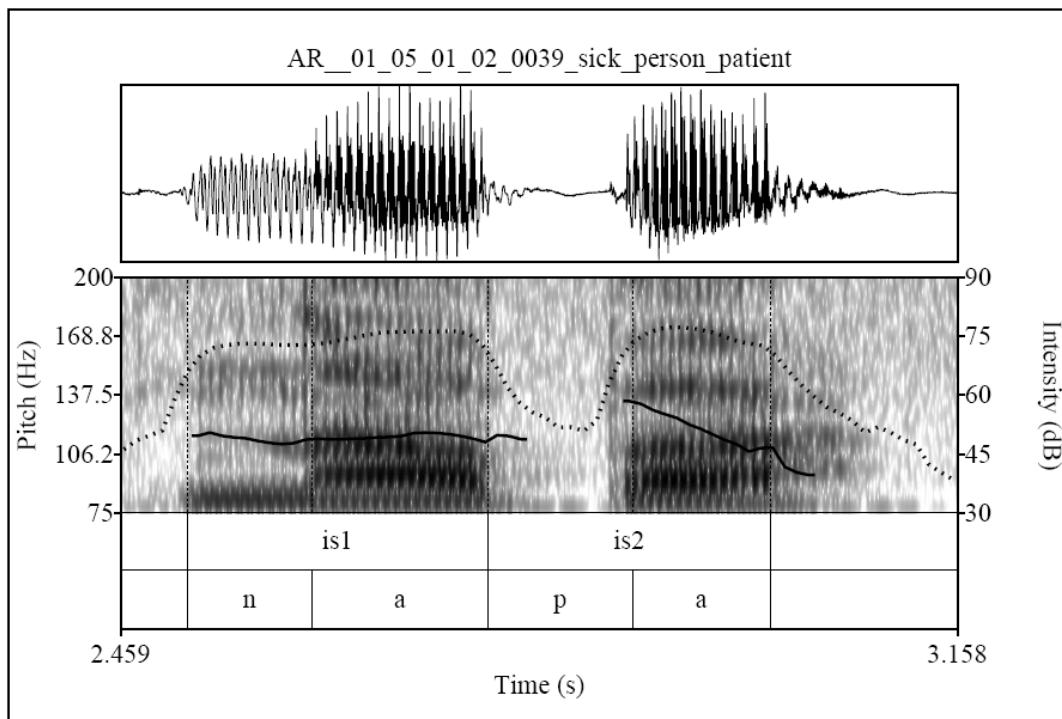
Figure 5.9 AR_05 / Nouns: Pitch vs. σ_2 closure



It is quite clear here that pitch is always higher than 130 Hz on σ_2 whenever σ_2 is closed. (The template of σ_1 is not relevant: about half of these filled circles have an [open.closed] structure and about half have a [closed.closed] structure, and they appear to be randomly distributed.) This correspondence between pitch and syllable closure accounts for the rightward “bulge” in the distribution of isolation forms in Figure 5.8. In monomorphemic nouns σ_2 is open, since it is generally a variant of [-pa] or [-ma]. Thus all of the words with a closed σ_2 are compounds.

The reason pitch on the σ_2 vowel is relatively high in these words has to do with the fact that pitch slope is an important correlate of σ_2 stress in Amdo nouns (as discussed in detail in section 5.3.) When σ_2 is open, pitch leaps to an early peak on the vowel, and then declines through to the end of the word. This is illustrated in the pitch trace for [na.'pa] *nad.pa* ‘sick person, patient’ in Figure 5.10 below⁴³. Pitch for the vowels in σ_1 and σ_2 was measured (over the medial 50%) as 116 Hz and 120 Hz, respectively, so the increase in mean pitch across syllables is a mere 4 Hz.

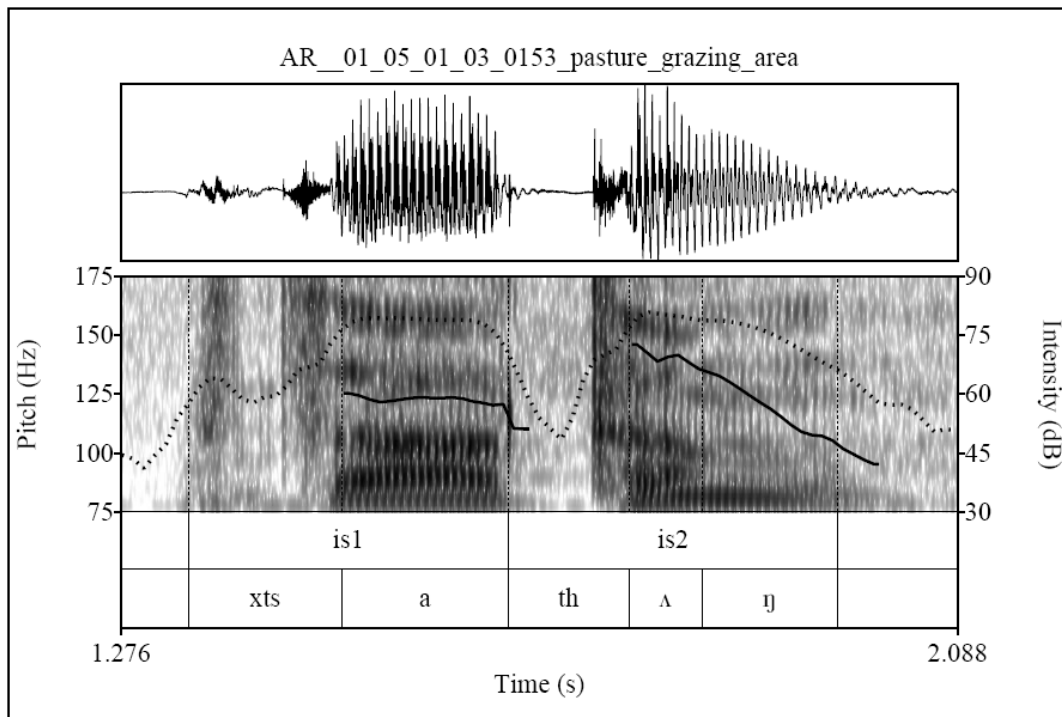
Figure 5.10 AR_05 / Nouns / Isolation / Pitch trace: [na.'pa] *nad.pa* ‘sick person, patient’



⁴³ In this and other figures, the solid line shows pitch, and the dashed line shows intensity.

On the other hand, when σ_2 is closed, pitch slope is borne not just by the vowel, but by the entire rhyme. As illustrated for the word [xtsa.'tʰʌŋ] *rtswa.thang* ‘pasture, grazing area’ in Figure 5.11 below, when so much of the decline is carried by the coda, the vowel itself carries a relatively high mean pitch.

Figure 5.11 AR_05 / Nouns / Isolation / Pitch trace: [xtsa.'tʰʌŋ] *rtswa.thang* ‘pasture, grazing area’

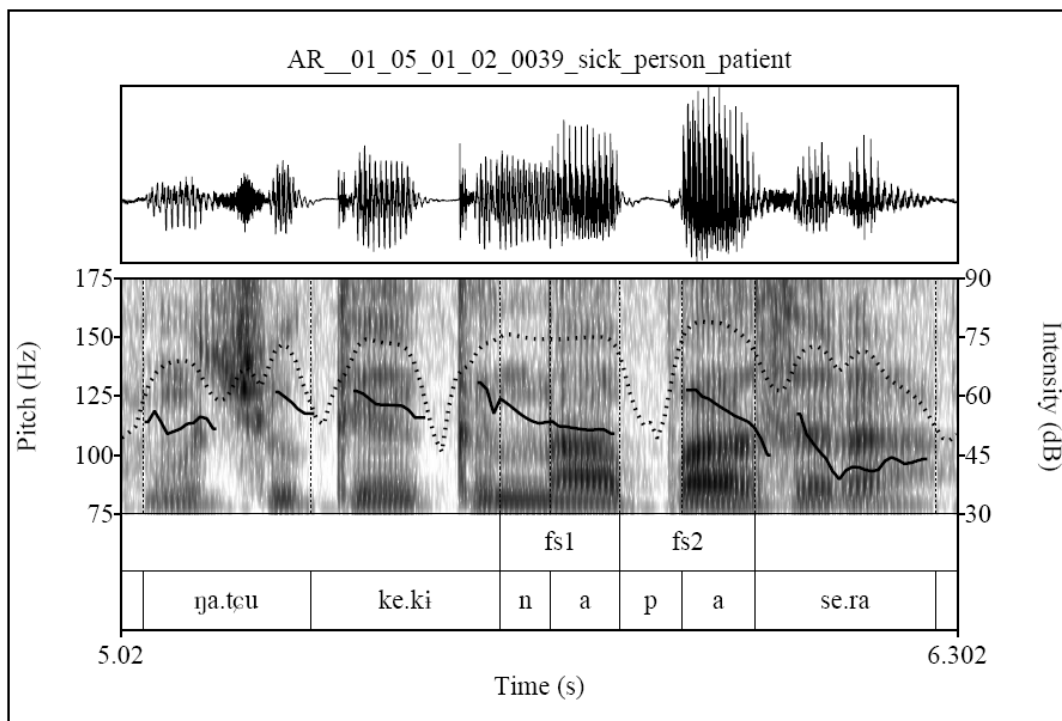


In this case, when the pitch measured for σ_1 (123 Hz) is compared to the pitch measured over the full rhyme of σ_2 (126 Hz), the increase across syllables is a trivial 3 Hz, comparable to the 4 Hz difference calculated for [na.'pa] above. But when the

pitch measured for σ_1 is compared to the pitch measured for only the vowel of σ_2 (140 Hz), the increase across syllables is an exaggerated 17 Hz.

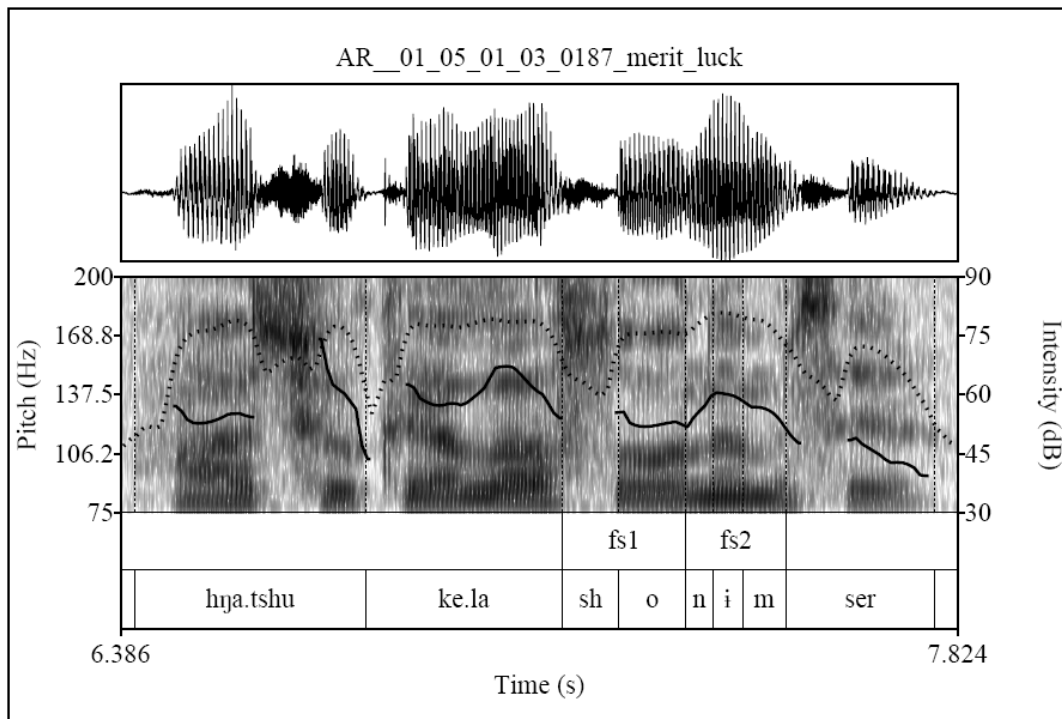
This same effect occurs in the frame forms, as illustrated by comparing the word [na.'pa] *nad.pa* ‘sick person, patient’, where σ_2 is open (Figure 5.12) ...

Figure 5.12 AR_05 / Nouns / Frame / Pitch trace: [na.'pa] *nad.pa* ‘sick person, patient’



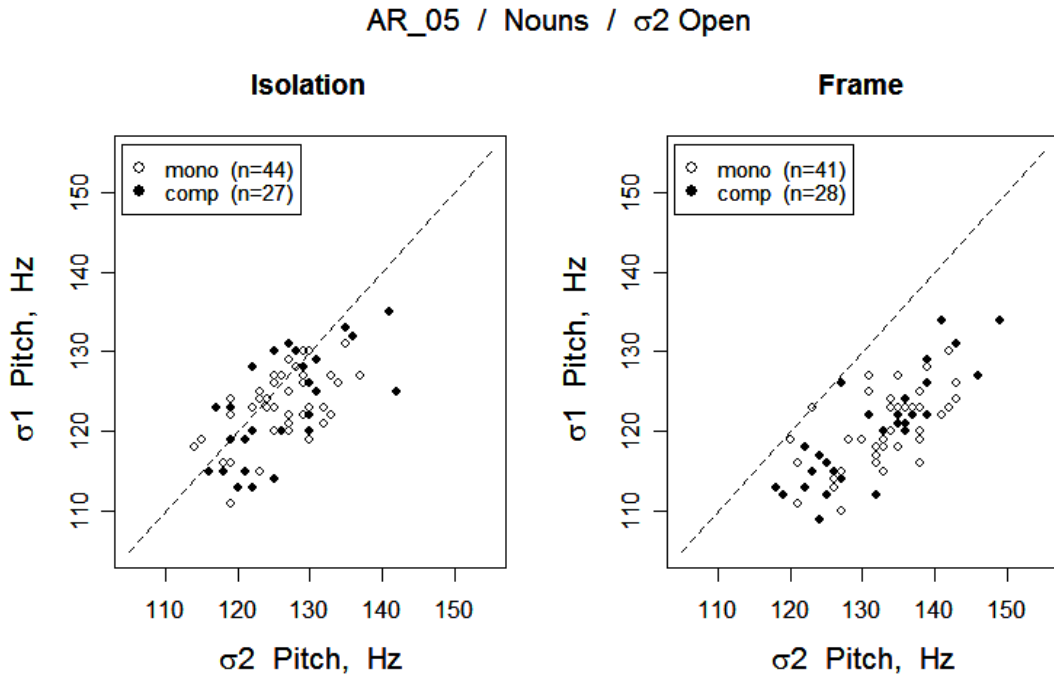
... with the word [s^ho.'nim] *bsod.nams* ‘merit, luck’, where σ_2 is closed (Figure 5.13).

Figure 5.13 AR_05 / Nouns / Frame / Pitch trace: [s^ho.'nim] *bsod.nams* ‘merit, luck’



All nouns with a closed σ_2 are influenced by this factor, because all of the codas are pitch-bearing voiced continuants. What is clear from this is that it would be preferable, in future studies, to measure pitch (and pitch slope) for the entire rhyme in such cases, rather than for just the nuclear vowel in each syllable. For the present, though, this factor can be controlled for by focusing the analysis on those nouns in which σ_2 is open. These are plotted in Figure 5.14 below.

Figure 5.14 AR_05/ Nouns / $\sigma 2$ Open: Pitch



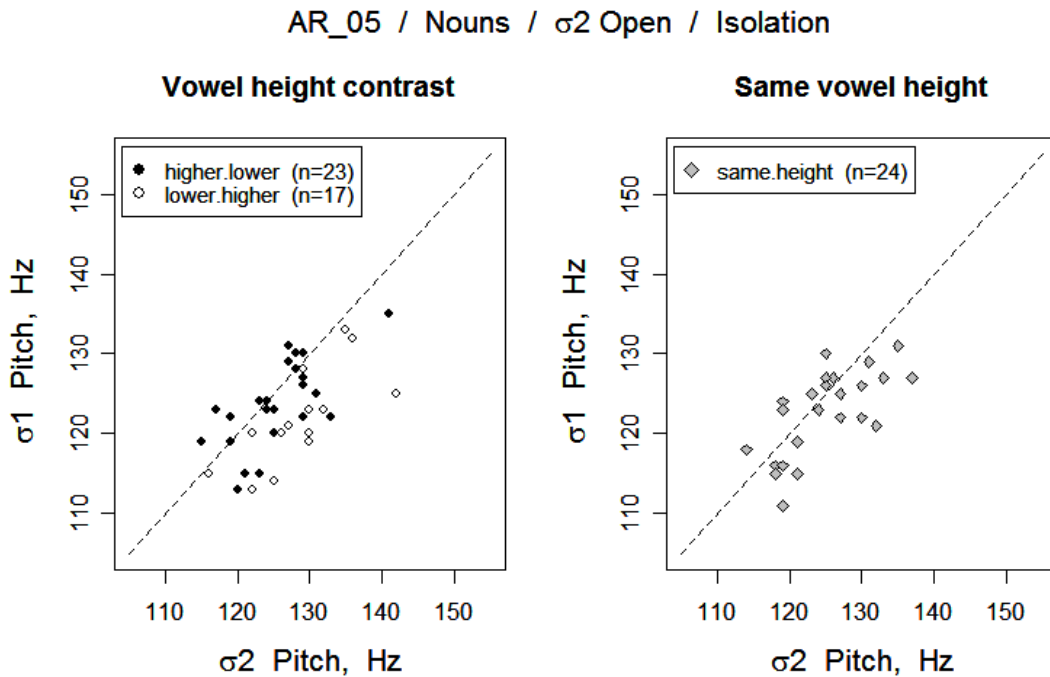
Having thus controlled for the incidental effects of $\sigma 2$ closure, the potential influence of vowel height contrast can be considered.

Isolation forms of nouns with $\sigma 2$ open are plotted in terms of vowel height in Figure 5.15 below. (The seven ‘ $\sigma 2$ -open’ words which have a diphthong in $\sigma 1$ are excluded from these plots.) When there is a contrast in vowel height across syllables, there is some tendency for pitch to be higher on whichever syllable has the higher vowel, as one might predict following Lehiste. As shown, all of the [lower.higher] nouns – predicted to have a higher pitch on $\sigma 2$ – indeed do fall to the right of the dashed line. Furthermore, the *only* nouns which have a higher pitch on $\sigma 1$ – falling to the left of the dashed line – are those which have a higher vowel in that syllable, the [higher.lower] nouns. However, other [higher.lower] nouns have a higher pitch on $\sigma 2$,

which may reflect some degree of correlation with stress. For this subset, one could say that stress “outranks” vowel height contrast as the factor governing pitch.

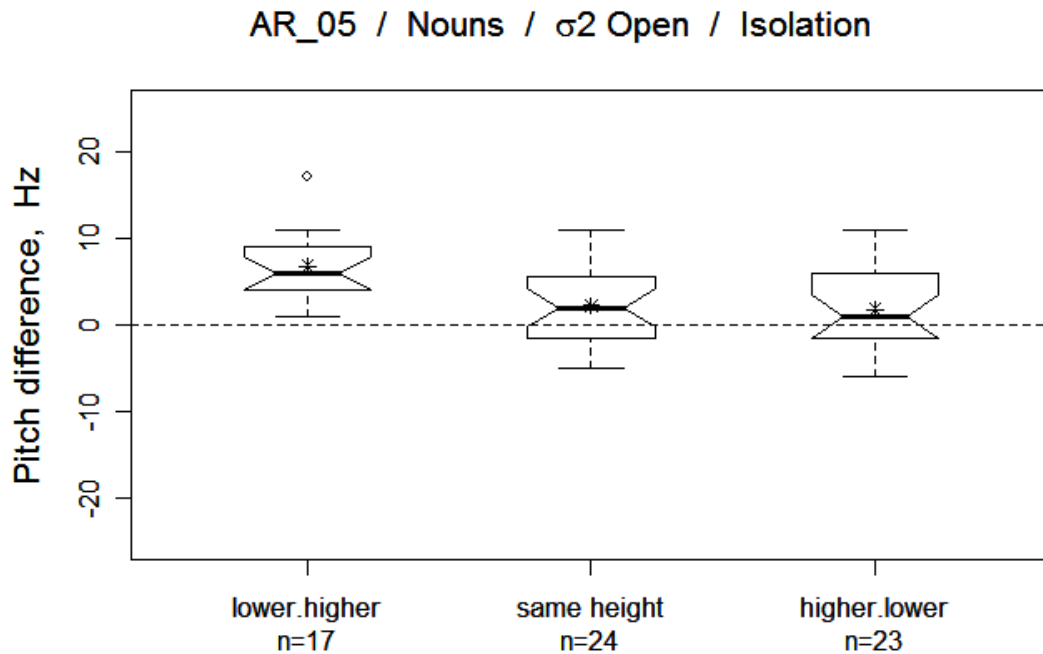
Nouns with vowels of the same height in both syllables fall to either side of the dashed line, though skewed to the right.

Figure 5.15 AR_05 / Nouns / σ 2 Open / Isolation: Pitch vs. vowel height



The magnitude of the pitch contrast across syllables is thus somewhat influenced by the contrast in vowel height; a steady trend is illustrated in the box-and-whisker plot in Figure 5.16 below. For the control group – those with vowels of the same height in both syllables – the notch representing the 95% confidence interval about the median skims the dashed “0” line.

Figure 5.16 AR_05 / Nouns / $\sigma 2$ Open / Isolation: Pitch differences vs. vowel height



Paired-sample t-tests, summarized in Table 5.3 below, indicate that the increase in pitch across syllables is barely statistically significant for the control group, with $p = 0.0278$. The difference is certainly not perceptually salient, with a mean difference of only 2.2 Hz, and a lower 95% confidence limit of 0.2 Hz. This reflects the fact that, in about one third of the nouns in the sample population, pitch is actually higher on $\sigma 1$, the unstressed syllable.

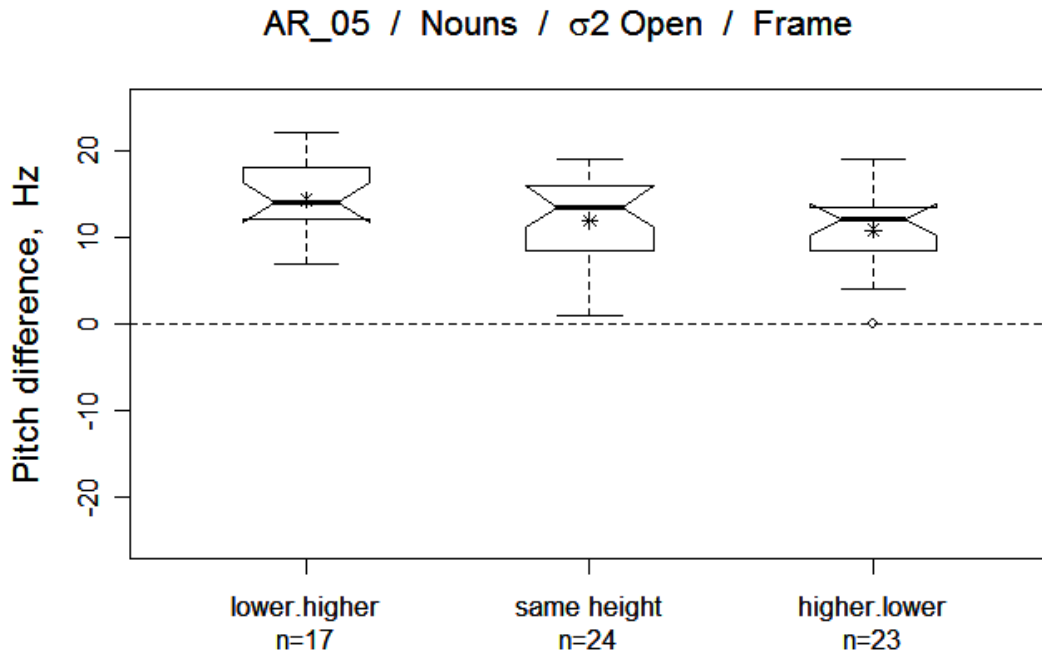
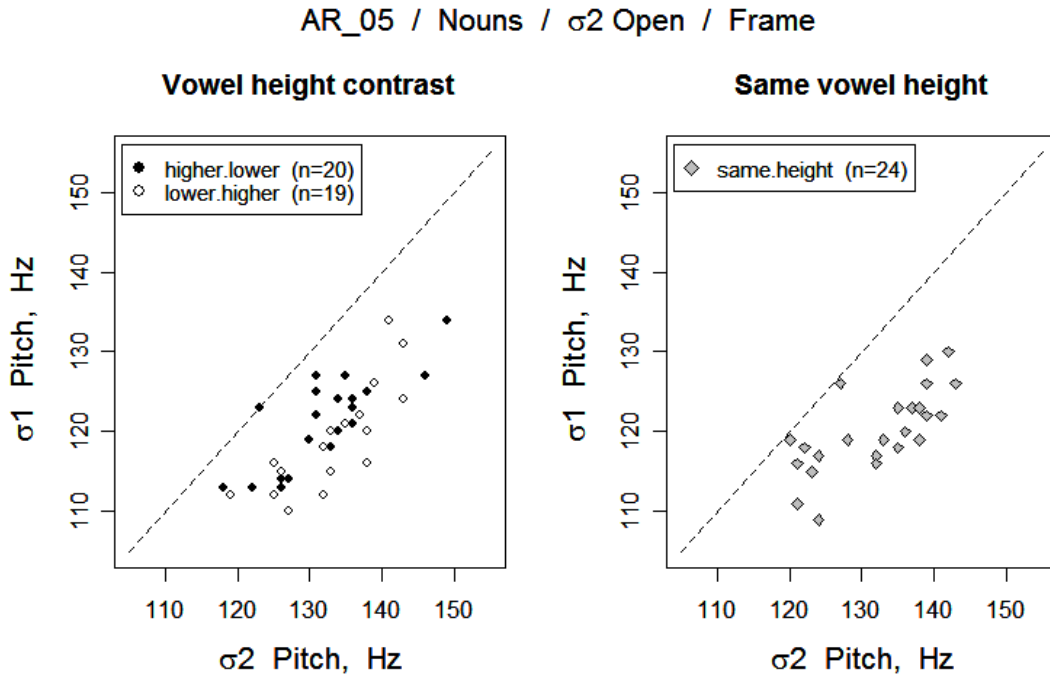
Table 5.3 AR_05 / Nouns / σ 2 Open / Isolation: Pitch by vowel height contrast: paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
					lower	upper
lower.higher	6.6312	16	5.772e-06	6.8	4.6	8.9
same height	2.3488	23	0.0278	2.2	0.2	4.1
higher.lower	1.8947	22	0.07135	1.8	-0.17	3.7

In conclusion, once the incidental effects of syllable closure and vowel height have been controlled for, we can see that pitch shows a statistically significant correlation with σ 2 stress, but that the pitch contrast across syllables cannot be considered perceptually significant. In fact, a statistically meaningful contrast may be smaller than the just-noticeable difference of 1 Hz (Lehiste 1970). Therefore pitch cannot be considered a correlate of stress for AR_05 isolation nouns.

The frame forms with σ 2 open manifest a considerably different pattern. (See also Figure 5.8, page 341.) As illustrated in Figure 5.17 below, pitch is only weakly influenced by the contrast in vowel height across syllables – with some distinction between the distributions of [lower.higher] and [higher.lower] groups – but all nouns (with one exception) still fall to the right of the dashed line, with a higher pitch on σ 2.

Figure 5.17 AR_05 / Nouns / $\sigma 2$ Open / Frame: Pitch vs. vowel height contrast



Paired-sample t-tests for the frame forms, summarized in Table 5.4 below, indicate that the difference in pitch across syllables is highly significant for all height groups, with $p \ll 0.05$. When the full set of frame forms with an open σ_2 is considered, the mean difference in pitch is 12 Hz, and the 95% confidence limit is fairly narrow.

Table 5.4 AR_05 / Nouns / σ_2 Open / Frame: Pitch by vowel height contrast: paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
					lower	upper
lower.higher	14.2869	17	6.69e-11	14	12	16
same height	10.9494	23	1.344e-10	12	9.7	14
higher.lower	10.6991	19	1.753e-09	11	8.7	13
all	21.642	68	< 2.2e-16	12	11	14

Thus for nouns produced in the sentence frame – unlike those produced in isolation – the contrast between the mean pitch on σ_1 and the mean pitch on σ_2 is both statistically and perceptually significant, indicating a correlation between pitch and perceived σ_2 stress. Contrasts in σ_2 closure have an incidental effect on pitch measurements, while contrasts in vowel height do not.

5.2.2.2 AR_05 Adjectives

The pitch patterns for AR_05 adjectives are similar to those described above for nouns. As illustrated in Figure 5.18 below, for the frame forms, the correlation between pitch and stress is quite clear. All points indicate a higher pitch on σ_2 , and even plot to the right of the dotted “+10 Hz” reference line. A paired-sample t-test,

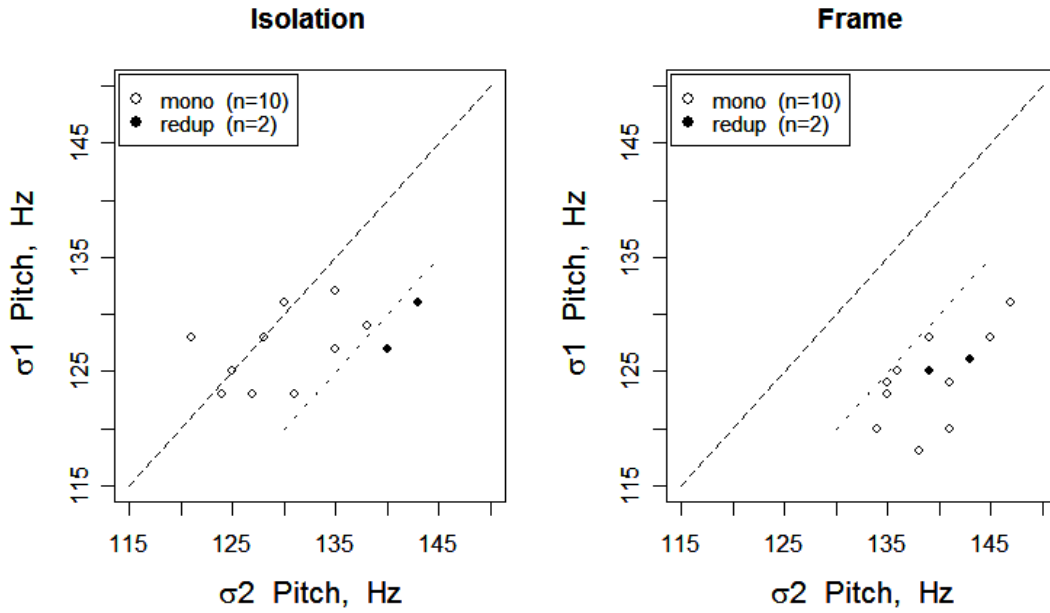
summarized in Table 5.5 below, shows that the difference across syllables is both statistically and perceptually significant ($p \ll 0.05$; mean difference = 15 Hz).

Table 5.5 AR_05 / Adjectives / Frame / Pitch: Results of paired-sample t-tests (two-tailed)

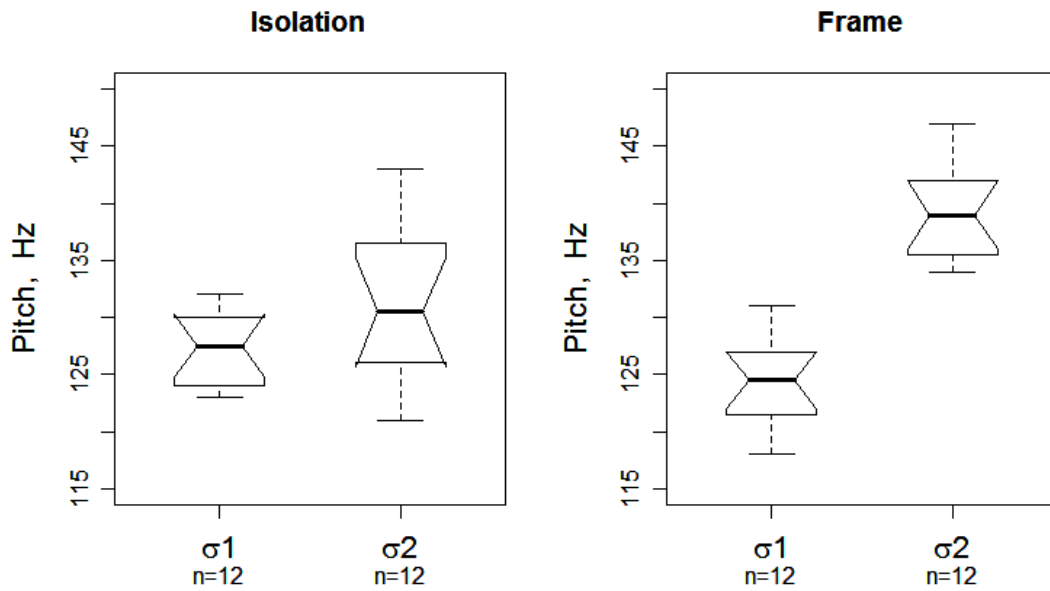
Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
AR_05	frame	15.0312	11	1.115e-08	15	12.87	17.29

Figure 5.18 AR_05 / Adjectives: Pitch

AR_05 / Adjectives

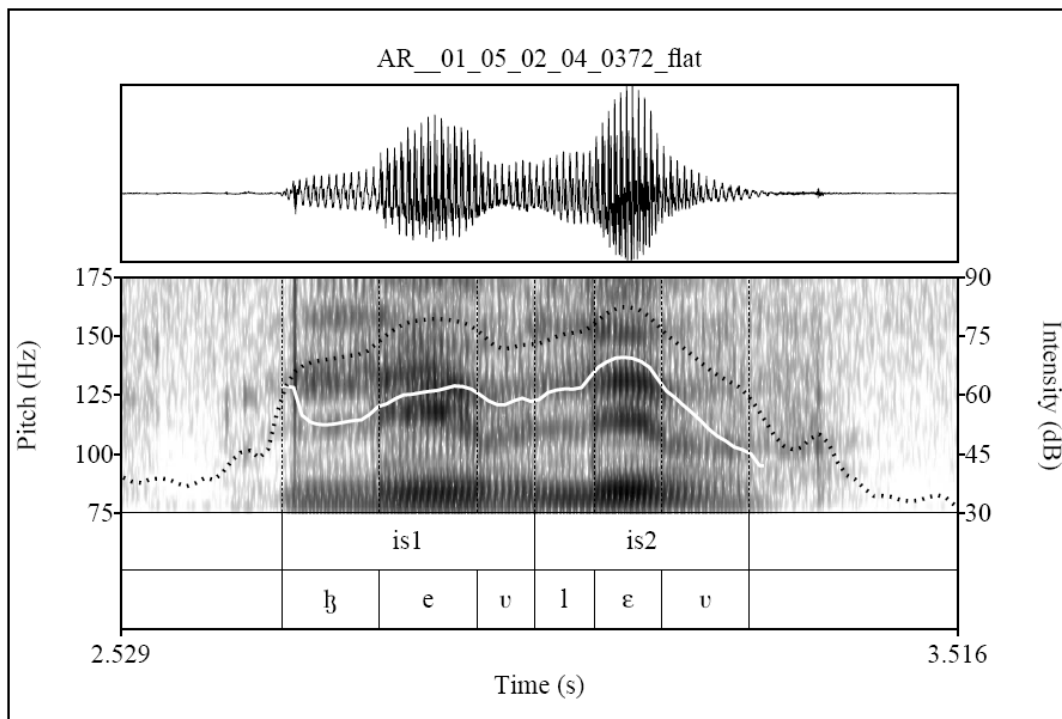


AR_05 / Adjectives



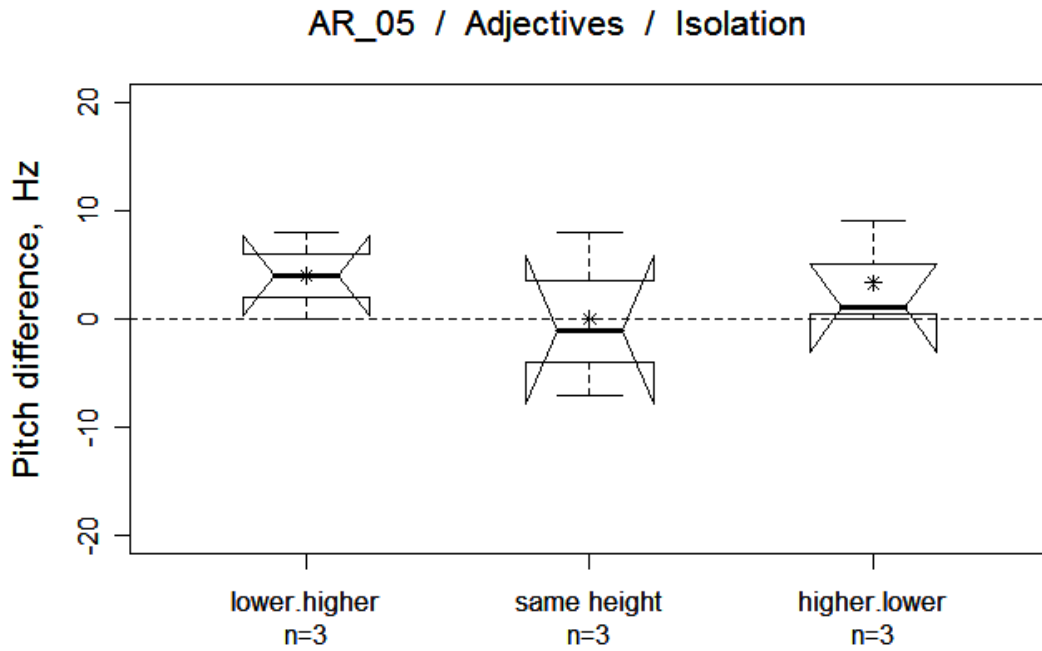
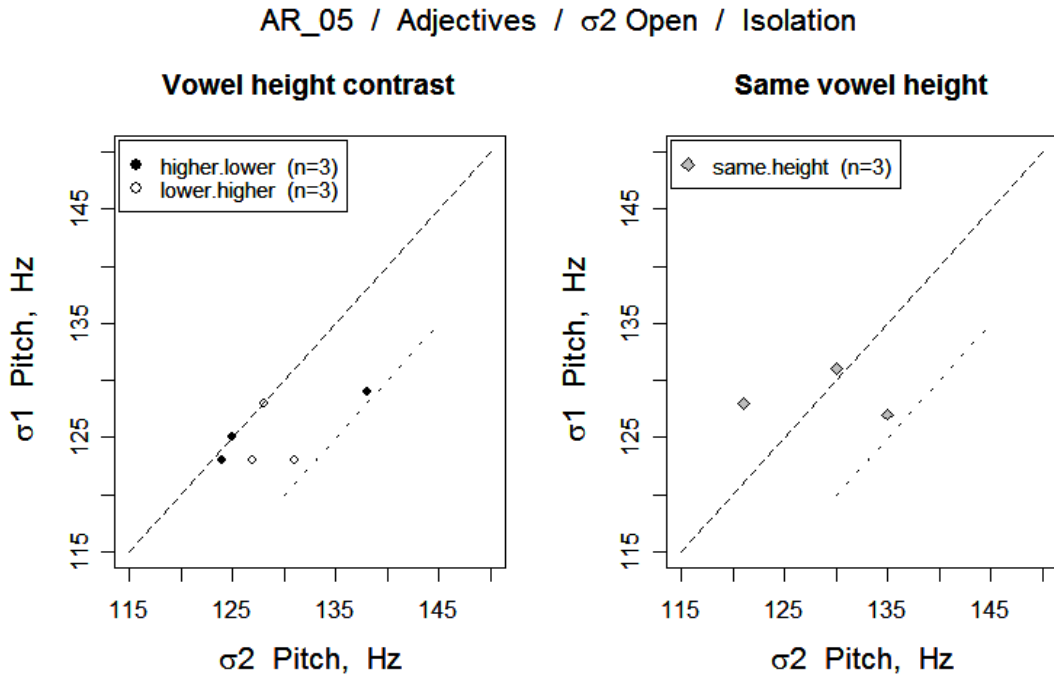
For the isolation forms, there is no such clear correlation. Four of the twelve tokens show no difference in pitch across syllables, and a fifth – [no.'χo] *nag.po* ‘black’ – has a distinctly higher pitch on $\sigma 1$. The two reduplicated forms which plot to the right of the dotted “+10 Hz” reference line – [ʒev.'lɛv] *leb.leb* ‘flat’ and [tɕʰuŋ.'tɕʰuŋ] *chung.chung* ‘small’ – are also the only two words with a closed $\sigma 2$. As was the case with nouns, since much of the pitch slope is carried by the coda when $\sigma 2$ is closed, the pitch measured over the medial 50% of the vowel is exaggerated. This is illustrated in Figure 5.19 below.

Figure 5.19 AR_05 / Adjectives / Isolation / Pitch trace: [ʒev.'lɛv] *leb.leb* ‘flat’



Excluding these two cases with a closed σ_2 – and also the one case with a diphthong in σ_1 – the remaining adjectives are plotted in terms of vowel height contrast in Figure 5.20 below. The plot is too sparse to draw any firm conclusions about the potential influence of vowel height contrast on pitch. In the boxplot at the bottom of the figure, the notch representing the control group intersects the “0” line, indicating that the median pitch difference is not significantly different than zero. In fact, the notches for the three height groups overlap one another; it is not clear that there is any real basis for distinguishing between them.

Figure 5.20 AR_05 / Adjectives / σ 2 Open / Isolation: Pitch vs. vowel height contrast



A paired-sample t-test for the set of nine σ_2 -open adjectives is summarized in Table 5.6 below. Here, $p > 0.05$ and the 95% confidence interval includes zero. This means that the null hypothesis – that there is no significant difference in pitch across syllables – cannot be rejected.

Table 5.6 AR_05 / Adjectives / σ_2 Open / Isolation / Pitch: Results of paired-sample t-tests (two-tailed)

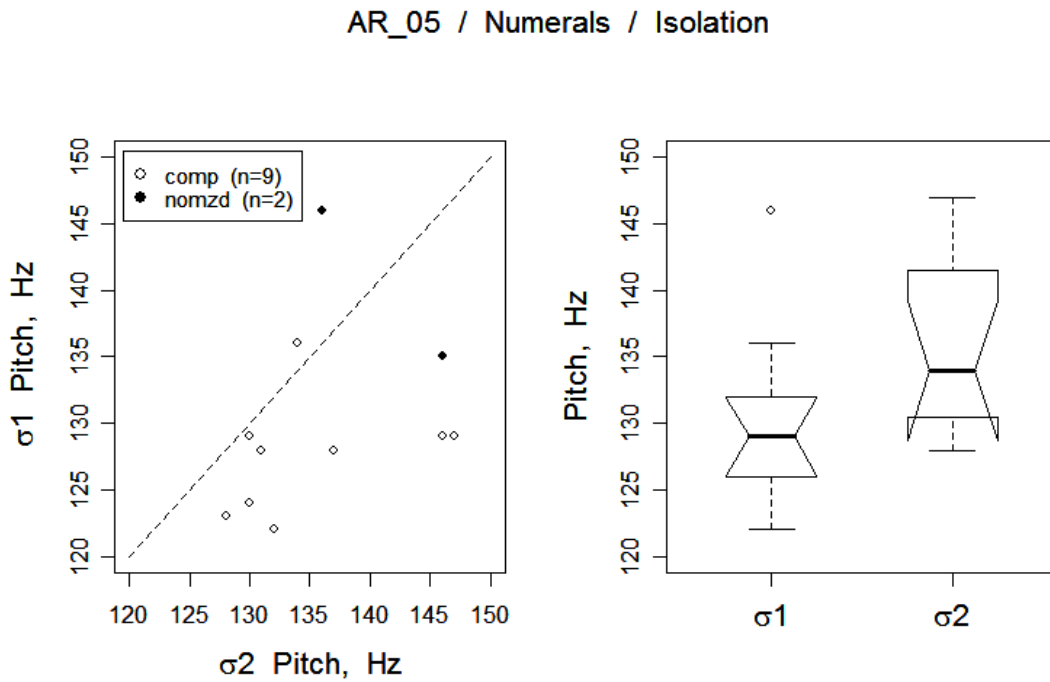
Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
AR_05	isolation	1.3914	8	0.2016	2.4	-1.6	6.5

In conclusion, for AR_05 adjectives, pitch is a robust cue for stress in the frame forms, but is not a correlate of stress at all for the isolation forms. Nonetheless, as will be shown in section 5.3.2.2, pitch slope is a robust correlate of stress in exactly those cases in which pitch is not.

5.2.2.3 AR_05 Numerals

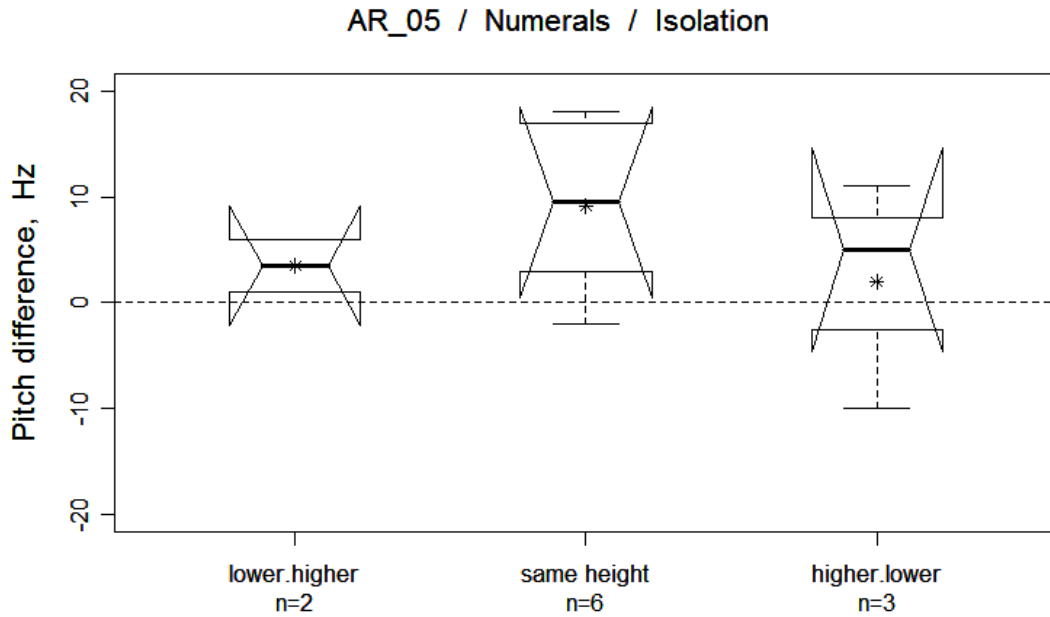
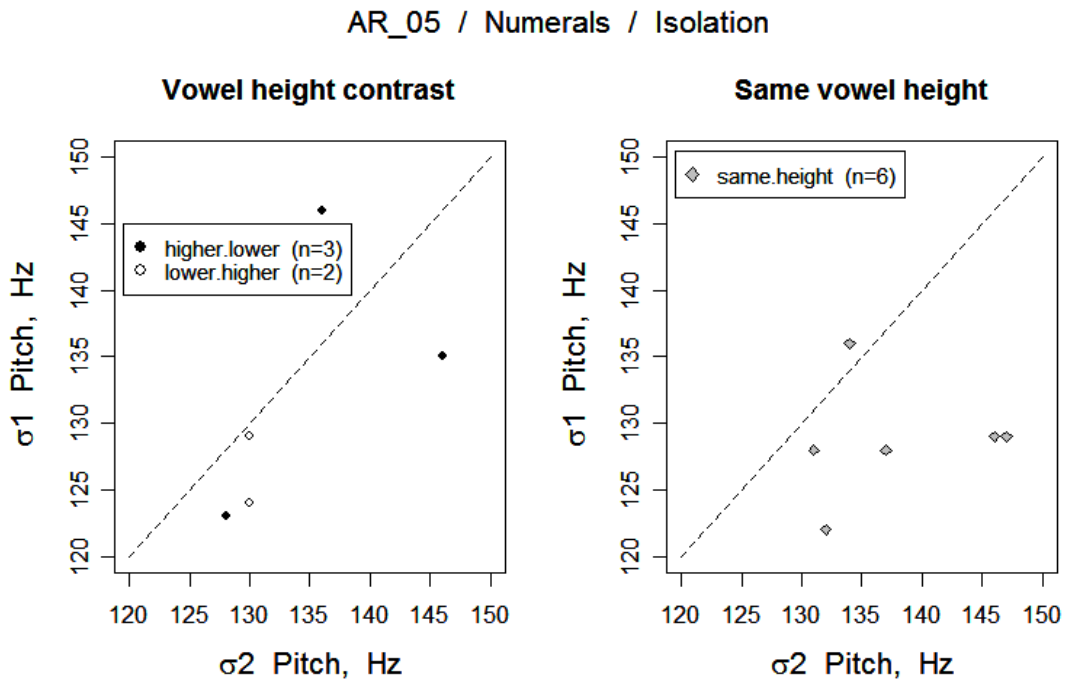
Pitch contrasts for all the numerals which could be measured are plotted in Figure 5.21 below. If pitch were a robust and reliable correlate of stress, all points would fall well to the right of the dashed line. As shown, most of the numerals do, but some are quite close to the dashed line – indicating minimal contrast across syllables – and two even fall to the left of the line – including the nominalized form [ksum.'ba] *gsum.pa* ‘third’ at (σ_2 , σ_1) coordinates (136,146). The median values for the two syllables are not significantly different from one another, as illustrated by the overlap of notches in the boxplot to the right below.

Figure 5.21 AR_05 / Numerals: Pitch



With such a small sample, it is particularly important to determine if the distribution is skewed by incidental factors, such as a contrast in vowel height across syllables. As shown in Figure 5.22, however, the pattern seems to be dominated by a correlation between pitch and stress, rather than by a correlation between pitch and vowel height. Several of the [higher.lower] numerals have a higher pitch on σ_2 – plotting to the right of the dashed line – even though the vowel height contrast would predict otherwise. And the [same.height] numerals fall mostly to the right, rather than being evenly distributed to either side of the dashed line.

Figure 5.22 AR_05/ Numerals / Isolation: Pitch vs. vowel height contrast



As shown in the box-and-whisker plot above, the notches for the [lower.higher] and [higher.lower] groups span the “0” line, indicating that the median difference in pitch across syllables is not significantly different than zero. Indeed, paired-sample t-tests for these subsets – summarized in Table 5.7 below – yield very large 95% confidence intervals, which is not surprising given the small sample sizes. The contrast in pitch across syllables is not statistically significant ($p > 0.05$) for either of these groups.

Table 5.7 AR_05 / Numerals / Pitch by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowel ht.	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.4	1	0.3949	3.5	-28	35
same height	2.8852	5	0.03438	9.2	1.0	17
higher.lower	0.3203	2	0.7791	2.0	-25	29
all	2.5055	10	0.03115	6.2	0.68	11

For the control group, a paired-sample t-test shows that the difference in pitch across syllables is statistically significant ($p < 0.05$), with a mean of 9.2 Hz. However, as indicated by the lower 95% confidence interval, a statistically significant mean difference may be as small as 1.0 Hz, which is just at the threshold of perceptibility. When the entire set of eleven numerals is considered, the lower 95% confidence limit of 0.68 Hz is smaller than the just-noticeable difference of 1 Hz (Lehiste).

In conclusion, then, a limited and controlled subset of the data shows a statistically significant but perceptually trivial correlation with the σ_2 stress perceived

in AR_05 numerals. Based on these findings, I do not consider pitch to be a cue for stress for this lexical category.

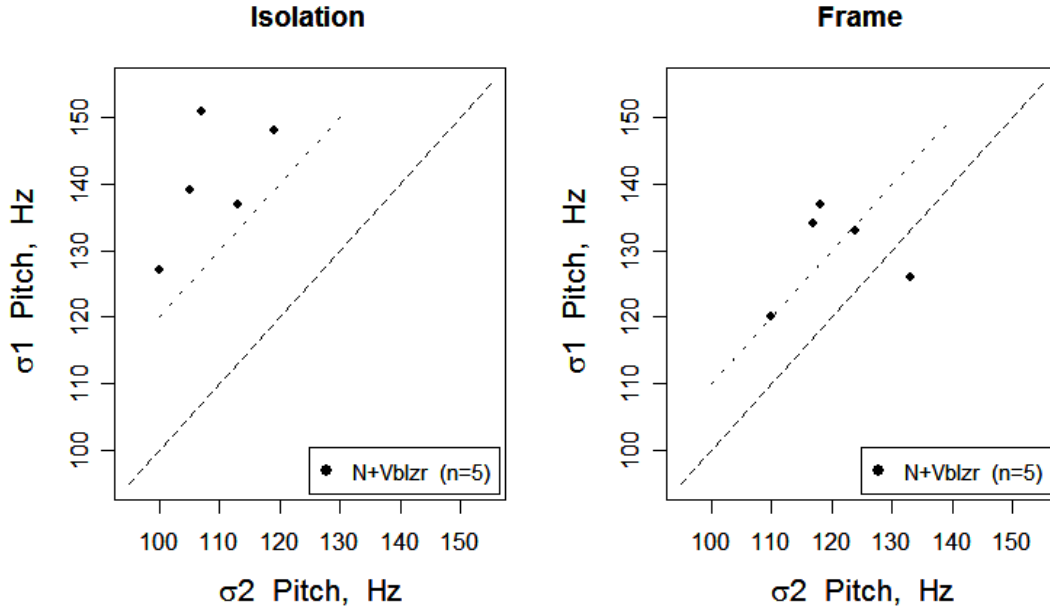
5.2.2.4 AR_05 Verbs

For AR_05 verbs, the isolation forms exhibit a consistent and strong correlation between pitch and σ_1 stress, despite the small sample size, as illustrated in Figure 5.23 on the following page.

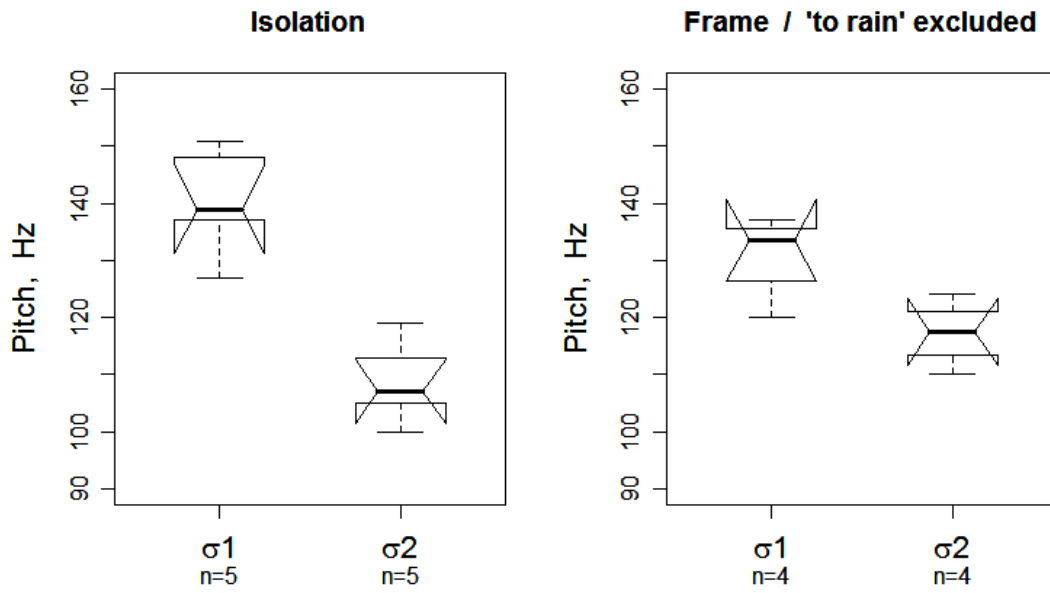
Frame forms too, show a clear correlation between pitch and stress, though the contrast across syllables is more subdued, with points plotting closer to the dashed line in Figure 5.23 . Also, as discussed in section 5.1.2, the verb [ˈnʌm.ˈbav] *gnam.babs* ‘to rain’ was produced with equal prominence on both syllables. This seems to have been an anomalous pronunciation, compared to the pronunciation of this same word in isolation and in the speaker’s own short, spontaneous sentence, and also compared to all of the other verbs both in isolation and in the sentence frame. I have included this verb in the distribution graphs below, for reference, but I have excluded it from box-and-whisker plots and statistical tests.

Figure 5.23 AR_05 / Verbs: Pitch

AR_05 / Verbs



AR_05 / Verbs



As was the case with numerals, since the sample of verbs is small, it is particularly important to check for any potential skewing due to incidental factors. That is, if all of the verbs here had a [higher.lower] vowel contrast, then the distribution observed could be entirely attributable to the predicted correlation between pitch and vowel height, and we would not be able to draw any firm conclusions about stress. However, this can be ruled out. The set of isolation forms includes only one [higher.lower] verb: [ʼvu.dzəp] *wu.brgyab* ‘to shoot a gun’; there are also three [same.height] forms, and one [lower.higher] form, and all of these clearly have a higher pitch on σ_1 , as well. For frame forms, the only verb with a [higher.lower] vowel contrast happens to be [ʼnɬm.ʼbau] *gnam.babs* ‘to rain’, which – as noted above – seemed to be an anomaly.

Having thus ruled out the potential influence of vowel height, we can conclude that the speaker has manipulated pitch in order to convey stress. Paired-sample t-tests, summarized in Table 5.8 below, confirm that the contrast across syllables is statistically significant ($p < 0.05$) for both isolation and frame forms. For the isolation forms, this contrast is also clearly perceptually significant: on average, pitch in σ_1 is 32 Hz higher than pitch in σ_2 , and the minimum significant mean difference (the upper 95% confidence interval) is -22 Hz. For the frame forms, the mean difference across syllables is -14 Hz, and the minimum probable mean difference is only -5.8 Hz.

Table 5.8 AR_05/ Verbs / Pitch: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz)	95 % conf. limits	
						lower	upper
AR_05	isolation	-9.0249	4	0.000835	-32	-41	-22
	frame *	-5.5092	3	0.01177	-14	-22	-5.8

* The verb 'to rain' has been excluded.

In conclusion, for the isolation forms of AR_05 verbs, pitch is a robust cue for σ_1 stress. The contrast across syllables is so strong that a listener could probably identify the stress pattern without reinforcement by any other acoustic cues. For the frame forms, pitch is a consistent correlate of stress, but the contrast across syllables is not as dramatic as for the isolation forms. Nonetheless, the pattern is clear, and is clearly distinct from the pitch pattern observed on non-verbs.

5.3 Pitch slope

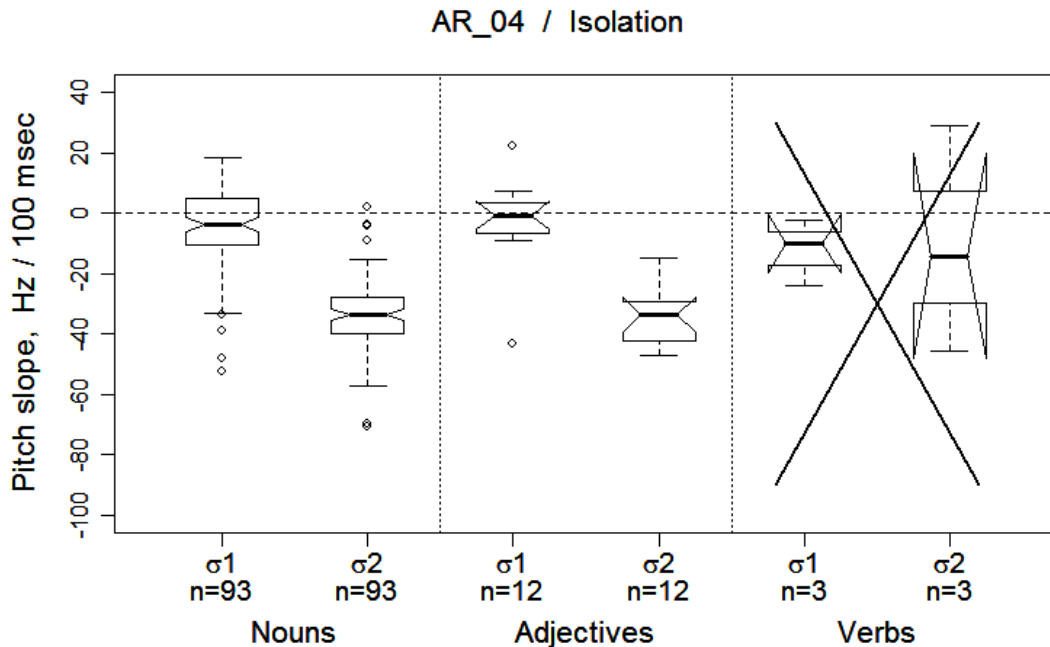
The very first time that I heard Rebkong Amdo Tibetan spoken, I was immediately struck by a dramatic fall in pitch on the second syllable of disyllabic nouns. This was such a highly perceptible feature – and it seemed so clearly to serve as a cue for stress – that it was at once obvious to me that a study of the acoustic correlates of stress in Rebkong Amdo could not be complete without considering this factor. I subsequently noted this same dramatic pitch slope in the speech of all of the other Rebkong Amdo speakers I worked with – in Xining, in Rebkong, and in Kathmandu.

Pitch slope data for speakers AR_04 and AR_05 is considered in detail in sections 5.3.1 and 5.3.2 below.

5.3.1 Pitch slope for speaker AR_04

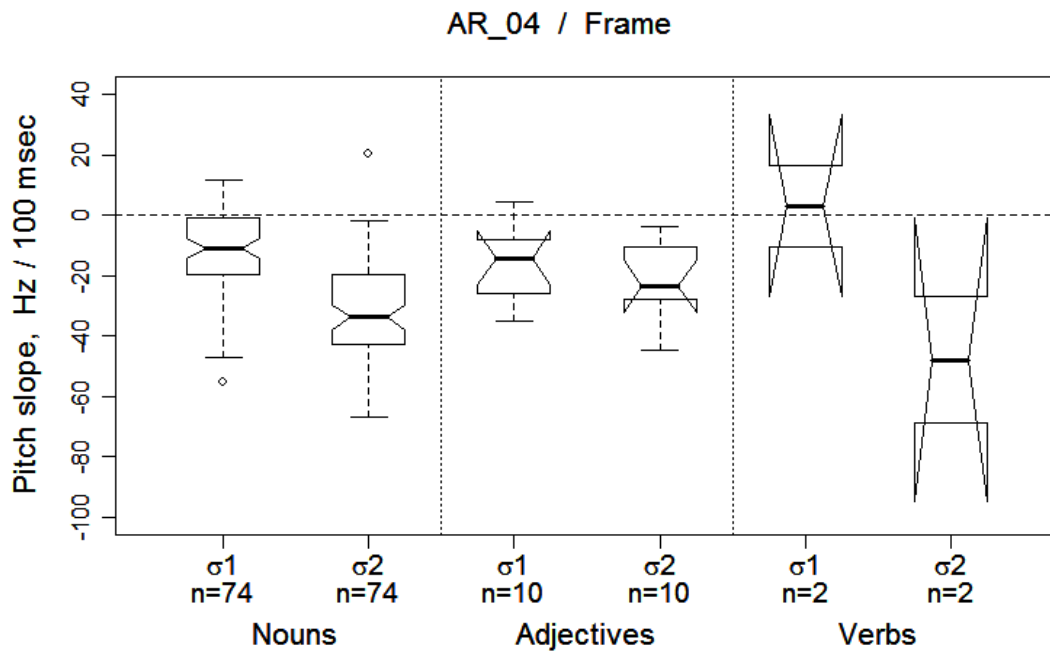
Pitch slope contrasts for the isolation forms of AR_04 nouns, adjectives, and verbs are illustrated by the box-and-whisker plots in Figure 4.21 above. For nouns and adjectives, pitch in $\sigma 1$ sometimes slopes upward, and sometimes slopes downward. In $\sigma 2$, pitch always slopes downward. There is a considerable contrast in pitch slope across syllables, with no overlap of the notches – representing the 95% confidence interval about the median – or the boxes – representing the interquartile range. As discussed previously, this speaker was not able to produce verbs in isolation in a natural way, so I consider this small set of data invalid and have X-ed through it here.

Figure 5.24 AR_04 / Isolation / Pitch slope contrasts



In the frame forms, illustrated in Figure 5.25 below, nouns and adjectives show similar patterns, though more subdued. For adjectives, measurements on the two syllables fall within much the same range, as evidenced by the overlap of notches and boxes. The two verbs analyzed also exhibit a “more downward” slope in σ_2 than in σ_1 , but this does not reflect a correlation with σ_1 stress; rather, it is a direct outcome of the pitch contrast across syllables, as discussed below.

Figure 5.25 AR_04 / Frame / Pitch slope contrasts



Pitch slope data for AR_04 nouns, adjectives, and verbs are considered in greater detail in sections 5.3.1.1 through 5.3.1.3 below.

In the distribution plot at left above, nearly all points fall to the right of the dashed line, indicating that pitch slope is “more downward” in σ_2 – the stressed syllable – than in σ_1 . That is, even in a word in which pitch slopes downward in σ_1 , it slopes downward more steeply in σ_2 . There are only a handful of exceptions to this pattern, including the compound form [dɔʁ.'s^ha] *'brog.sa* ‘nomad area’ at (σ_2 , σ_1) coordinates (-19, -52) and the monomorphemic form [xkɔ.'pa] *skud.pa* ‘thread’ at (-28, -39).

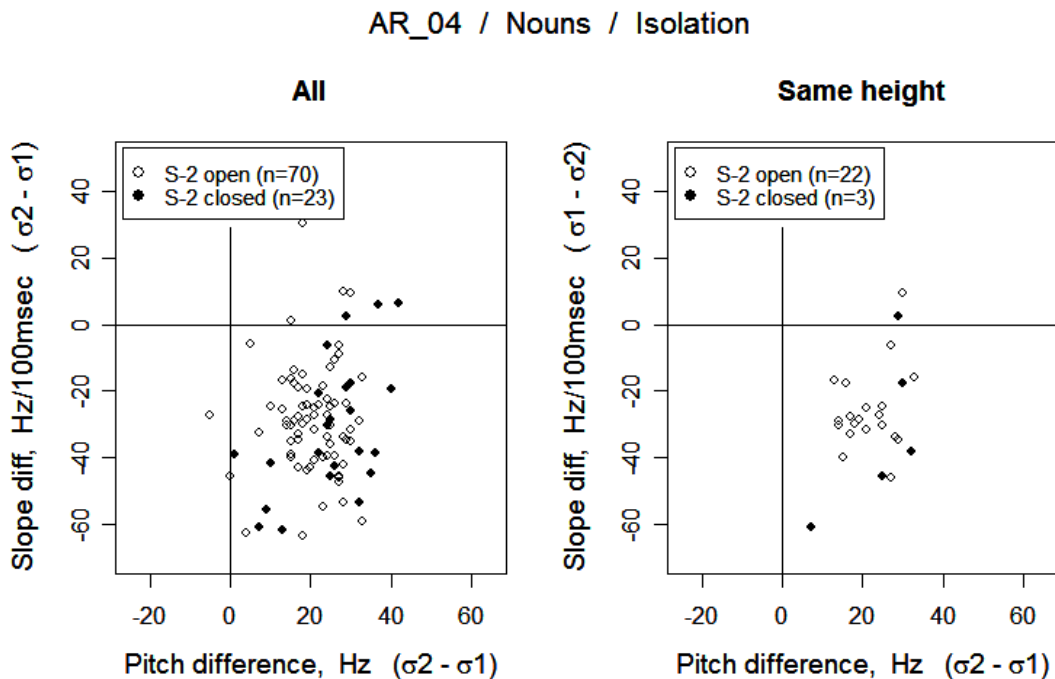
The statistical significance of the contrast in slope across syllables is confirmed by a paired-sample t-test. As indicated in Table 5.9 below, $p \ll 0.05$. The mean slope difference is -29 Hz/100msec, with a 95% confidence interval ranging from -33 Hz/100 msec to -25 Hz/100msec. I know that differences of this magnitude are perceptually highly significant, since I was struck by the contrast upon first hearing the language in the field, and also in these recordings. (There is no other basis by which to judge perceptual significance, since pitch slope has not been widely considered as an acoustic correlate of stress and there is no experimental evidence available.)

Table 5.9 AR_04 / Nouns / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

Setting	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_04	Isolation	-16.3817	92	< 2.2e-16	-29	-33	-25

Finally, what makes pitch slope such a prominent cue for stress in the AR_04 isolation forms is its contextual significance. The downward pitch slope in σ_2 contrasts with – and, in fact, is highlighted by – the overall increase in pitch across the word, from σ_1 to σ_2 . This contrast between pitch and pitch slope is illustrated in Figure 5.27 below. These graphs show the difference in pitch across syllables *vs.* the difference in pitch slope across syllables. The plot on the right includes only those words which have vowels of the same height in both syllables: this factor can have an intrinsic effect on pitch. Here, though, there is no remarkable difference between this graph and the one on the left, which includes all nouns regardless of vowel height.

Figure 5.27 AR_04 / Nouns / Isolation: Pitch difference *vs.* pitch slope difference



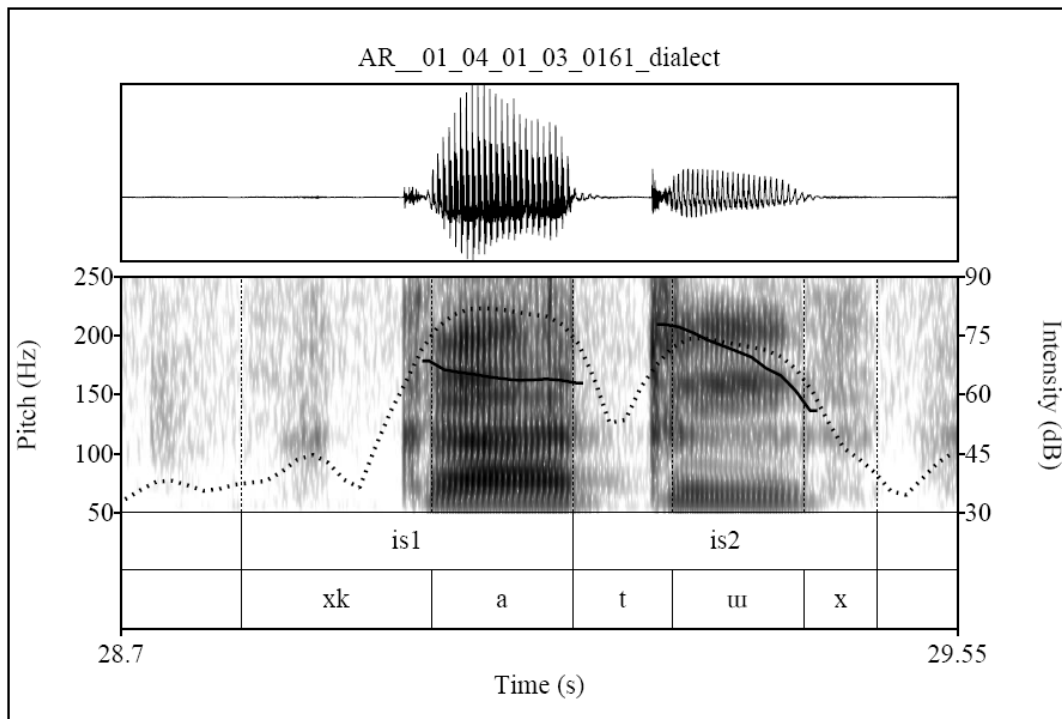
Almost all of the nouns fall in the lower right quadrant of the graph. Points plot to the right of the y-axis – meaning that pitch rises from σ_1 to σ_2 – and below the

x-axis – meaning that pitch slopes downward within σ_2 . This contrast – a downward slope set against a background rise – lends particular salience to the pitch slopes measured in σ_2 .⁴⁴

The fact that most nouns fall well within the lower right quadrant reflects the fact that pitch and pitch slope usually reinforce one another as correlates of stress. A prototypical example is the compound noun [xka.'tux] *skad.lugs* ‘dialect’, shown in Figure 5.28 below. In the vowel in σ_1 , pitch slopes downward at -10 Hz/100msec. In σ_2 pitch slopes downward much more steeply, at -51 Hz/100msec. The difference in slope is calculated as $(-51) - (-10) = -41$ Hz/100msec.

⁴⁴ This is quite different from the pattern observed in BSh_03 nouns, where – rather than contrasting with the pitch pattern – pitch slope merged with the pitch pattern and could not be distinguished from it. In Figure 3.39, BSh_03 nouns plot in the upper right quadrant of the graph.

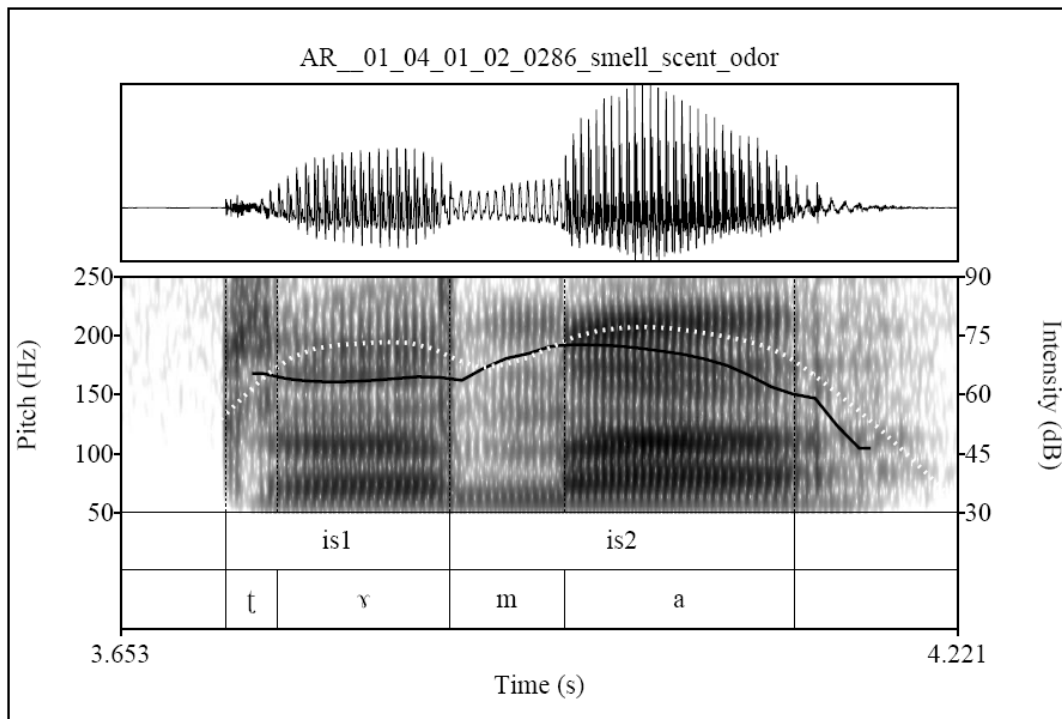
Figure 5.28 AR_04 / Noun / Pitch trace: [xka.'tux] *skad.lugs* 'dialect'



In addition, though, the downward-sloping pitch in σ_2 takes place at a higher overall pitch level compared to that in σ_1 : the average pitch over the medial 50% of the vowel was measured as 164 Hz in σ_1 , and as 186 Hz in σ_2 , for an increase across syllables of 22 Hz. Thus both pitch and pitch slope are more prominent in σ_2 . In fact, in this word, pitch and pitch slope are the only parameters that convey σ_2 stress. Intensity does not play a role: intensity is higher on σ_1 , not σ_2 (Δ intensity = -9 dB). Nor does vowel length play a role: the vowel in σ_1 is longer than that in σ_2 (Δ duration = -10 msec). Finally, vowel quality does not contribute to stress, either: the vowel in σ_1 is full [a] while that in σ_2 is reduced to [u].

Another representative example is the monomorphemic noun [tʃ.'ma] *dri.ma* ‘smell, scent, odor’, shown in Figure 5.29 below. The difference in slope across syllables is $(-28.29) - (-3.85) = -24 \text{ Hz}/100\text{msec}$; the difference in pitch is 22 Hz. In this case, intensity also corresponds with stress ($\Delta\text{intensity} = 4 \text{ dB}$), and the vowel is longer in σ_2 ($\Delta\text{duration} = 39 \text{ msec}$). This example also shows that when the consonant at the syllable boundary is a sonorant, the pitch contour exhibits some transitional curves, and the slope on σ_2 is not so distinctly straight and constant as it was in Figure 5.28.

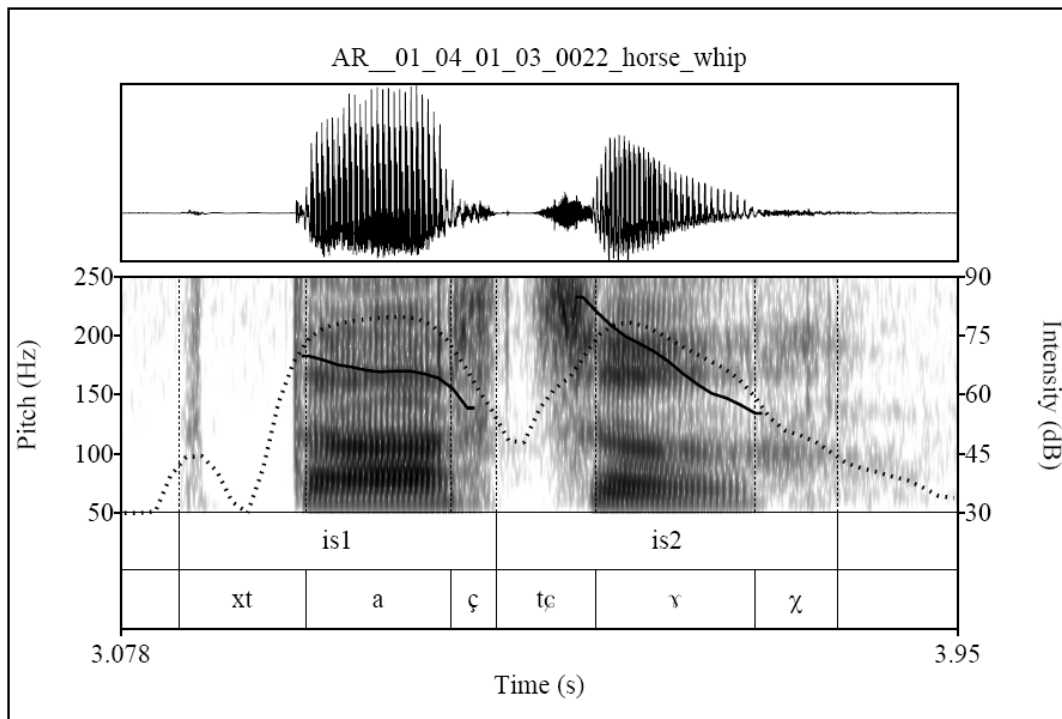
Figure 5.29 AR_04 / Noun / Pitch trace: [tʃ.'ma] *dri.ma* ‘smell, scent, odor’



There are a few nouns in which pitch slope alone lends prominence to σ_2 ; these are points that fall very close to the y-axis in Figure 5.27. An example is the compound noun [xtax.'tɕɤχ] *rta.lcag* 'horse whip', shown in isolation in Figure 5.30 below. Pitch declines gently over the vowel in σ_1 , with slope measured as -8.3 Hz/100msec. In σ_2 pitch declines much more steeply, at a rate of -56 Hz/100msec. The difference in pitch slope across syllables is audibly (and here, visibly) quite striking; the difference in slope is calculated as -48 Hz/100msec.

Average pitch itself does not contribute to the perception of σ_2 stress in this word: the average pitch across the medial 50% of the vowel was measured as 171 Hz in σ_1 and as 172 Hz in σ_2 , for an increase across syllables of only 1 Hz. (This is very different than what we observed in nouns produced by Balti speaker BSh_03, in section 4.23. There, pitch slope was an epiphenomenon of the pitch increase from σ_1 to σ_2 .) And intensity is actually higher in σ_1 , the syllable which is *not* stressed: intensity over the medial 50% of the vowel was measured as 79 dB in σ_1 and 74 dB in σ_2 , for a decrease across syllables of 5 dB. (Of course, this intensity contrast may in part be attributed to the contrast in vowel height across syllables, as will be discussed in greater detail in section 5.4.) The vowel is longer in σ_2 (166 msec) than in σ_1 (151 msec), but it is not possible to determine whether this reflects a correlation with stress or not, since vowels are often lengthened in the final syllable of an utterance. Finally, vowel quality cannot be a factor contributing to the perception of stress, since the vowel in σ_2 is reduced to [ɤ], while the vowel in σ_1 is full [a].

Figure 5.30 AR_04 / Noun / Pitch trace: [xtaɕ.'tɕɣɕ] *rta.lcag* ‘horse whip’

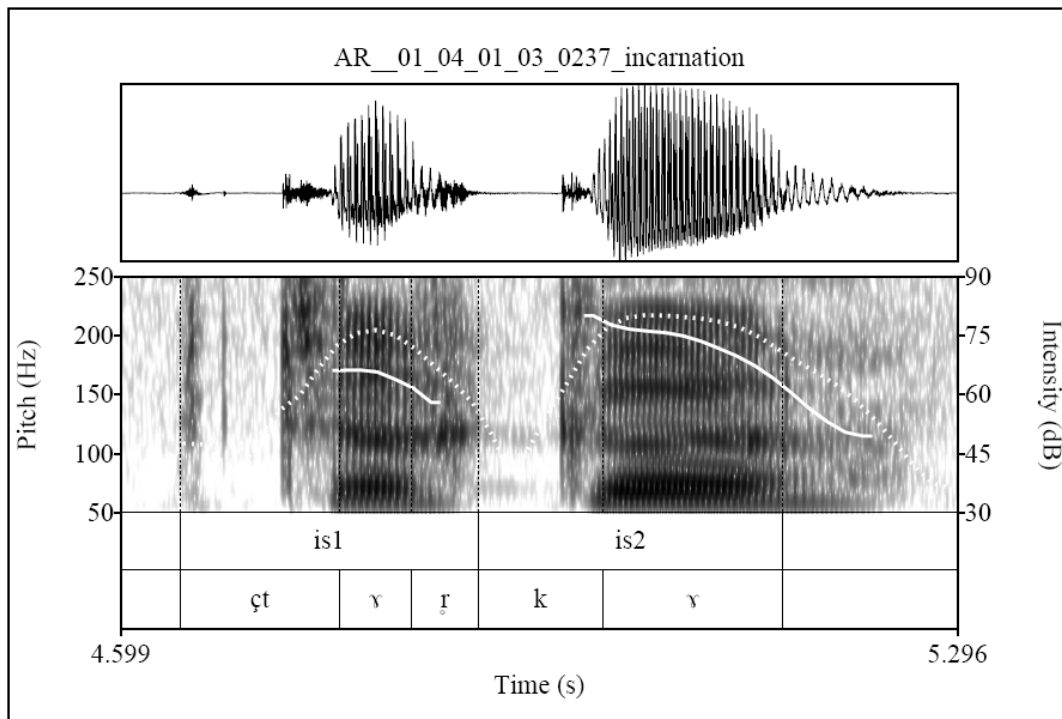


This example thus demonstrates that pitch slope alone can be manipulated in order to convey stress. In fact, one notable feature of the pitch traces in Figure 5.28 and Figure 5.30 is that the downward slope in σ_2 is straight and constant. There is no leveling off of the curve at any point, no stage at which the speaker maintains a steady target pitch, however brief. Rather, it appears that the slope itself is the speaker's target.

There are also a few nouns in which pitch plays a greater role than pitch slope in conveying σ_2 stress; these are points that fall close to the y-axis in Figure 5.27. In [ɕtɣɕ.'kɣ] *sprul.sku* ‘incarnation, incarnate lama’, for example, the contrast in pitch slope across syllables is only -7 Hz/100msec, but the contrast in pitch across syllables

is 27 Hz. (In this case there is no difference in vowel quality across syllables, and intensity increases 2 dB from σ_1 to σ_2 .) The pitch trace for this word is shown in Figure 5.31 below.

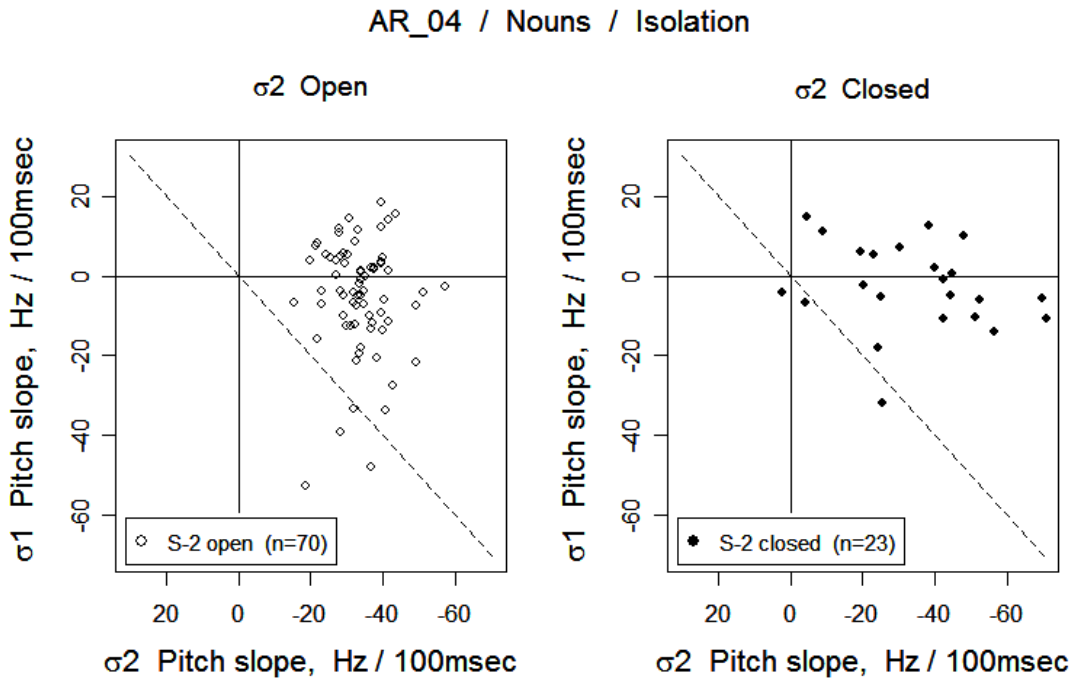
Figure 5.31 AR_04 / Noun / Pitch trace: [çtɣr̥.'kɣ] *sprul.sku* ‘incarnation, incarnate lama’



The examples above included cases in which σ_2 was open, and cases in which σ_2 was closed. As illustrated in Figure 5.32 below, pitch slope shows the same general pattern for both templates. The only slight difference is that, when σ_2 is closed, pitch in the vowel shows greater variation, with points dispersed a bit more widely along the x-axis. But in both structural types, with only a handful of exceptions, points fall to the right of the dashed line, indicating that slope is more

prominent in σ_2 . This reflects the fact that the speaker is consistently – though unconsciously – manipulating pitch in order to create this distinctive slope, which plays such an important role in conveying stress.

Figure 5.32 AR_04 / Nouns / Isolation: Pitch slope vs. σ_2 closure

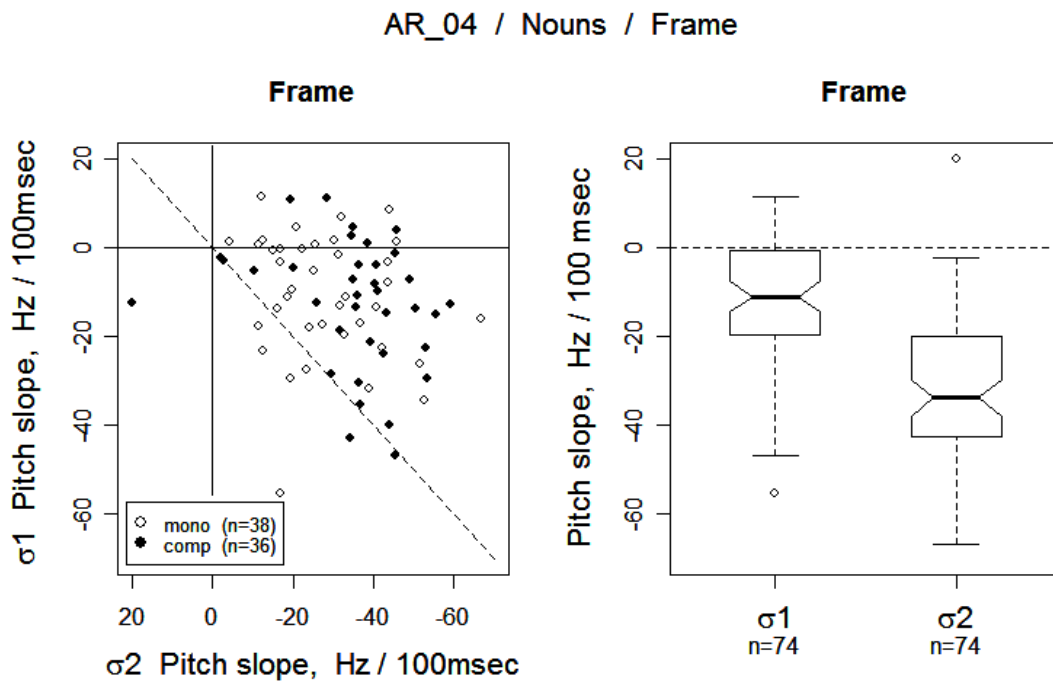


These patterns are different from what we observed in nouns produced in isolation by speaker BM_01. In that case, I attributed the “more downward” pitch slope in σ_2 to a falling terminal intonation contour. This downward slope was sometimes blocked in words with σ_2 closed, which accounts for the random distribution of points to either side of the dashed line in Figure 4.59. That is not what happens in the case of Rebkong Amdo: because pitch slope is crucial in conveying stress, vowels are manipulated so that pitch slopes prominently downward even when σ_2 is closed.

In summary, for AR_04 nouns produced in isolation, pitch slope contrasts across syllables are statistically, perceptually, and contextually significant. In most nouns, pitch and pitch slope both play a role in conveying stress. In a handful of cases, though, pitch is a cue for stress where pitch slope is not, and in another handful of cases, pitch slope is a cue for stress where pitch is not. In this sense, pitch and pitch slope function as complementary reflexes of fundamental frequency.

Pitch slope also shows a robust correlation with σ_2 stress in AR_04 nouns produced within the carrier sentence [ʔa.tʃu 'kɛ.ki X 'se.'ra] 'In our language, we X say'. As illustrated in Figure 5.33 below, in the frame forms pitch in σ_1 usually slopes downward, but sometimes slopes upward. In σ_2 , pitch always slopes downward, and usually fairly steeply. In the box-and-whisker plot to the right, there is no overlap of either the notches or the boxes, indicating that the range of measurements for the two syllables are fairly distinct.

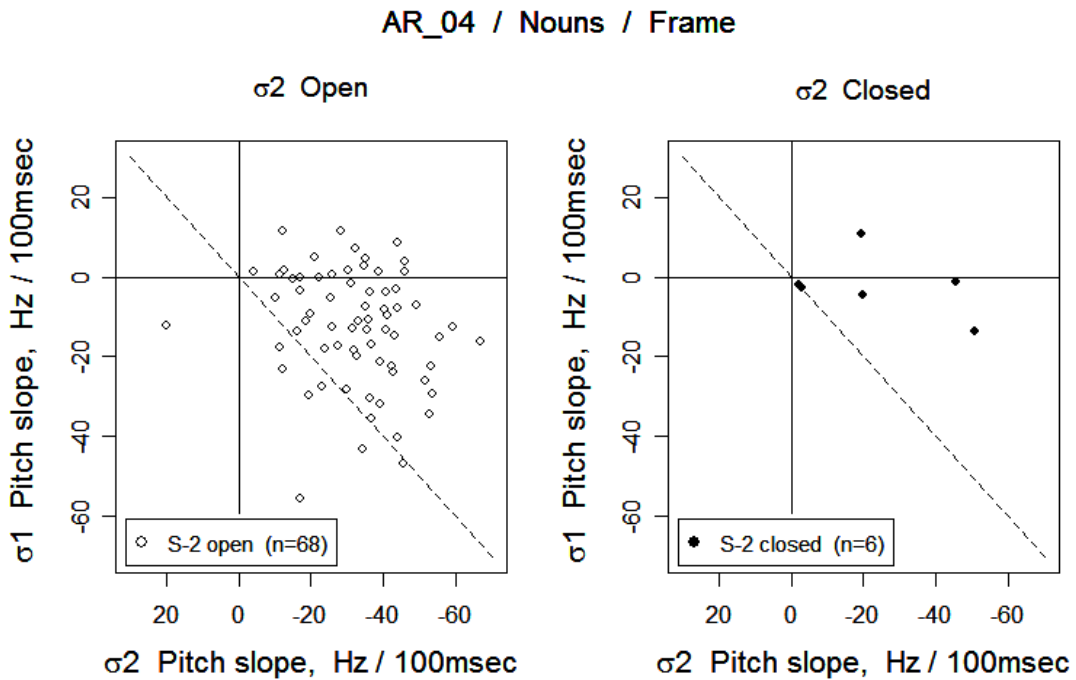
Figure 5.33 AR_04 / Nouns / Frame: Pitch slope



As was the case with the isolation forms, in the graph to the left above nearly all points fall to the right of the dashed line, again indicating that pitch slope is “more downward” in σ_2 – the stressed syllable – than in σ_1 . There are only a handful of exceptions to this pattern. Aside from [xkɤ.'pa] *skud.pa* ‘thread’ at (-17, -55) and [kɤ.'ʒi] *sku.lus* ‘body.HON’ at (20, -12), most of these exceptions fall quite close to the dashed line, indicating that the contrast in slope across syllables is minimal. [kɤ.'ʒi] *sku.lus* ‘body.HON’ is the only token plotting to the left of the y-axis, with a steep upward slope in σ_1 .

Figure 5.34 below suggests that slope patterns are the same whether σ_2 is open or closed.

Figure 5.34 AR_04 / Nouns / Frame: Pitch slope vs . σ_2 closure



The statistical significance of the difference in slope across syllables is confirmed by paired-sample t-tests, summarized in Table 5.10 below: $p \ll 0.05$. The mean difference in slope across syllables is -20 Hz/100msec, with a 95% confidence interval ranging from -24 to -15 Hz/100msec.

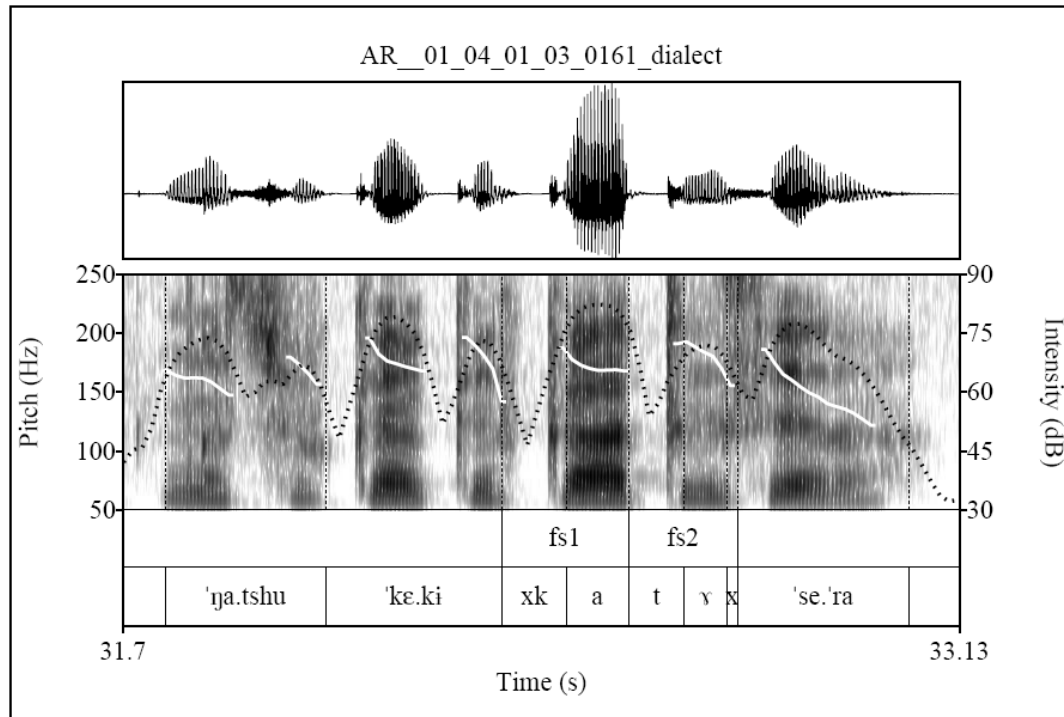
Table 5.10 AR_04 / Nouns / Pitch slope: Results of paired-sample t-tests (two-tailed)

Setting	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_04	Frame	-9.224	73	7.087e-14	-20	-24	-15

Thus pitch slope shows a robust and consistent correlation with σ_2 stress for the frame forms of nouns produced by speaker AR_04, as it did for the isolation forms.

A fairly typical example of the frame forms is [xka.'tux] *skad.lugs* ‘dialect’, illustrated in Figure 5.35 below.

Figure 5.35 AR_04 / Noun / Frame / Pitch trace: [xka.'tux] *skad.lugs* ‘dialect’



Measurements of acoustic parameters in the isolation and frame forms of [xka.'tux] / [xka.'tux] *skad.lugs* ‘dialect’ are compared in Table 5.11 below.⁴⁵ Even though the rate of speech is much faster in the frame form – the total length of the word was 651 msec in isolation, and only 402 msec in the sentence frame – the pitch and pitch slope contrasts are similar. In both settings, pitch and pitch slope are more prominent in σ_2 , the stressed syllable. Both intensity (discussed in section 5.4) and

⁴⁵ The isolation form was illustrated in Figure 5.28.

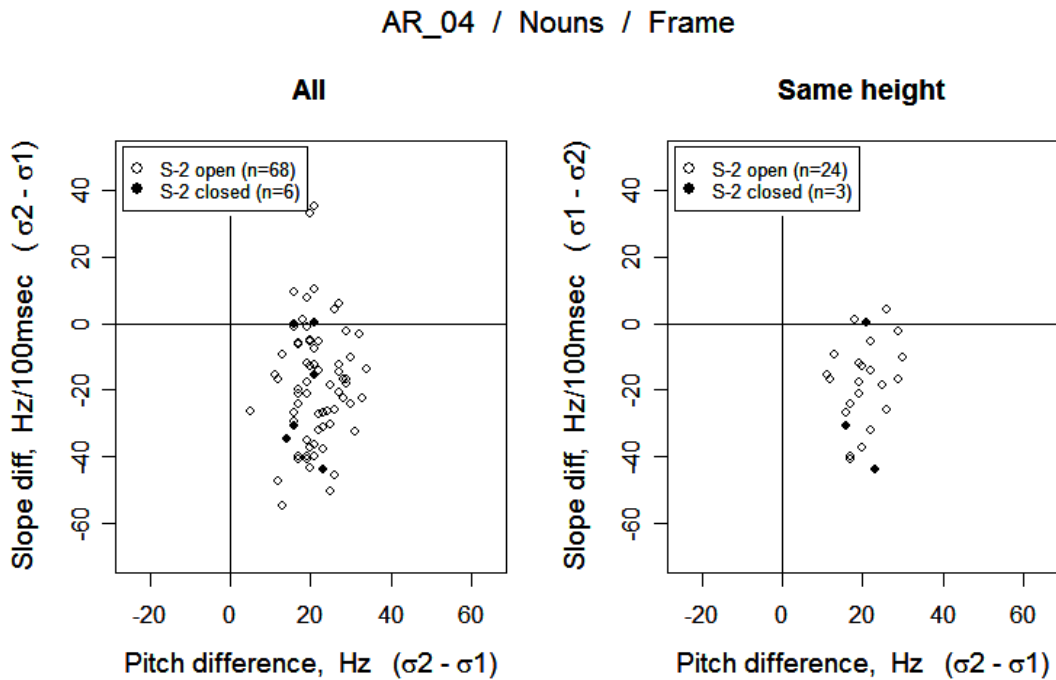
vowel duration (discussed in section 5.5) are more prominent in σ_1 , the unstressed syllable. (For both parameters, this is consistent with what one might predict based on intrinsic variation as a function of vowel quality.)

Table 5.11 AR_04 / Nouns / Potential stress correlates: [xka.'tuux] *skad.lugs* ‘dialect’

Parameter	Isolation			Frame		
	σ_1	σ_2	Δ	σ_1	σ_2	Δ
Pitch (Hz)	164	186	22	170	184	14
Slope (Hz/100msec)	-10	-51	-41	-14	-51	-37
Intensity (dB)	82	73	-9	82	72	-10
Duration (msec)	145	135	-10	105	73	-32

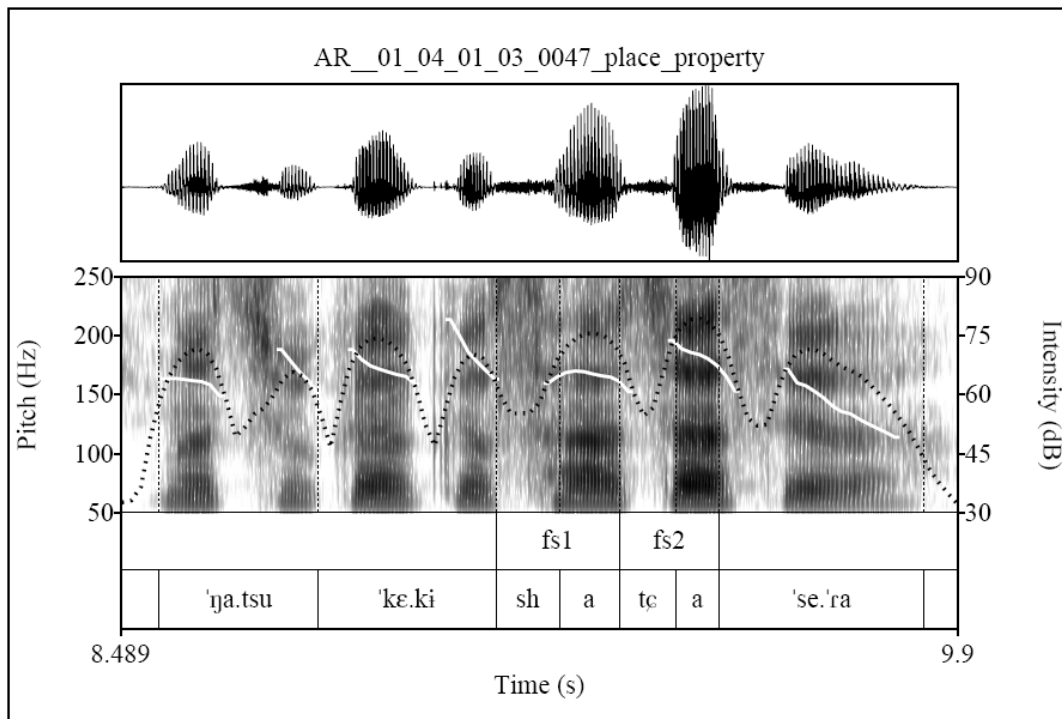
The pitch and pitch slope differences measured for [xka.'tuux] *skad.lugs* ‘dialect’ are fairly representative of the frame forms. This word plots at coordinates (14, -35) in the graph at left in Figure 5.36 below, which illustrates the relationship between these two reflexes of fundamental frequency. (The graph at right includes only those nouns with vowels of the same height in both syllables. There is no major difference between the two plots.)

Figure 5.36 AR_04 / Nouns / Frame: Pitch difference vs. pitch slope difference



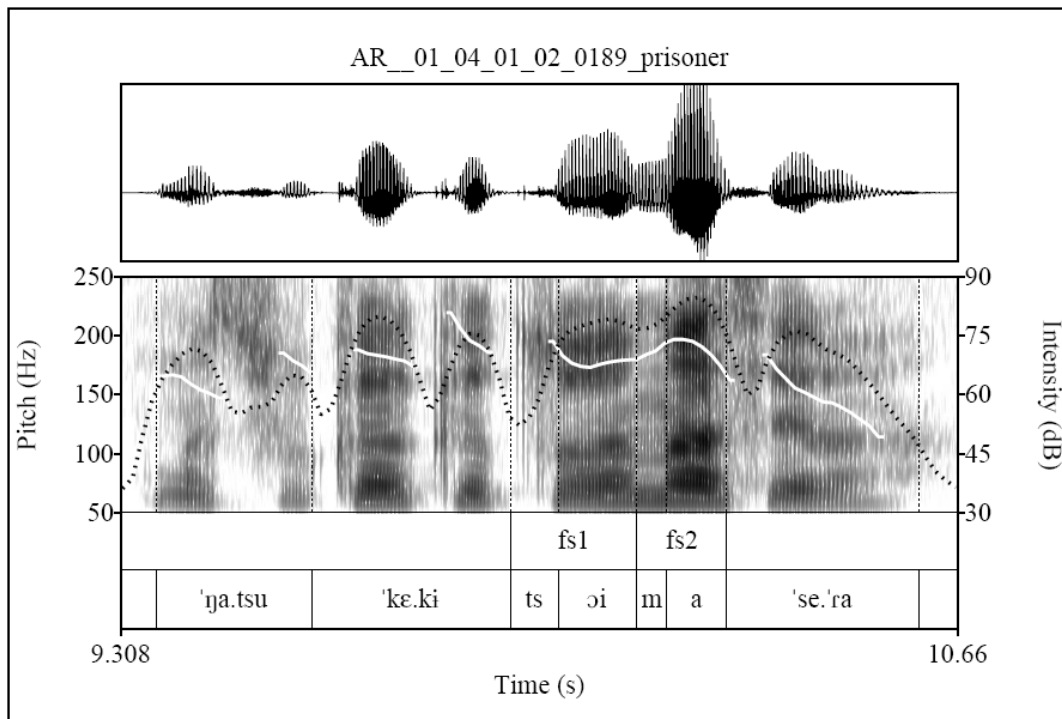
As was the case with the isolation forms (Figure 5.27), most of the frame forms plot in the lower right quadrant of the graph. This indicates that pitch and pitch slope are usually both more prominent in σ_2 , reinforcing one another as correlates of σ_2 stress. A typical example – the compound noun [s^ha.'tʃa] *sa.cha* ‘land, place, property’ – is shown in Figure 5.37 below. For this word, Δ pitch = 16 Hz, Δ pitch slope = -28 Hz/100msec, Δ intensity = 4 dB (all more prominent on σ_2); and Δ vowel duration = -29 msec (more prominent on σ_1).

Figure 5.37 AR_04 / Noun / Frame / Pitch trace: [s^ha.'tɕa] *sa.cha* ‘land, property’



An example with a continuant at the syllable boundary is the monomorphemic noun [tsɔi.'ma] *btson.ma* ‘prisoner’, shown in Figure 5.38 below. The contrasts across syllables are: Δ pitch = 16 Hz; Δ pitch slope = -29 Hz/100msec; Δ intensity = 5 dB; Δ vowel duration = -31 msec.

Figure 5.38 AR_04 / Noun / Frame / Pitch trace: [tsɔi.'ma] *btson.ma* ‘prisoner’

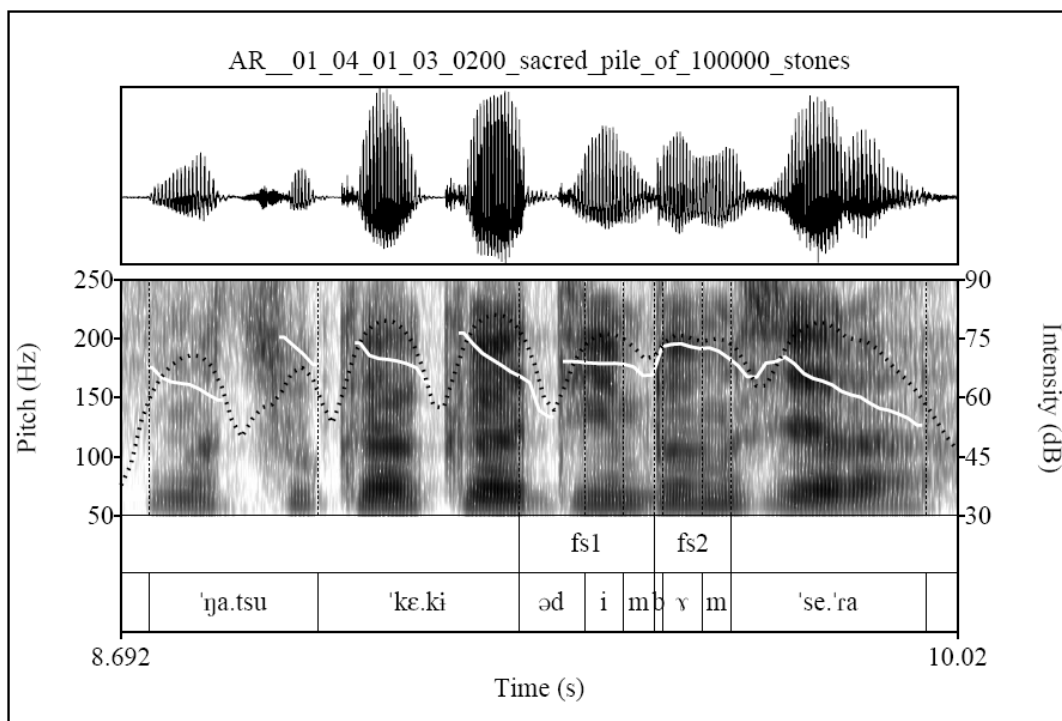


There were a few tokens among the isolation forms in which pitch slope was a strong cue for stress but pitch was not (i.e., points which plot close to the y-axis in Figure 5.27); this is never the case for the frame forms. Here, pitch always correlates with stress; all points fall to the right of the y-axis.

However, there are a few frame forms in which pitch is a strong cue for stress but pitch slope is not; these are the points plotting above the x-axis in Figure 5.36 above. (The two nouns up at the top, with a pitch slope difference of ~35 Hz/10msec, are [xkɣ.'pa] *skud.pa* ‘thread’ and [kɣ.'ʒi] *sku.lus* ‘body.HON’.) One such noun is [ʔdim.'bɣm] *rdo. 'bum* ‘sacred pile of 100,000 stones’, shown in Figure 5.39 below. In

this case, the difference in slope across syllables is only 0.05 Hz/100msec, but the difference in pitch across syllables is 16 Hz. The intensity difference across syllables for this word is 0 dB; the vowel duration difference is 2 msec; pitch is thus the sole cue for stress.

Figure 5.39 AR_04 / Noun / Frame / Pitch trace: [ˈdim.ˈbʊm] *rdo.ˈbum* ‘sacred pile of 100,000 stones’



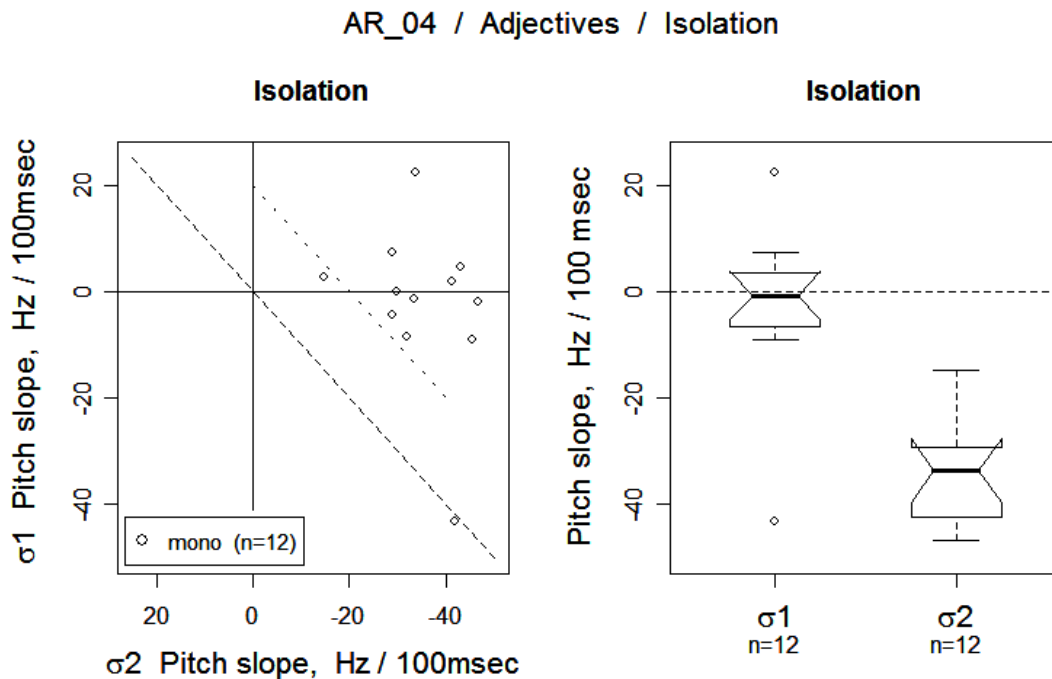
These patterns are completely different than those exhibited in frame forms produced by speaker BSh_03. In Figure 4.39 in the previous chapter, nearly all points fall in the upper right quadrant. In that case, pitch slope on σ_2 was “more upward” than on σ_1 , and this was a direct consequence of the increase in pitch across syllables.

5.3.1.2 AR_04 Adjectives

For the adjectives produced by speaker AR_04, pitch slope is a strong acoustic correlate of stress for the isolation forms, but not for the frame forms.

Slope contrasts across syllables for the isolation forms are illustrated in Figure 5.40 below. In the graph on the left, points are distributed both above and below the x-axis: sometimes pitch slopes upward in σ_1 , and sometimes it slopes downward. All points fall to the right of the y-axis, indicating that pitch always slopes downward in σ_2 . With the exception of [chɤ.'ko] *phyug.po* 'rich', at (σ_2, σ_1) coordinates (-43, -42), pitch slopes "more downward" in σ_2 than in σ_1 . In fact, the difference in slopes is generally more than -20 Hz/100msec, as indicated by the dotted reference line.

Figure 5.40 AR_04 / Adjectives / Isolation: Pitch slope contrasts



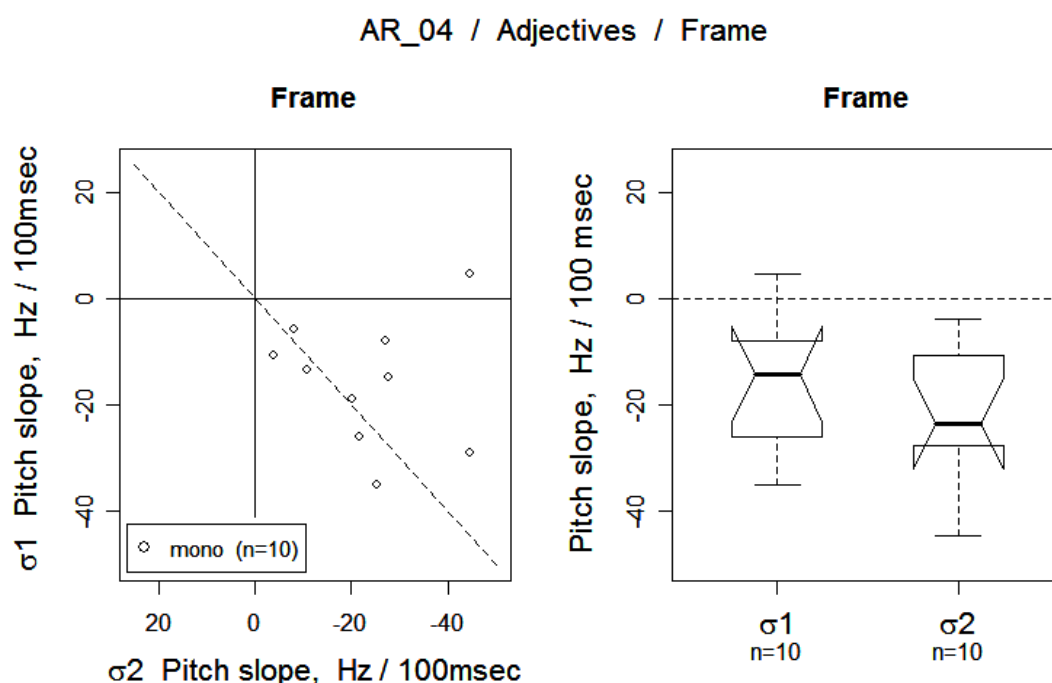
A paired-sample t-test, summarized in Table 5.12 below, confirms that the slope contrast is statistically significant ($p \ll 0.05$), with a mean difference across syllables of -32 Hz/100msec.

Table 5.12 AR_04 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_04	isolation	-7.3033	11	1.536e-05	-32	-42	-23

Slope contrasts for the frame forms of AR_04 adjectives are plotted in Figure 5.41 below. With one exception – [tshan.di] *tsha. 'di* ‘hot’ – all points fall in the lower right quadrant of the graph, indicating that pitch slopes downward in both syllables of each word. Furthermore, points fall close to and to either side of the dashed line, indicating that sometimes slope is slightly “more downward” in σ_1 , and sometimes slightly “more downward” in σ_2 . In the box-and-whisker plot to the right, both the notches and boxes overlap considerably. That is, the ranges of values measured in the two syllables are quite similar.

Figure 5.41 AR_04 / Adjectives / Frame: Pitch slope contrasts



Indeed, as summarized in Table 5.13 below, a paired-sample t-test yields a p-value > 0.05 and a 95% confidence interval which spans zero. This means that the null hypothesis – that there is no difference in slope across syllables – cannot be rejected.

Table 5.13 AR_04 / Adjectives / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed)

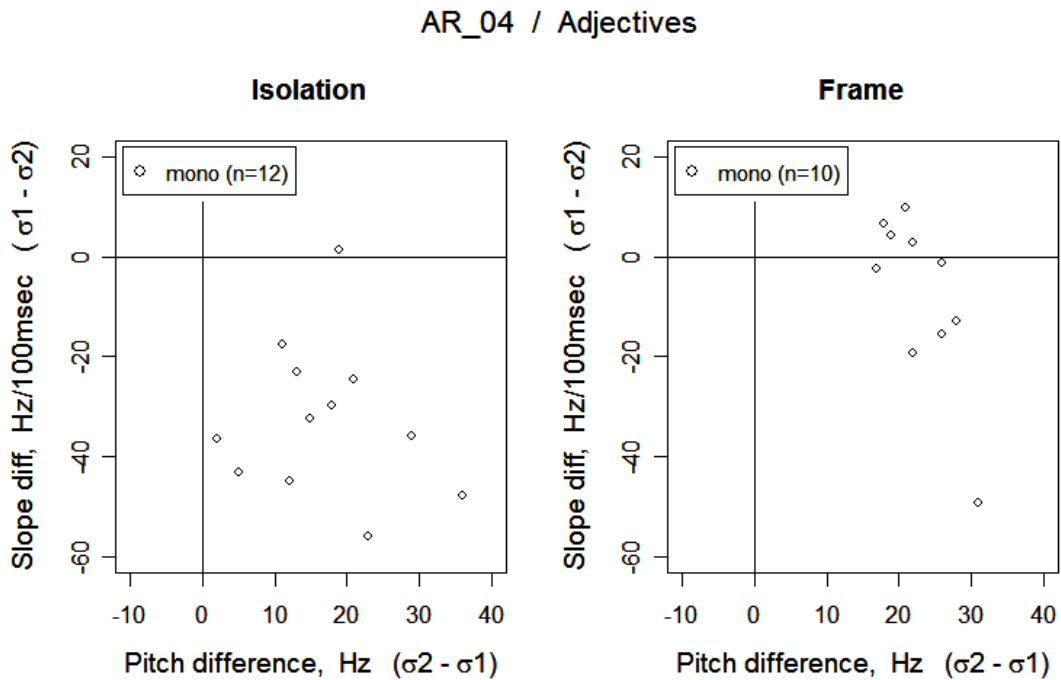
Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_04	frame	-1.3817	9	0.2004	-7.7	-20	4.9

As was the case for nouns, pitch and pitch slope are complementary correlates of stress for AR_04 adjectives, in the sense that pitch is a reliable stress cue exactly

when pitch slope is not, and pitch slope is a reliable stress cue exactly when pitch is not. For the majority of isolation forms, the two parameters reinforce one another in conveying σ_2 stress.

Pitch differences and pitch slope differences across syllables for AR_04 adjectives are plotted in Figure 5.42 below. For the isolation forms, nearly all adjectives fall in the lower right quadrant: both pitch and pitch slope are prominent on σ_2 , and are redundant cues for stress. In the case of [çhɣ.'ko] *phyug.po* 'rich', which plots just above the x-axis, slope is not a correlate of stress, but pitch is 19 Hz higher on σ_2 than on σ_1 . Conversely, on those words which plot closest to the y-axis – indicating that pitch is a weak cue – pitch slope is a very strong cue, with differences of more than -30 Hz/100msec across syllables.

Figure 5.42 AR_04 / Adjectives: Pitch difference vs. pitch slope difference

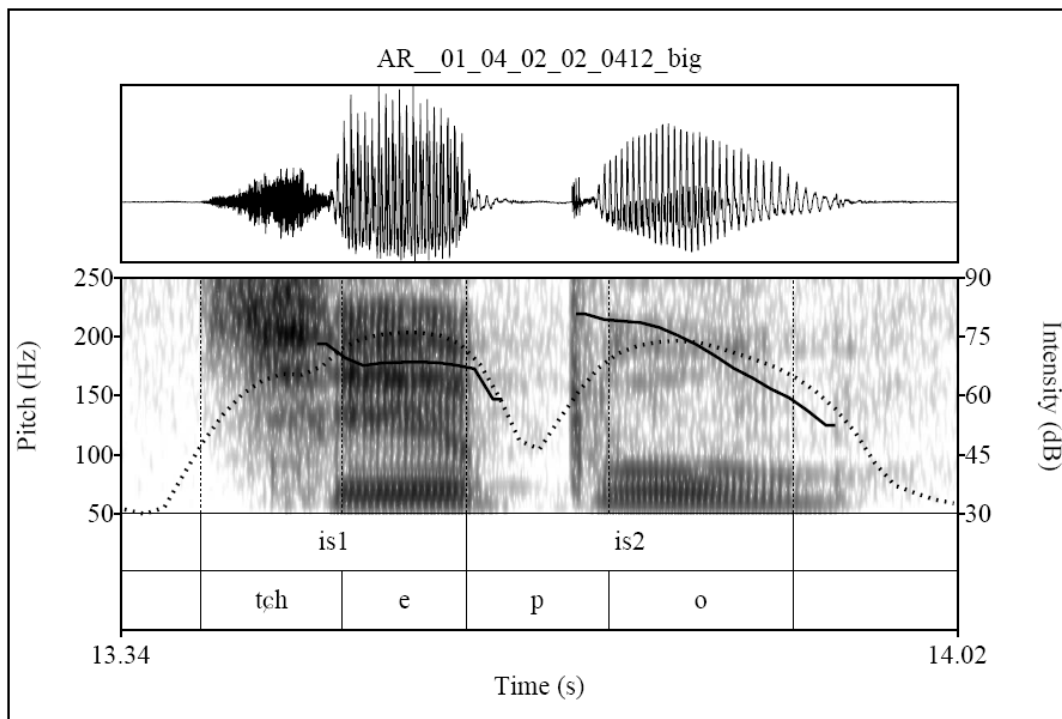


For the frame forms, shown on the right above, pitch is a robust cue for stress in all cases, so it does not matter that pitch slope sometimes is not. (The words which plot above the x-axis are [tʰɣ.'ko] *phyug.po* ‘rich’, [ka.'ru] *dkar.po* ‘white’, [sʰo.'ma] *so.ma* ‘new’, and [xkɔim.'bu] *dkon.po* ‘rare’.)

The pitch trace for the adjective [tʰe.'po] *chen.po* ‘big’ in Figure 5.43 below provides an illustration. For this word, since the consonant at the syllable boundary is a stop, pitch leaps upward from σ_1 to σ_2 , and the downward slope in σ_2 is fairly straight and constant. Pitch slope was measured as -2.1 Hz/100msec in σ_1 , and -47 Hz/100msec in σ_2 ; the difference in slope -45 Hz/100msec. Pitch is also a correlate of stress for this word: pitch was measured over the medial 50% of the vowel as 178 Hz

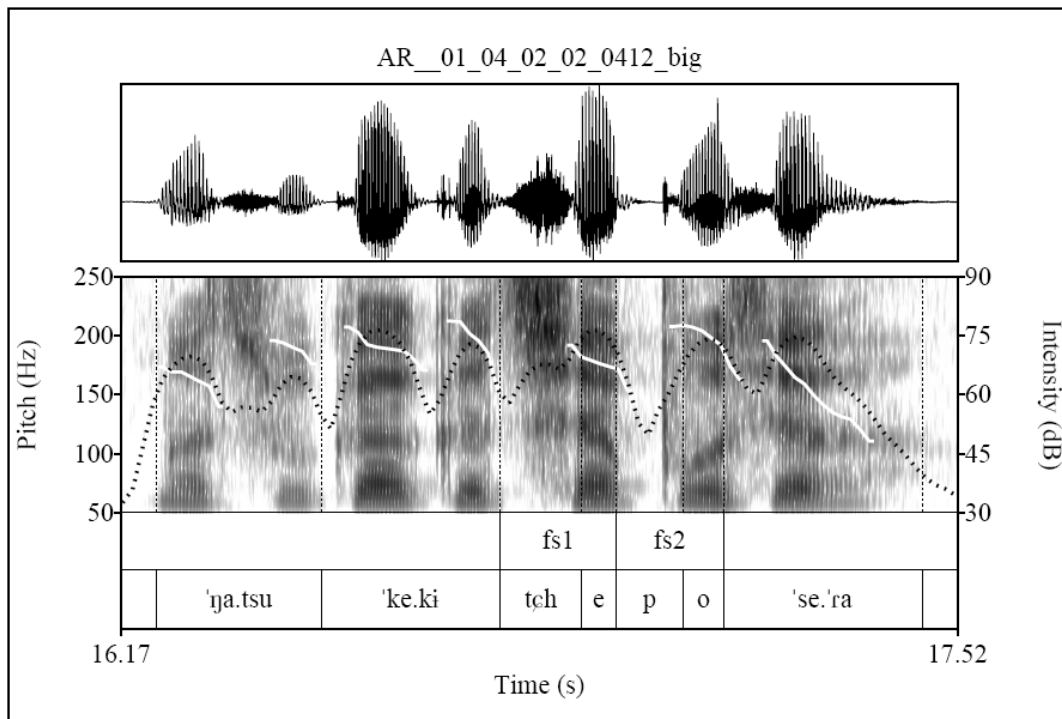
in σ_1 and 190 Hz in σ_2 , for an increase in pitch of 12 Hz. This word thus plots in the lower right quadrant in the graph on the left above, at coordinates (12, -44).

Figure 5.43 AR_04 / Adjectives / Pitch trace: [tʰe.'po] *chen.po* 'big'



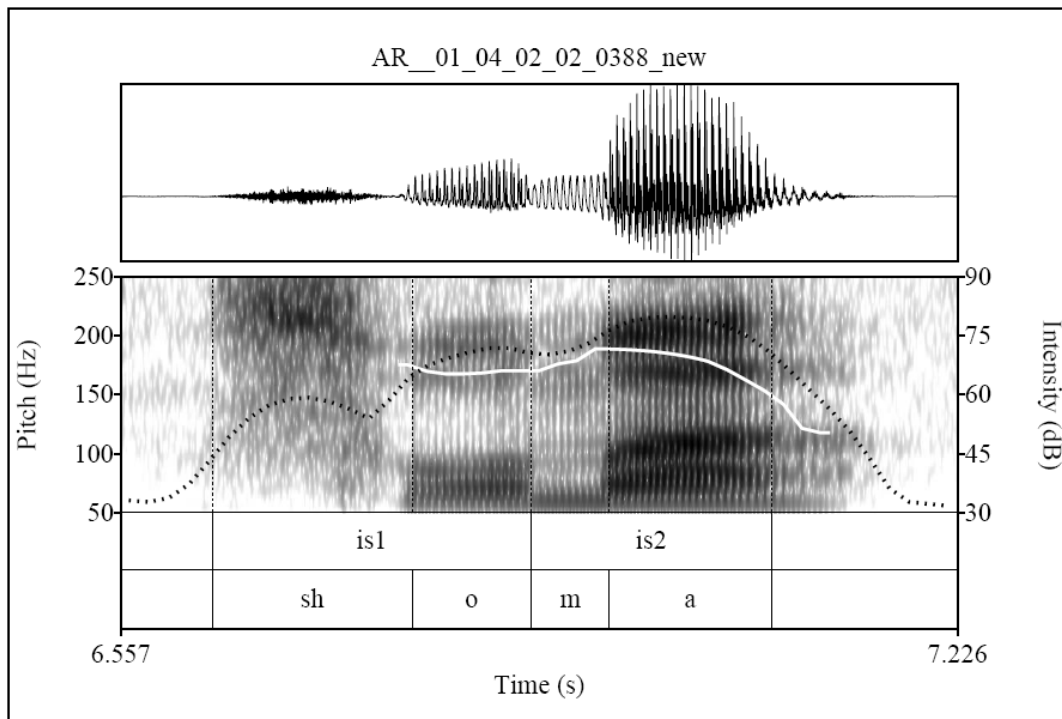
When [tʰe.'po] *chen.po* 'big' is produced in the sentence frame, the contrast in slope is more subdued, and the contrast in pitch is more pronounced. Here, $\Delta\text{pitch} = 26$ and $\Delta\text{pitch slope} = -16$, so this point still plots in the lower right quadrant; both pitch and pitch slope are correlates of stress.

Figure 5.44 AR_04 / Adjectives / Frame / Pitch trace: [tɕ^he.'po] *chen.po* 'big'



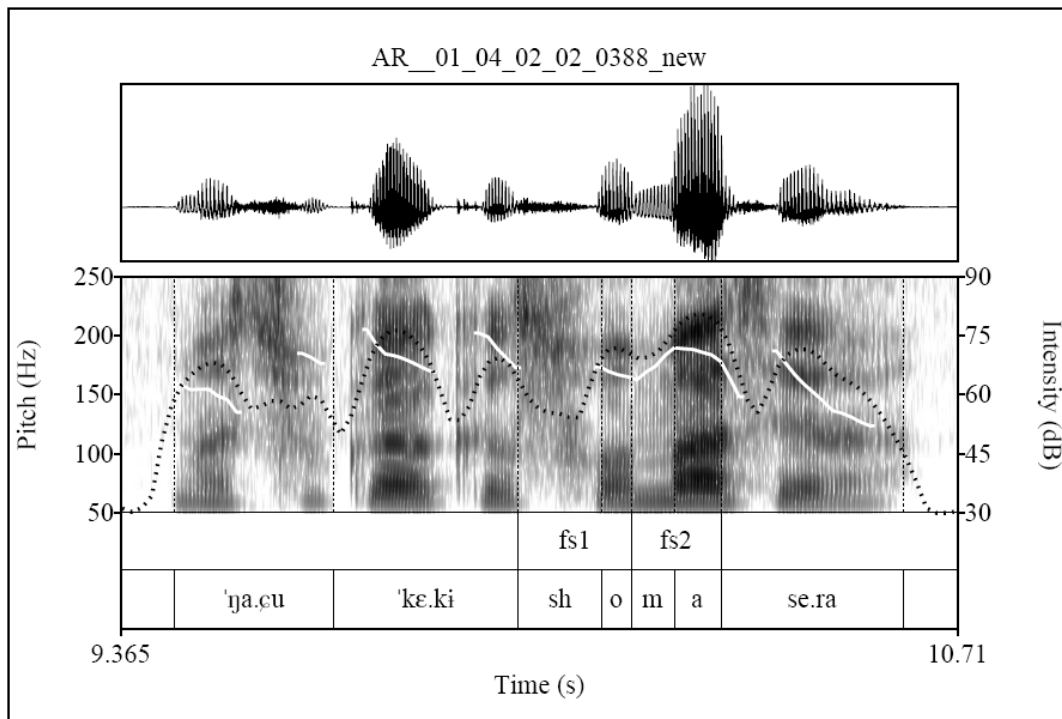
For [s^ho.ma] *so.ma* 'new', on the other hand, there is a greater difference between isolation and frame forms. In isolation – as shown in Figure 5.45 below – pitch slope was measured as -3.6 Hz/100msec in σ_1 , and -28 Hz/100msec in σ_2 . The contrast in slopes across syllables is -25 Hz/100msec. The contrast in pitch across syllables for this word is 13 Hz, so it, too, falls in the lower right quadrant in Figure 5.42, at coordinates (13, -25)

Figure 5.45 AR_04 / Adjectives / Pitch trace: [s^ho.ma] *so.ma* ‘new’



When this same word occurs in the sentence frame, the slopes are actually steeper in both syllables than in the isolation forms, and the relationship between them is different. Pitch slope was measured as $-18 \text{ Hz}/100\text{msec}$ in σ_1 , and $-15 \text{ Hz}/100\text{msec}$ in σ_2 , for a slope difference of $3 \text{ Hz}/100\text{msec}$. The positive difference means that pitch slopes “more downward” in σ_1 than in σ_2 ; this does not indicate a correlation with σ_2 stress. Pitch, on the other hand, is more prominent on the stressed syllable: $\Delta\text{pitch} = 19 \text{ Hz}$. This point falls in the upper right quadrant in the graph on the right in Figure 5.42 above, at coordinates (19, 3).

Figure 5.46 AR_04 / Adjectives / Frame / Pitch trace: [s^ho.ma] *so.ma* ‘new’



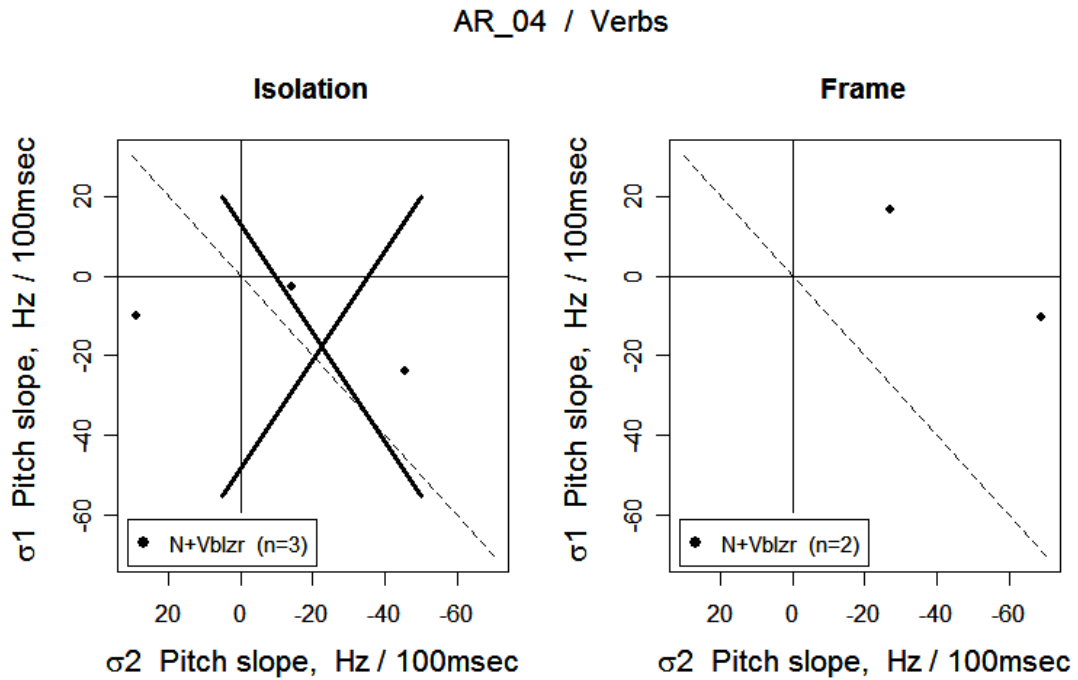
What these examples have shown is that pitch slope is a robust cue for σ_2 stress in isolation forms of AR_04 adjectives, but does not show a correlation with stress in the frame forms at all. In the frame forms, the other reflex of fundamental frequency – pitch – plays a stronger role.

5.3.1.3 AR_04 Verbs

As discussed previously, speaker AR_04 was not able to comfortably produce verbs in isolation, so I do not discuss them here. The two frame forms are plotted in the graph at right in Figure 5.47 below. Both verbs fall to the right of the y-axis, indicating that pitch always slopes downward in σ_2 . However, the sample is really

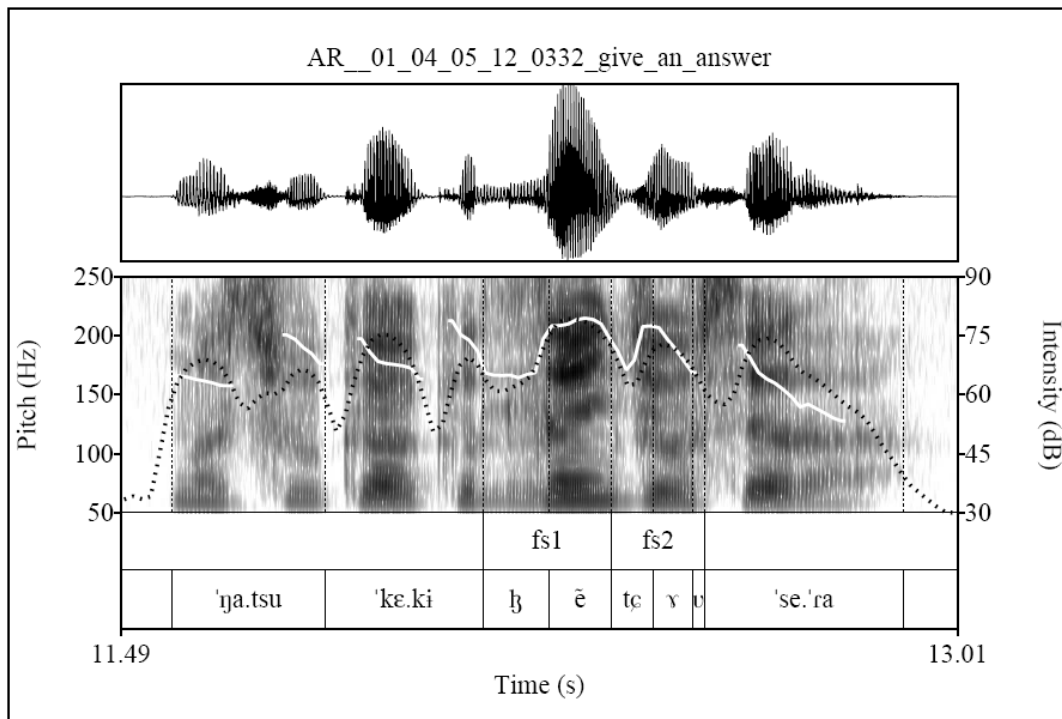
too small to draw any firm conclusions from this – and even if the sample were larger, the trend suggested here cannot be distinguished from the contrast in pitch across syllables.

Figure 5.47 AR_04 / Verbs: Pitch slope



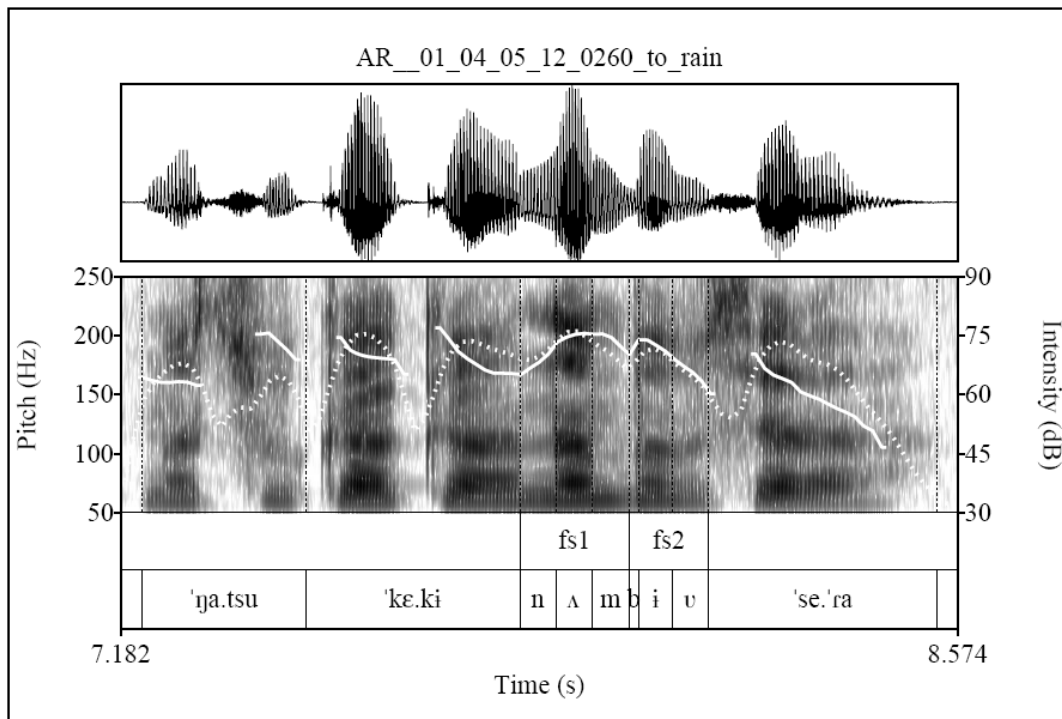
The point which plots below the x-axis is [ʔʒẽ.tɕɯ] *lan.brgyab* ‘to answer’, shown in Figure 5.48 below. For this verb, pitch slopes downward in both syllables, but downward more steeply in $\sigma 2$ than in $\sigma 1$.

Figure 5.48 AR_04 / Verbs / Frame / Pitch trace: ['ḷ̥ḷ̥.tɕʰʊ] *lan.brgyab* 'to answer'



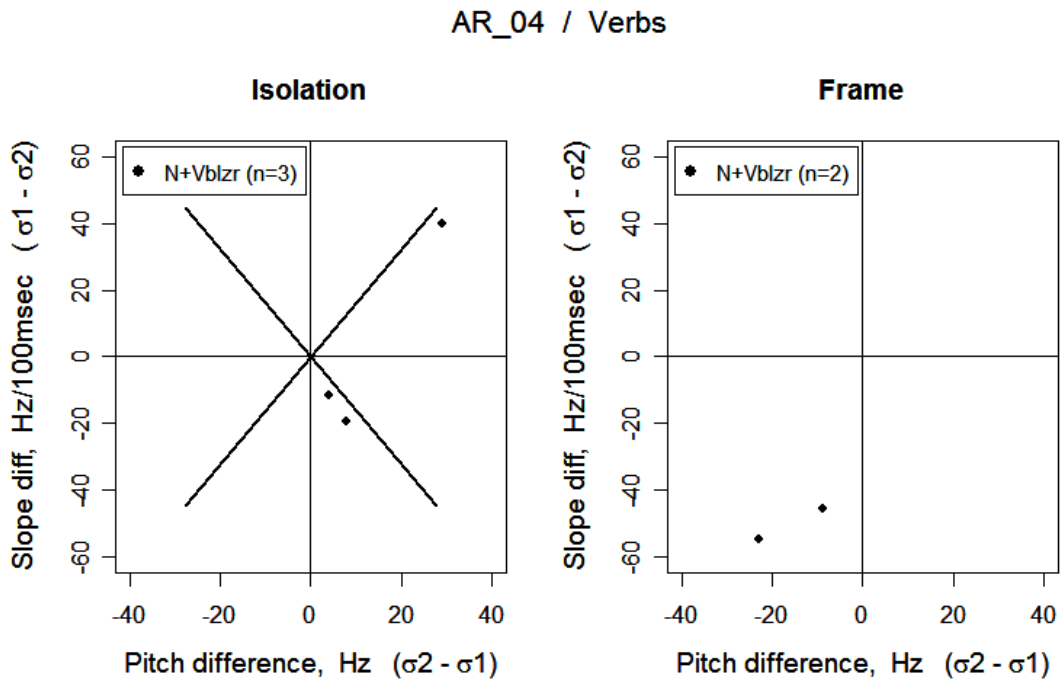
The point above the x-axis – with an upward slope in σ_1 and a downward slope in σ_2 is ['nɒm.biʊ] *gnam.babs* 'to rain', shown in Figure 5.49 below.

Figure 5.49 AR_04 / Verbs / Frame / Pitch trace: ['nʌm.biʊ] *gnam.babs* 'to rain'



In Figure 5.50 below, the difference in pitch slope across syllables is plotted against the difference in pitch across syllables. Both verbs fall in the lower left quadrant. As demonstrated in section 5.2.1.3, pitch is a robust correlate of σ_1 stress for AR_04 verbs, significantly higher in σ_1 than in σ_2 . Pitch slope is an indistinguishable outcome of this pitch contrast; as pitch falls from the first syllable to the second, pitch slopes downward.

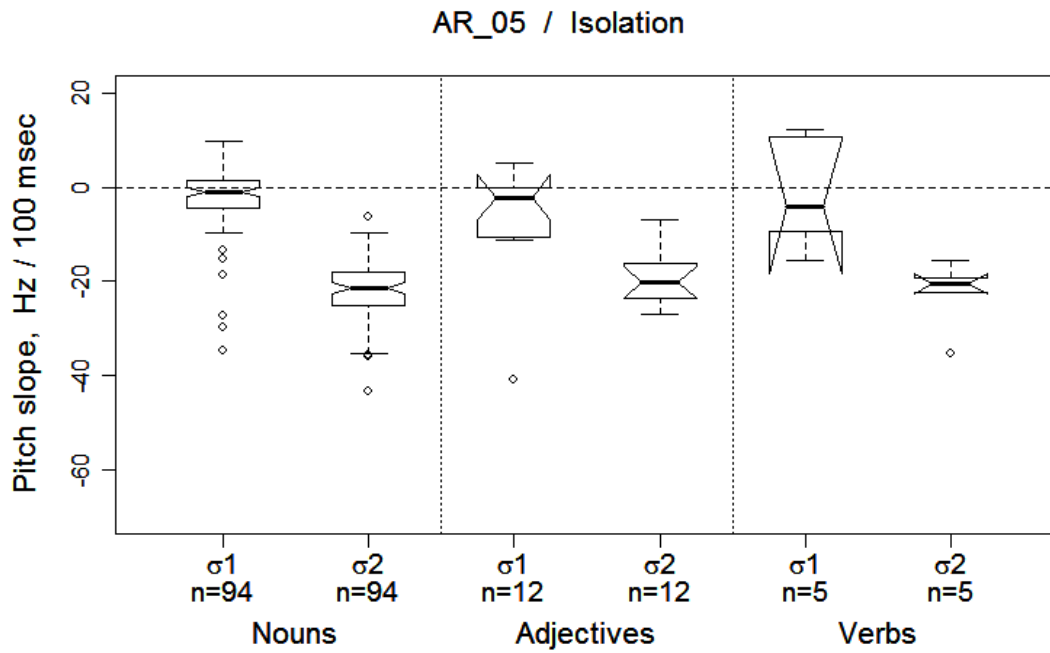
Figure 5.50 AR_04 / Verbs: Pitch difference vs. pitch slope difference



5.3.2 Pitch slope for speaker AR_05

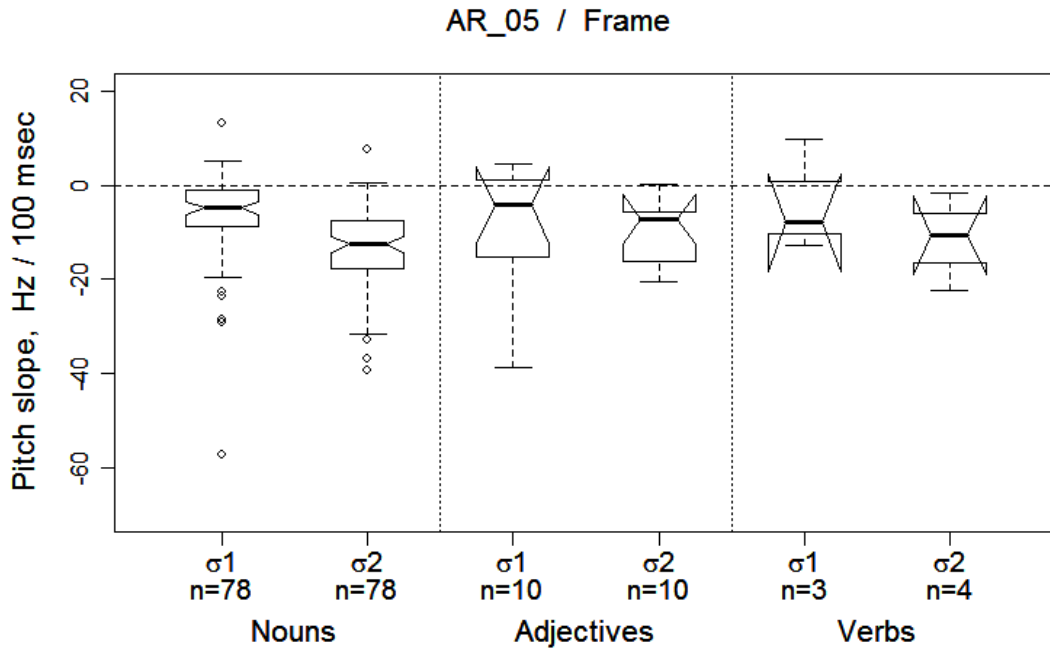
Pitch slope contrasts for the isolation forms of AR_05 nouns, adjectives, and verbs are illustrated by the box-and-whisker plots in Figure 5.51 below. For nouns and adjectives, pitch in σ_1 usually slopes downward, though it sometimes slopes upward. In σ_2 , pitch always slopes downward. As was the case with speaker AR_04, there is a considerable contrast in pitch slope across syllables, with no overlap of the notches or the boxes. Verbs show a similar pattern – though with more variation in σ_1 – but here it is an incidental effect of the drop in pitch across syllables.

Figure 5.51 AR_05 / Isolation / Pitch slope contrasts



In the frame forms, illustrated in Figure 5.52 below, the patterns are similar, but more subdued. For adjectives, measurements on the two syllables show considerable overlap.

Figure 5.52 AR_05 / Frame / Pitch slope contrasts



Pitch slope data for AR_05 nouns, adjectives, numerals, and verbs is considered in greater detail in sections 5.3.2.1 through 5.3.2.4 above.

5.3.2.1 AR_05 Nouns

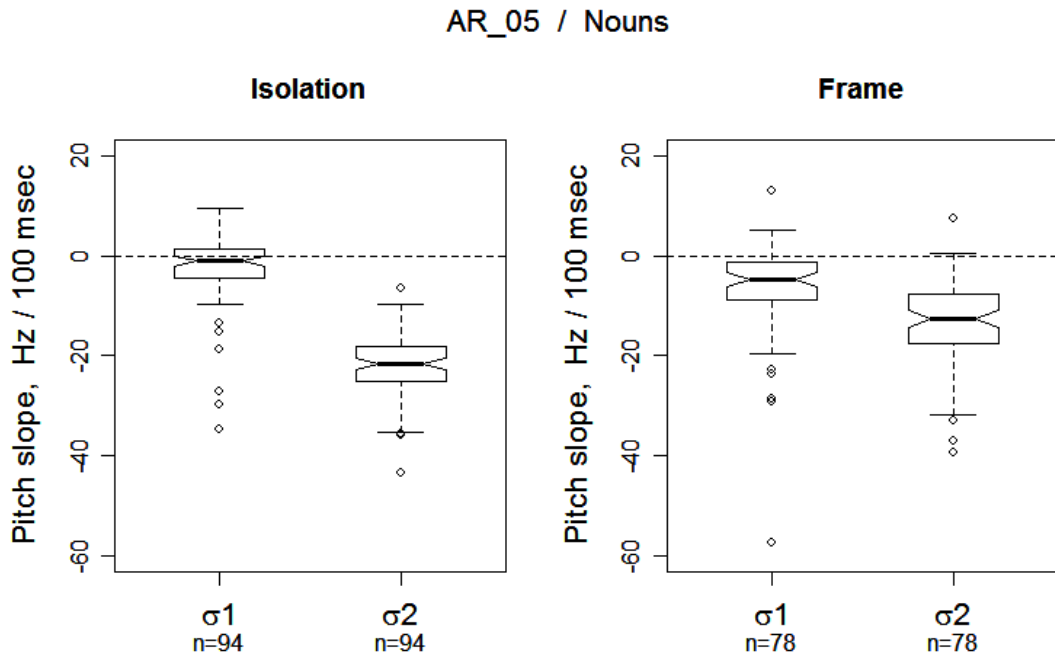
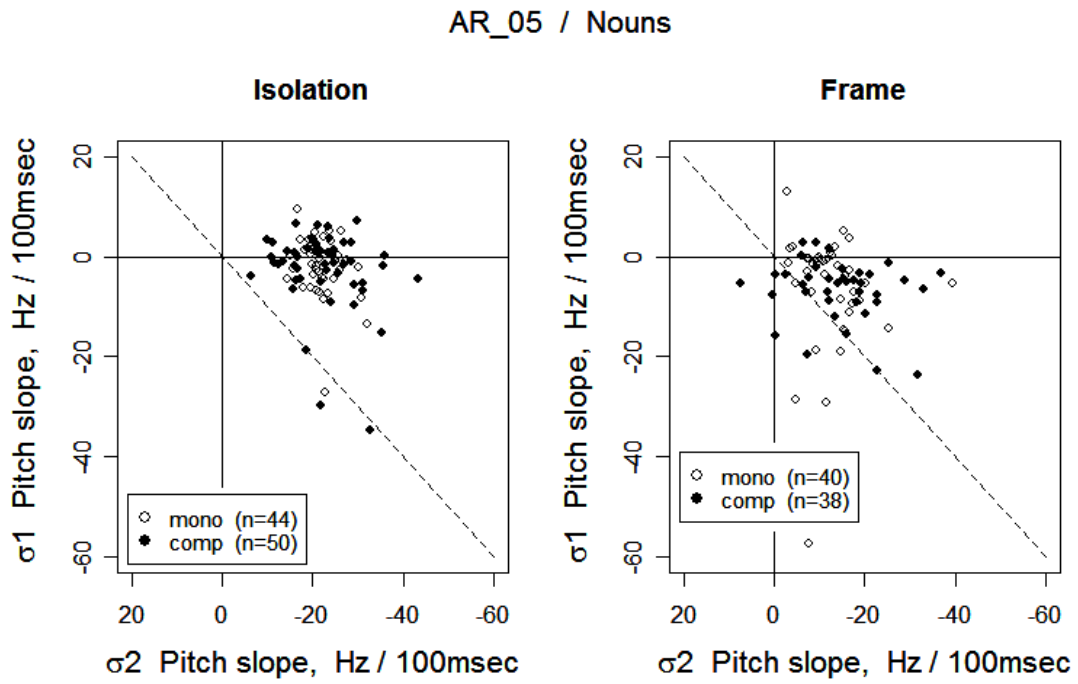
Pitch slope is a robust cue for σ_2 stress in AR_05 nouns produced in isolation, and a weaker correlate of stress for nouns produced in the sentence frame. This was also the case for nouns produced by speaker AR_04.

Slope measurements are plotted in Figure 5.53 below. The isolation forms define a distinct cluster, with nearly all points falling to the right of the dashed line and centered on the x-axis. In σ_1 , pitch slopes upward in about half the sample, and downward in about half the sample. In σ_2 , pitch always slopes downward, and always

“more downward” than in $\sigma 1$ (i.e., to the right of the dashed line). Only a few exceptions fall close to the dashed line, with a minimal slope contrast across syllables. (The three points closest to the dashed line are [xkʌŋ.'lʌm] ‘*rkang.lam* footpath’, [tɕʰor.'tʰin] *mchod.rten* ‘chorten’, and [xtiɾ.'kʌ] *sprul.sku* ‘incarnation, incarnate lama’.)

In comparison, points representing the frame forms – shown in the graphs on the right below – are more scattered, and there are fewer cases in which pitch slopes upward in $\sigma 1$ (i.e., fewer points above the x-axis). Still, most of the frame forms plot to the right of the dashed line, indicating that pitch slopes “more downward” in $\sigma 2$ than in $\sigma 1$. (The exceptions which fall to the left of the dashed line include [ʂo.'mʌŋ] *sro.ma* ‘nit, lice egg’ left of the y-axis, [kʌ.'pa] *skud.pa* ‘thread’ at ($\sigma 2$, $\sigma 1$) coordinates (-7.8, -57) and [tɕʰʌŋ.'go] *chu.mgo* ‘water source, spring’ at (0.45, -7.7).)

Figure 5.53 AR_05 / Nouns: Pitch slope



The difference in pitch slope across syllables is statistically significant for nouns in both settings, as indicated by paired-sample t-tests summarized in Table 5.14 below. In both cases, $p \ll 0.05$. For the isolation forms, the mean difference in slope across syllables is -20 Hz/100msec, with a narrow 95% confidence interval. Based on my own listening, this is a highly perceptible contrast. For the frame forms, the mean difference in slope across syllables is only -6.8 Hz/100msec; but a statistically significant mean slope difference can be as small as only -4.1 Hz/100msec (the upper 95% confidence limit), which is probably not of great perceptual significance. Thus I consider pitch slope to be a robust cue for σ_2 stress in the isolation forms, and a weaker correlate of stress for frame forms.

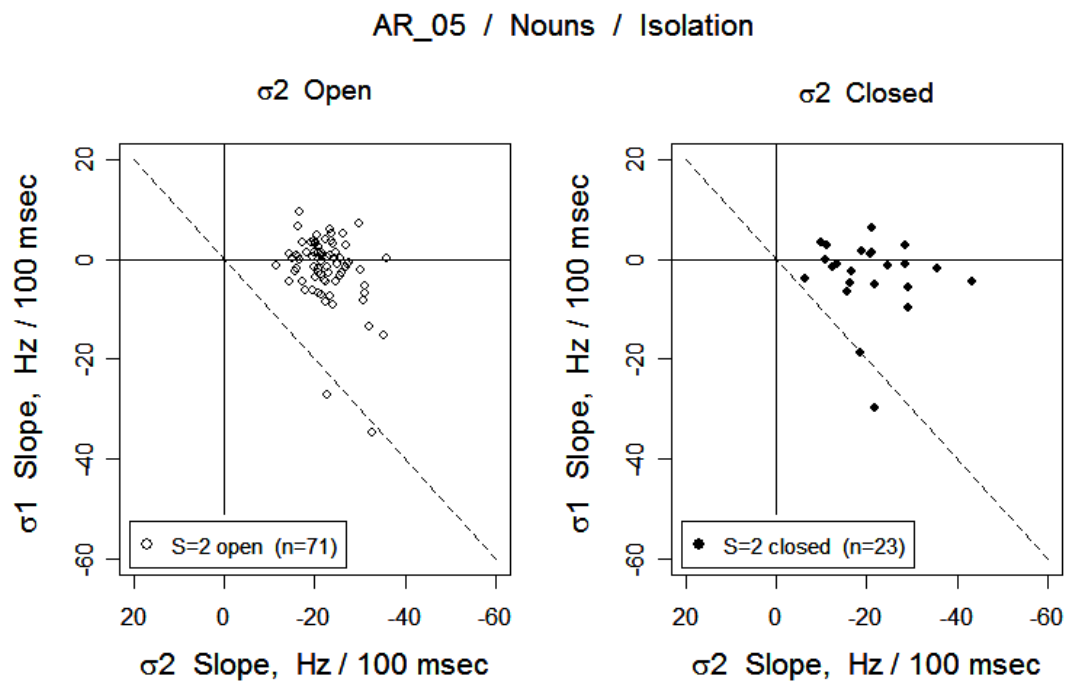
Table 5.14 AR_05 / Nouns / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_05	Isol.	-23.3503	93	< 2.2e-16	-20	-22	-18
	Frame	-5.04	77	3.005e-06	-6.8	-10	-4.1

In Figure 4.20 below, the isolation forms of AR_05 nouns are plotted in terms of σ_2 closure. As shown, σ_2 -open nouns define a fairly tight cluster centered on the x-axis, while σ_2 -closed nouns are dispersed more broadly. (This variation occurs because in some closed syllables, pitch slopes downward at the same rate over both the vowel and the following coda, while in other cases most of the slope may be carried by the coda.) In both cases, all points fall to the right of the dashed line. (As

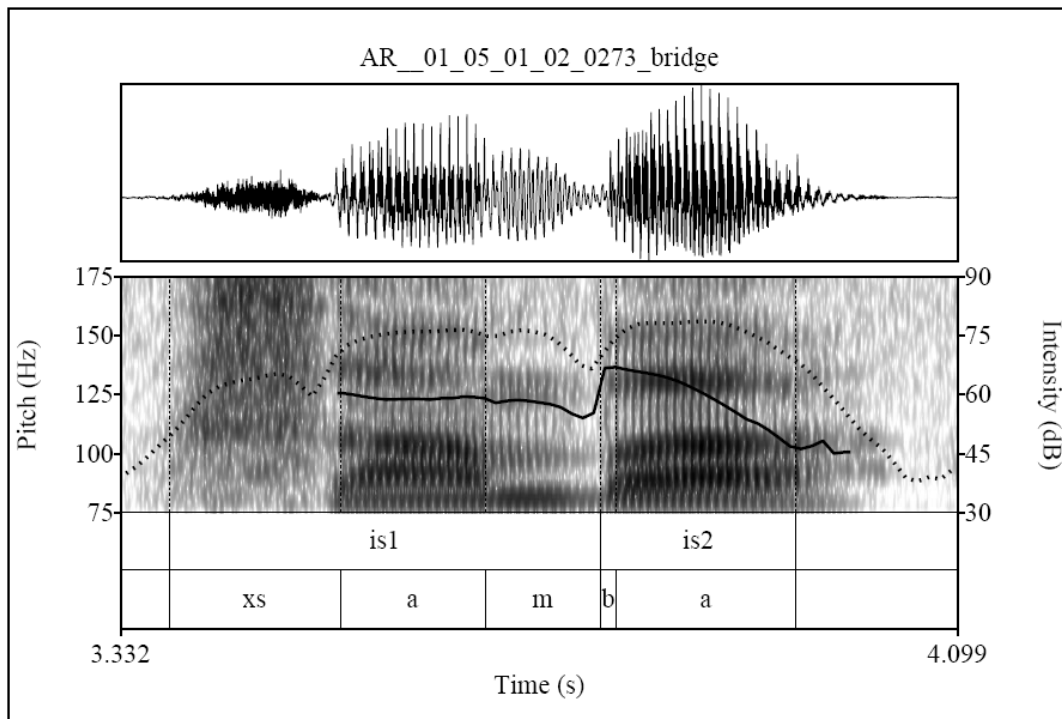
discussed in section 4.3.2.1 in the previous chapter, this contrasts with the patterns exhibited by BM_01. There, points representing nouns with σ_2 closed fell to either side of the dashed line; pitch slope there reflects a falling intonation contour, which is interrupted by a closed σ_2 .)

Figure 5.54 AR_05 / Nouns / Isolation: Pitch slope vs. σ_2 closure



A prototypical example of the σ_2 -open type is [xsam.'ba] *zam.ba* ‘bridge’; the pitch trace for this word is shown in Figure 5.55 below.

Figure 5.55 AR_05 / Nouns / Isolation / Pitch trace: [xsam.'ba] *zam.ba* 'bridge'

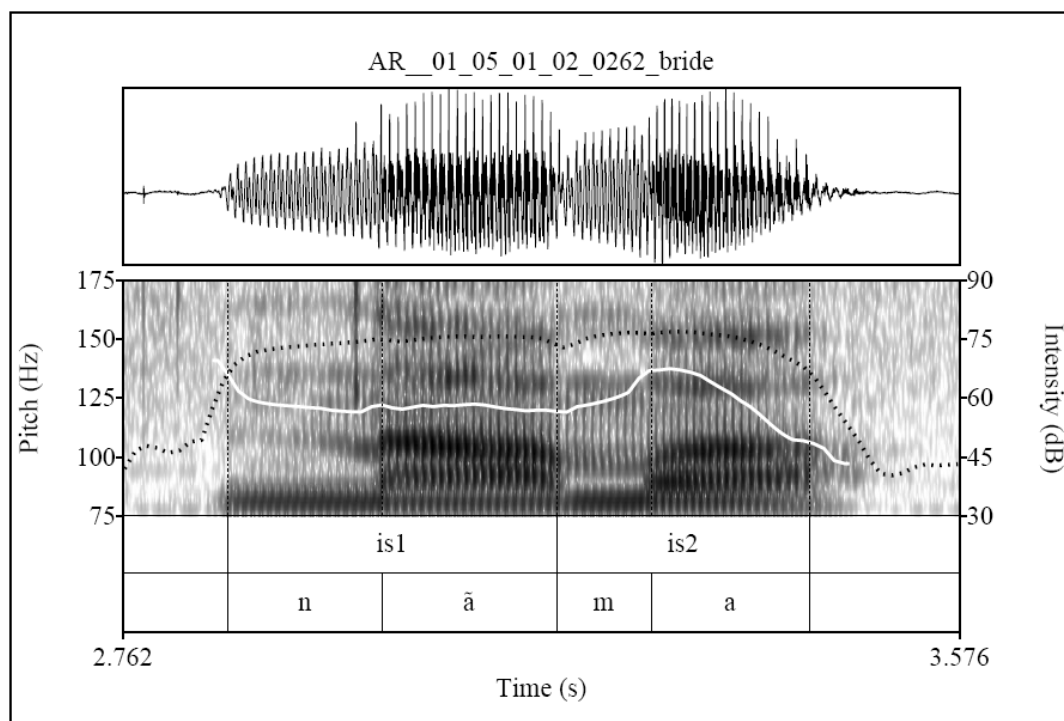


As was the case in AR_04 nouns, when the consonant at the syllable boundary is a stop, pitch leaps to a maximum at the beginning of σ_2 , and immediately begins to slope downward at a constant rate. In [xsam.'ba] *zam.ba* 'bridge', slope in σ_1 is -1.8 Hz/100msec, slope in σ_2 is -21 Hz/100msec, so the slope difference is -19 Hz/100msec. This is quite a noticeable contrast. Average pitch, on the other hand, shows almost no contrast across syllables: pitch in σ_1 is 123 Hz while pitch in σ_2 is 124 Hz, so the difference in pitch is a barely perceptible 1 Hz. (There is no contrast in vowel quality across syllables; Δ intensity is 2 dB, and Δ vowel duration is 32 msec.)

When the consonant at the syllable boundary is a sonorant, the pitch slope in σ_2 rises more gently, as illustrated for [na.'ma] *mna'.ma* ‘bride’ in Figure 5.56 below.

Here, the difference in pitch slope across syllables is -18 Hz/100 msec. (Δ pitch = 5 Hz, Δ intensity = 0 dB, Δ vowel duration = -16 msec.)

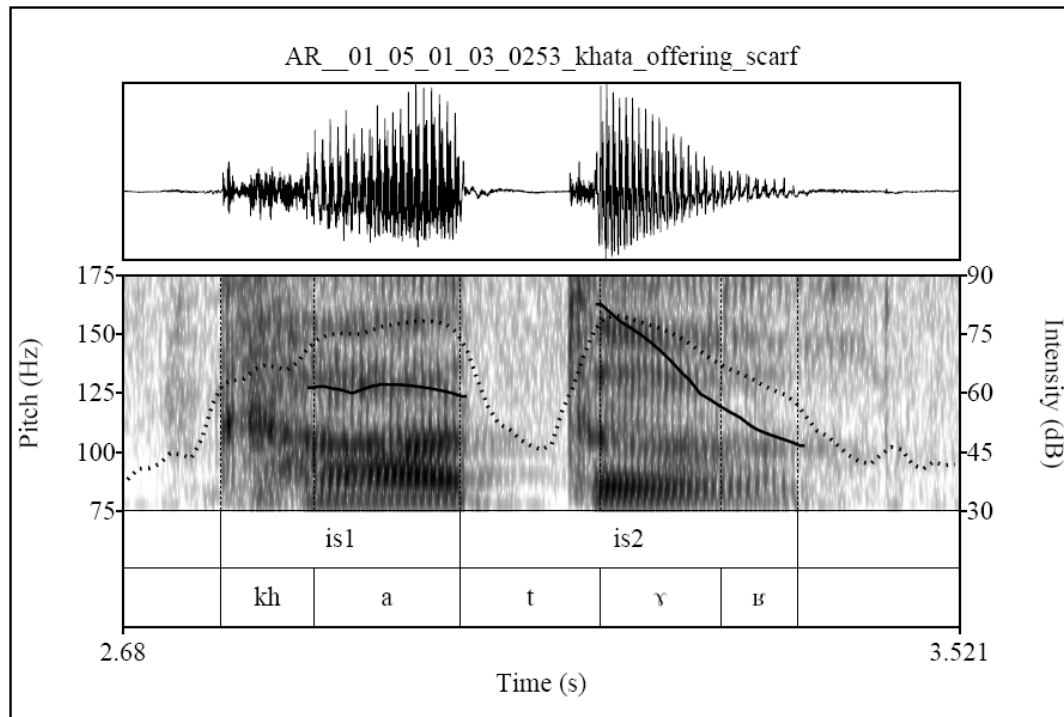
Figure 5.56 AR_05 / Nouns / Isolation / Pitch trace: [na.'ma] *mna'.ma* ‘bride’



When σ_2 is closed – which can only occur in compound nouns, plotted in the graph on the right in Figure 4.20 – pitch slope in σ_2 is more variable. In [k^ha.'tʁʁ] *kha.btags* ‘khata, offering scarf’, the slope in σ_2 is steeper than is typically seen in the monomorphemic / σ_2 open nouns. This is clear from the pitch trace in Figure 5.57

below. Here, pitch slope is only -1.8 Hz/100msec in σ_1 , but a notable -36 Hz/100msec in σ_2 , for a slope difference of -34 Hz/100msec.

Figure 5.57 AR_05 / Nouns / Isolation / Pitch trace: [k^ha.'tʏʏ] *kha.btags* ‘khata, offering scarf’

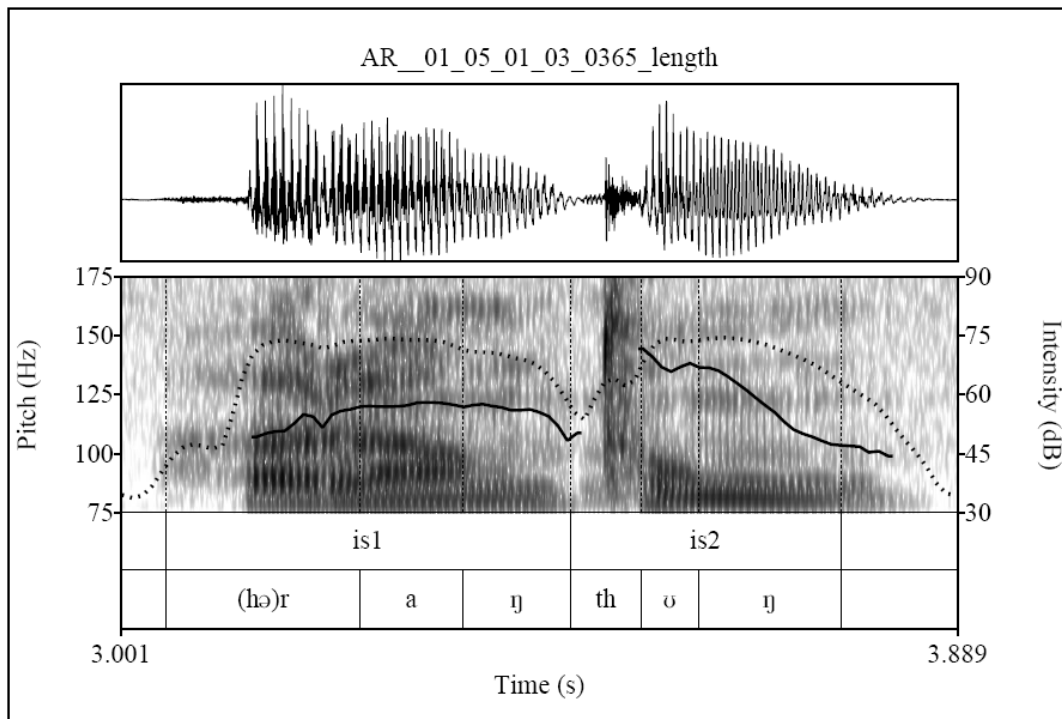


In this case, there is also a clear increase in average pitch across syllables, from 128 Hz in σ_1 to 142 Hz σ_2 , for an increase of 14 Hz. As discussed in section 5.2.2.1, pitch is usually exaggerated in σ_2 when it is closed. The explanation for this is clear from the pitch trace above: the downward pitch slope in σ_2 occurs over the entire span of the rhyme; if we look only at the vowel, only the highest portion of this slope is represented. (This may not always effect pitch *slope*, though: the vowel may – or may not – exhibit exactly the same slope as the entire rhyme. Even if the vowel

is short, the rate of change of pitch may be the same for this shorter span – the vowel – as it is for the longer span – the rhyme.) For the word here, the average pitch over the full rhyme is 130 Hz; this is nearly the same as the average pitch of 128 Hz in $\sigma 1$. (For this word Δ intensity = 0 dB and Δ vowel duration = -25 msec.)

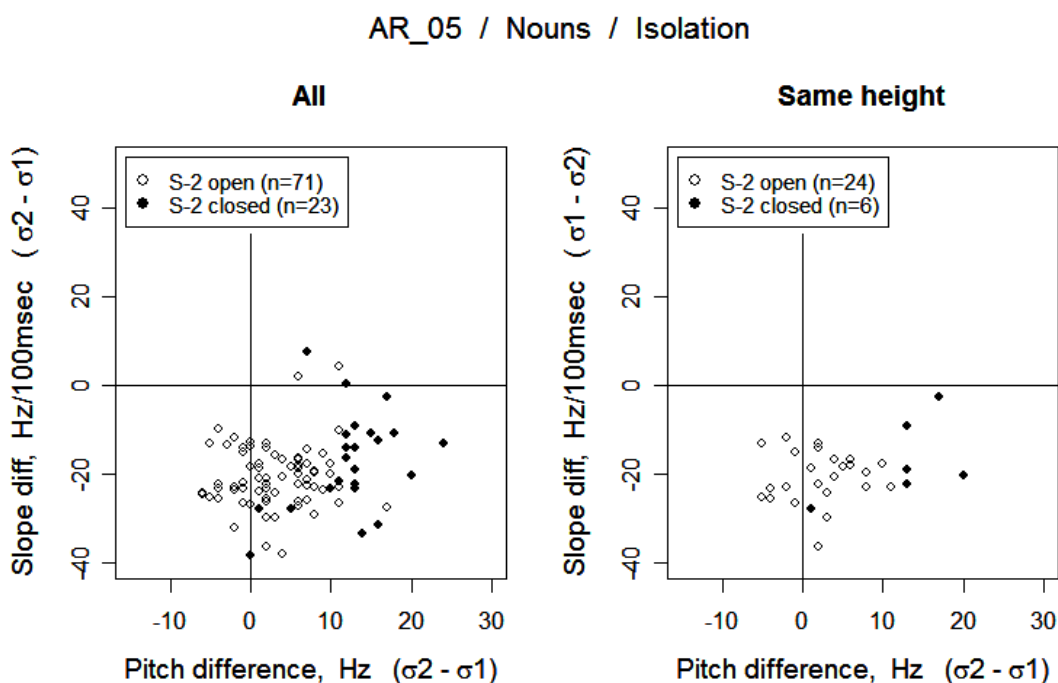
Another example is [^hraŋ.t^huŋ] *ring.thung* ‘length’, shown in Figure 5.58 below. Here, pitch is unstable in the $\sigma 2$ vowel, and most of the pitch slope is carried by the nasal coda. Thus slope in the $\sigma 2$ vowel is more gentle than is typically seen in the monomorphemic / $\sigma 2$ open nouns, measured as only -11 Hz/100msec. In $\sigma 1$, pitch slope is 0.10 Hz/100 msec, so the difference in slope is -11 Hz/100msec. The difference in average pitch across syllables is 15 Hz if the $\sigma 1$ and $\sigma 2$ vowels are compared, but only 4 Hz if the $\sigma 1$ vowel is compared to the $\sigma 2$ rhyme – which is the more appropriate comparison. (Δ intensity = 0 dB, Δ vowel duration = -47 msec.)

Figure 5.58 AR_05 / Nouns / Isolation / Pitch trace: [h^oraŋ.t^hoŋ] *ring.thung* ‘length’



The relationship between pitch slope and pitch for the AR_05 isolation forms is illustrated in Figure 5.65 below. The graph at right below shows only nouns which have vowels of the same height in both syllables, thus controlling for the potential incidental effect of vowel height contrasts on pitch. Within this group, the subset of nouns with σ_2 open – controlled for the variation in pitch which occurs when σ_2 is closed – is centered on the y-axis and falls distinctly below the x-axis. This graph shows that pitch *slope* is a robust and distinctive correlate of σ_2 stress, though *pitch* – as concluded in section 5.2.2.1 – is only a limited and very weak correlate.

Figure 5.59 AR_05 / Nouns / Isolation: Pitch difference vs. pitch slope difference



The fact that nouns in the control group cluster nearly symmetrically about the vertical “pitch difference = 0” line suggests that speaker AR_05 is really not even aiming to create a pitch contrast across syllables at all. Instead, the speaker’s aim is to create a contrast in pitch slope of between ~ -10 Hz/100msec and ~ -30 Hz/100msec. (Note that the pattern here is *not* the same as that in Figure 4.20. The graph there shows that, in isolation forms with σ_2 open, the speaker is aiming for a flat slope in σ_1 and a steep downward slope in σ_2 . The graph here shows that, in isolation forms with σ_2 open, the speaker is aiming for a neutral pitch contrast across syllables, but a strong pitch slope contrast.)

In summary, pitch slope is a robust correlate of σ_2 stress in the isolation forms of AR_05 nouns. Pitch and pitch slope do not reinforce one another as cues for stress

(which was the case for AR_04 isolation forms). Instead, the speaker seems to manipulate fundamental frequency in order to create a contrast in pitch slope across syllables, but not a contrast in pitch.

In the frame forms of AR_05 nouns, the role of pitch slope as an acoustic correlate of σ_2 stress is more muted. As shown in Table 5.14 on page 404, the difference in slope across syllables is, indeed, statistically significant ($p \ll 0.05$), but it is probably not perceptually highly salient: the mean slope difference is only -6.8 Hz/100msec (compared to a mean difference of -20 Hz/100msec for the isolation forms), and the 95% confidence limit ranges from -10 Hz/100msec to as little as -4.1 Hz/100msec. Fundamental frequency is thus manipulated to emphasize the contrast in pitch across syllables, rather than the contrast in pitch slope.

The frame forms of AR_05 nouns are plotted in terms of morphological composition and in terms of σ_2 closure, respectively, in Figure 5.60 and Figure 5.61 below. As was the case with the isolation forms, these two factors appear to have some effect on pitch slope; slopes in σ_2 seem to be steeper in the compound / σ_2 closed subsets than in the monomorphemic / σ_2 open subsets. (It seems, from Figure 5.61, that most of the nouns falling to the left of the dashed line have an open σ_2 . I am not sure why this should be so; it may be that more nouns with σ_2 closed would fall in this area, too, but the sample is too small to reveal this.)

In contrast to the isolation forms, most points plot below the x-axis, meaning that slope is only rarely upward in σ_1 . Still, most points fall to the right of the dashed

line, meaning that even if pitch slopes downward in both syllables, it slopes downward more steeply in σ_2 .

Figure 5.60 AR_05 / Nouns / Frame: Pitch slope vs. morphological composition

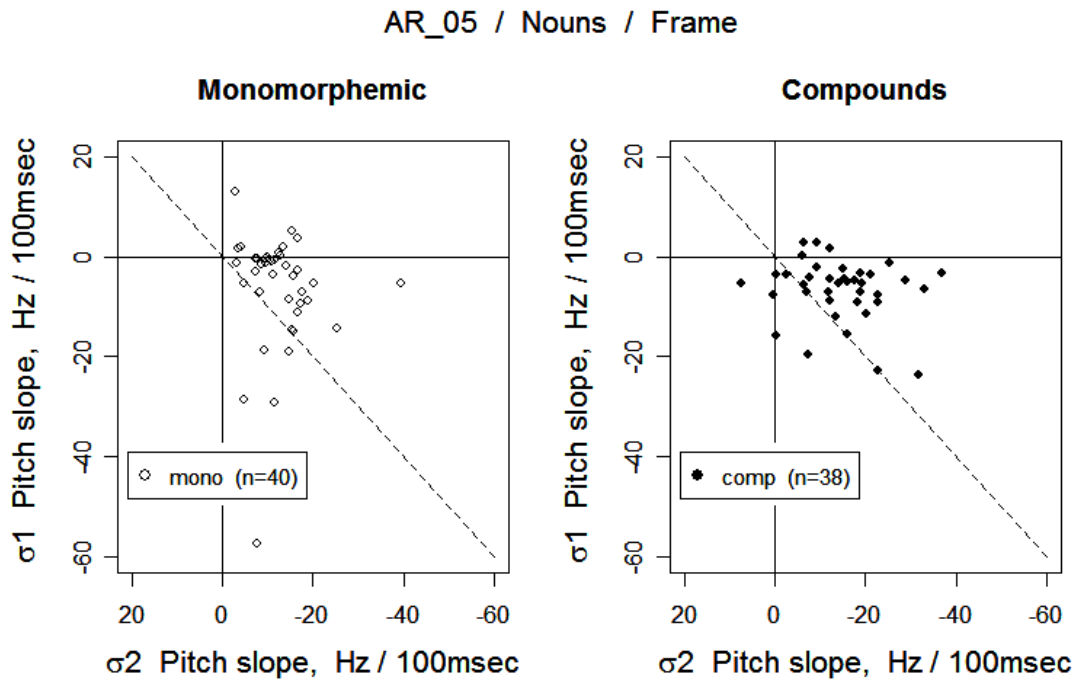
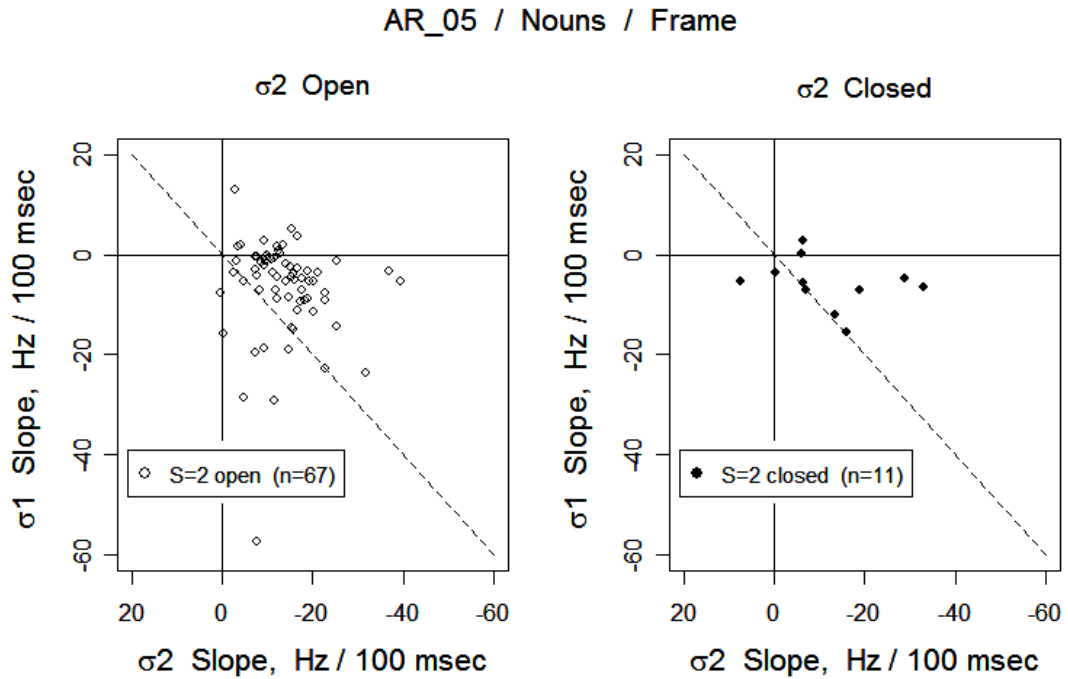
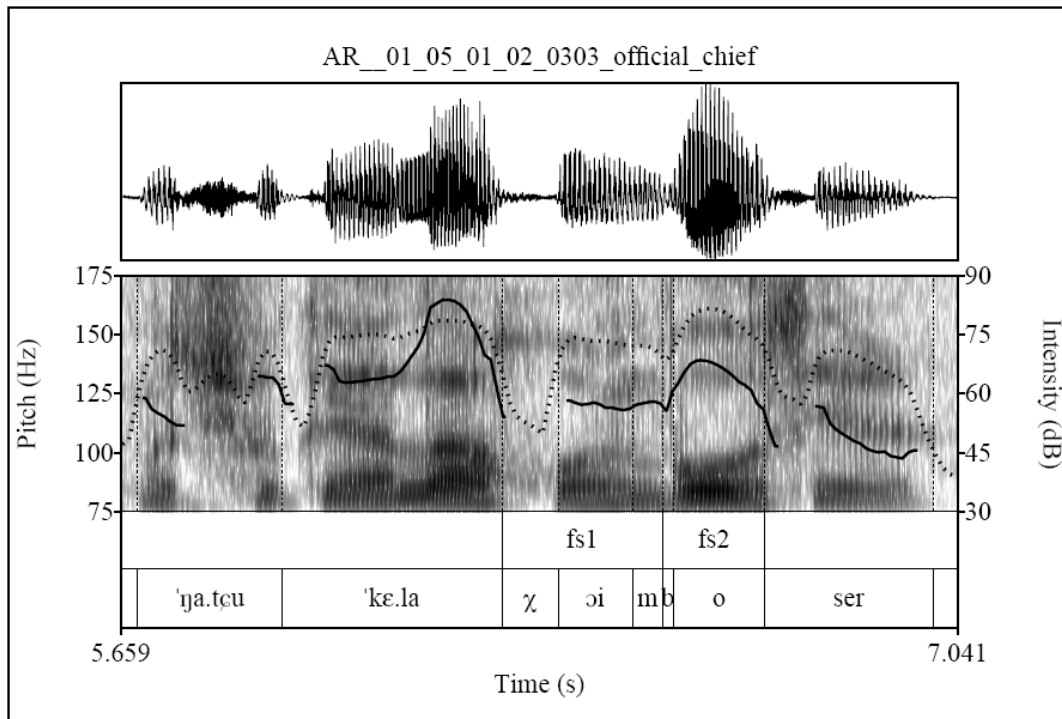


Figure 5.61 AR_05 / Nouns / Frame: Pitch slope vs. σ_2 closure



An example of the σ_2 open type is the monomorphemic noun [$\chi\omicron\text{im}.\text{'bo}$] *dpon.po* ‘official, chief’, shown in Figure 5.62 below. Here, pitch slope was measured as -1.8 Hz/100 msec in σ_1 , and -14 Hz/100msec in σ_2 , for a difference of -12 Hz/100msec. Pitch, intensity, and vowel duration also contribute to the perception of stress on σ_2 in this case: $\Delta\text{pitch} = 16$ Hz; $\Delta\text{intensity} = 7$ dB; $\Delta\text{vowel duration} = 26$ msec.

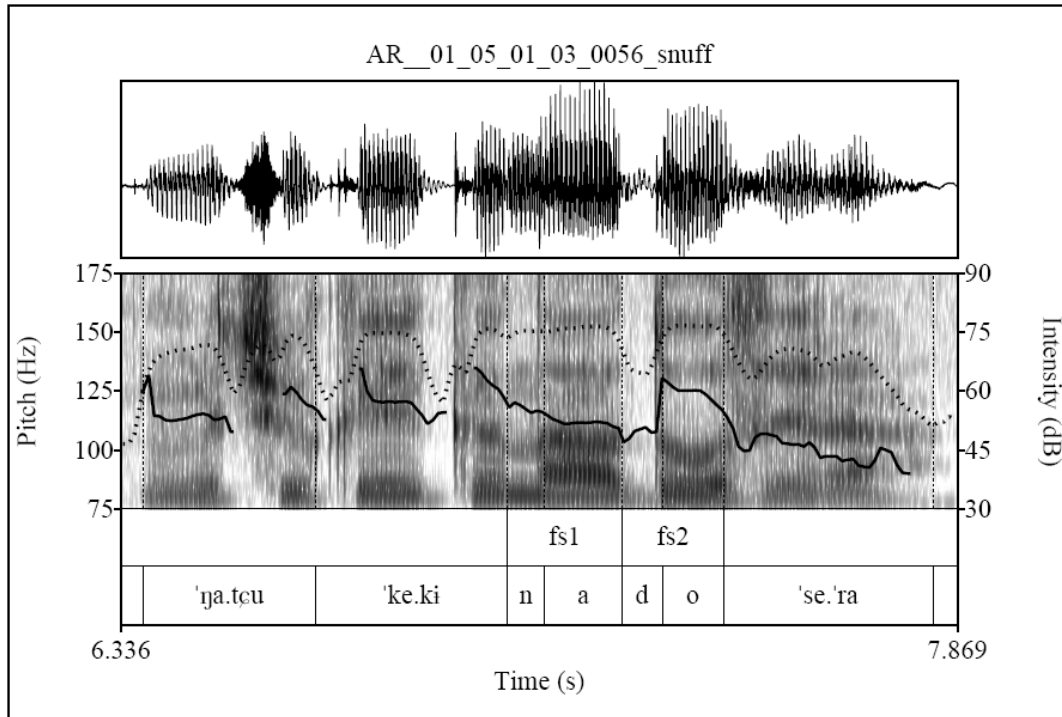
Figure 5.62 AR_05 / Nouns / Frame / Pitch trace: [$\chi\omicron\text{im}.\text{'bo}$] *dpon.po* ‘official, chief’



Another example is the compound noun [$\text{na}.\text{'do}$] *sna.tha? sna.tho? sna.rdul?*, shown in Figure 5.63 below. In this word, the difference in pitch slope across syllables is -7.9 Hz/100msec. Pitch and intensity also correlate with stress, though

vowel duration does not: $\Delta\text{pitch} = 13 \text{ Hz}$; $\Delta\text{intensity} = 1 \text{ dB}$; $\Delta\text{vowel duration} = -30 \text{ msec}$.

Figure 5.63 AR_05 / Nouns / Frame / Pitch trace: [na.'do] *sna.tha? sna.tho? sna.rdul?* 'snuff'

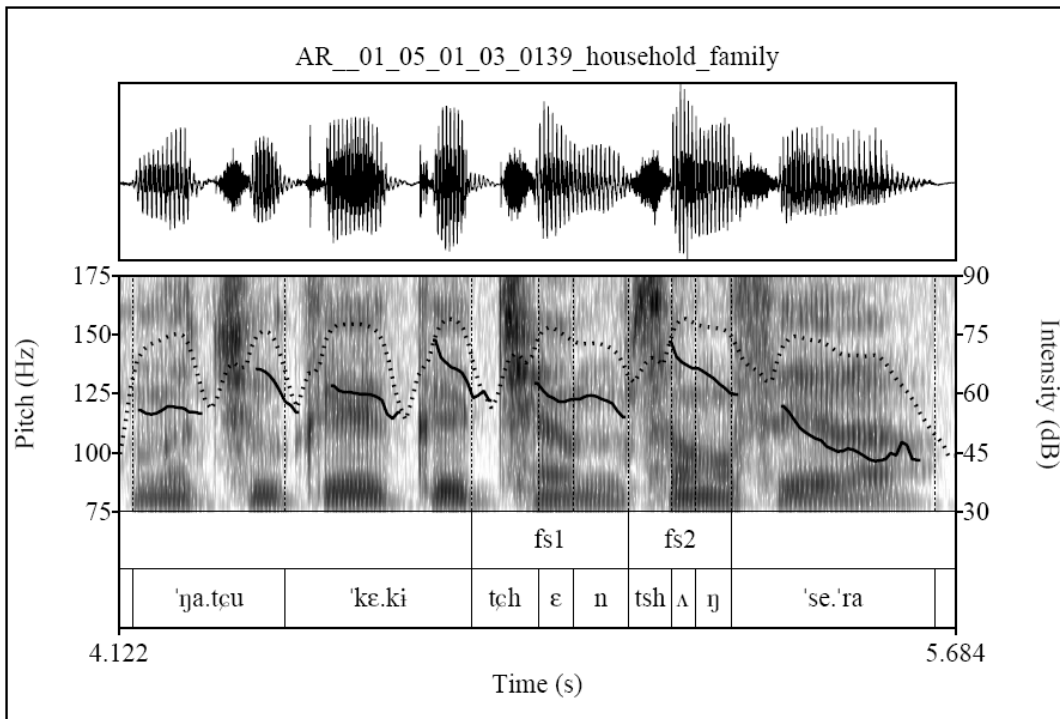


The compound noun [tɕɛn.'tsʰʌŋ] *khyim.tshang* 'household, family' is an example in which σ_2 is closed; the pitch trace for this word can be seen in Figure 5.64 below. Pitch slope is $-7.1 \text{ Hz}/100\text{msec}$ in σ_1 , and $-19 \text{ Hz}/100\text{msec}$ in σ_2 , for a slope difference of $-11 \text{ Hz}/100\text{msec}$. Unlike the isolation forms, here, the contrast in pitch across syllables reflects a correlation with stress, regardless of whether we consider the vowels or the rhymes: in σ_1 , pitch is 123 Hz in the vowel and 122 Hz over the

entire rhyme; in σ_2 , pitch is 138 Hz in the vowel, and 134 Hz over the entire rhyme. Either way, pitch is 12 to 15 Hz higher in σ_2 than in σ_1 .

Intensity also contributes to the perception of stress on σ_2 , though vowel duration does not: Δ intensity = 3 dB; Δ vowel duration = -22 msec.

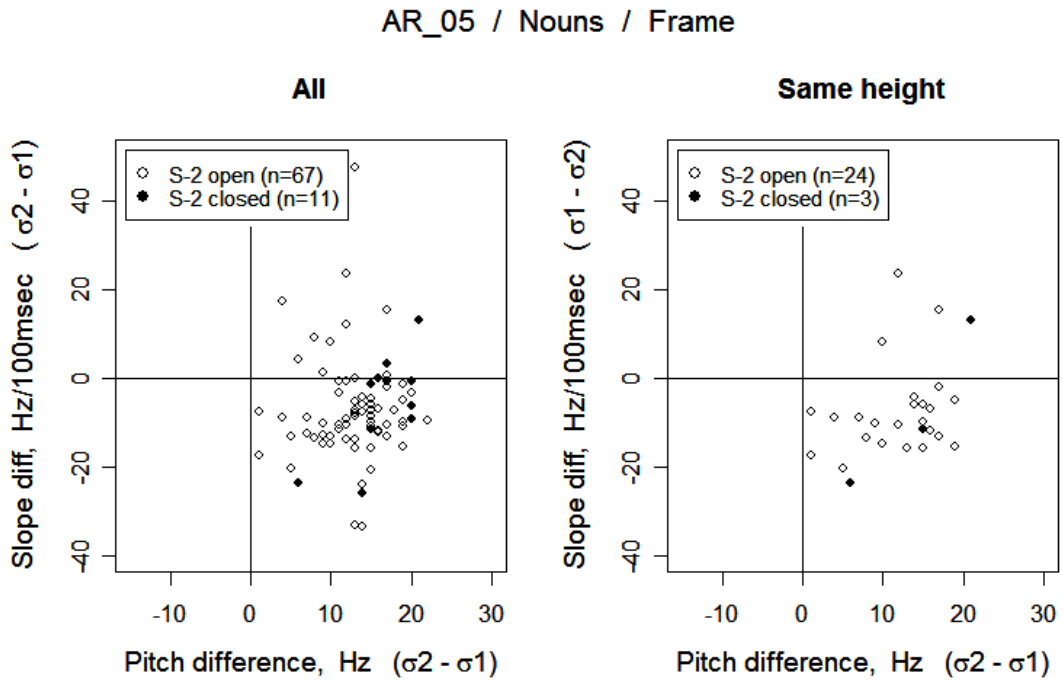
Figure 5.64 AR_05 / Nouns / Frame / Pitch trace: [tɕɛn.'tsʰʌŋ] *khyim.tshang* ‘household, family’



The relationship between pitch slope and pitch is illustrated in Figure 5.65 below. Since pitch can vary intrinsically as a function of vowel height, in the graph on the right below I show only nouns which have vowels of the same height in both syllables. For this speaker, vowel height does not significantly affect the distribution, compared to the graph on the left. In both graphs, most nouns plot in the lower right

quadrant. This means that pitch and pitch slope reinforce one another as acoustic correlates of stress.

Figure 5.65 AR_05 / Nouns / Frame: Pitch difference vs. pitch slope difference



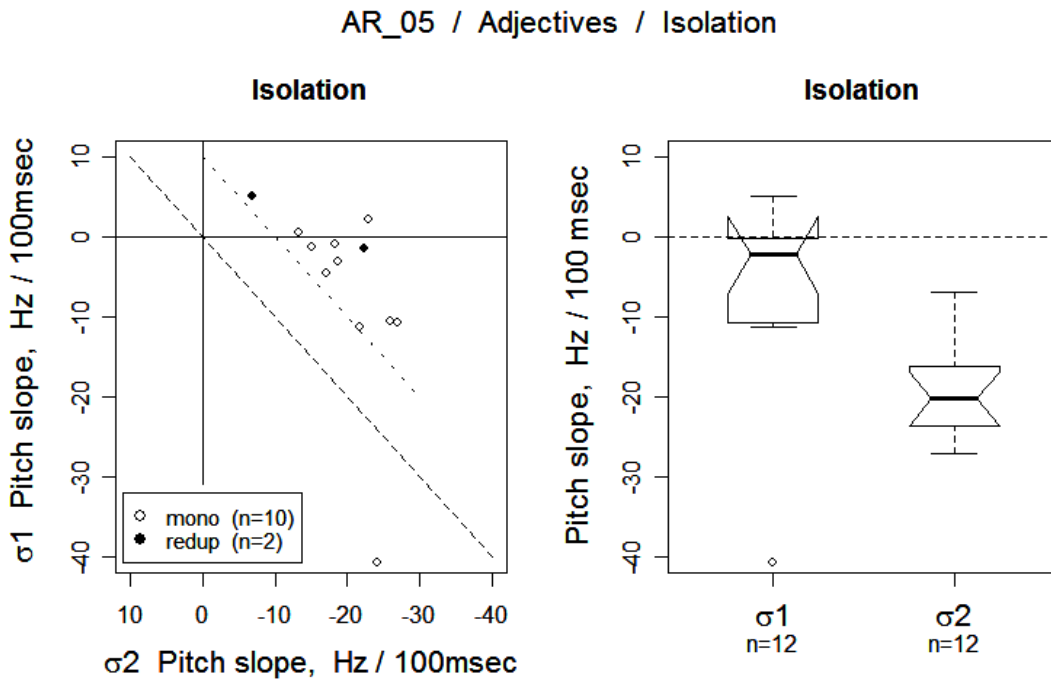
In summary, speaker AR_05 manipulates fundamental frequency in different ways in the isolation forms and frame forms of nouns. In the isolation forms, pitch slope is a robust cue for σ_2 stress, while the average pitch contrast across syllables is weak. In the frame forms, pitch slope is a weaker correlate of stress, but pitch plays a stronger role.

5.3.2.2 AR_05 Adjectives

As was the case for nouns, pitch slope is a robust cue for stress for the isolation forms of adjectives produced by speaker AR_05, but not for the frame forms.

The isolation forms are plotted in Figure 5.66 below. As shown, with the exception of [tʰɪ.ko] *phyug.po* ‘rich’, at (σ_2 , σ_1) coordinates (-24, -41), all points fall well to the right of the dashed line. This indicates that pitch slope is more prominent in σ_2 – i.e., slope is “more downward” in σ_2 than in σ_1 . The difference across syllables is always greater than -10 Hz/100msec, as indicated by the dotted reference line. In the boxplot on the right below, there is no overlap of either the notches – representing the 95% confidence interval about the median – or the boxes – representing the interquartile range.

Figure 5.66 AR_05 / Adjectives / Isolation: Pitch slope contrasts



A paired-sample t-test, summarized in Table 5.15 below, confirms that the difference in slope across syllables is statistically significant, with $p \ll 0.05$. The mean difference is -13 Hz/100msec; however, the lower 95% confidence limit of only -6.6 Hz/100msec may not be highly perceptible.

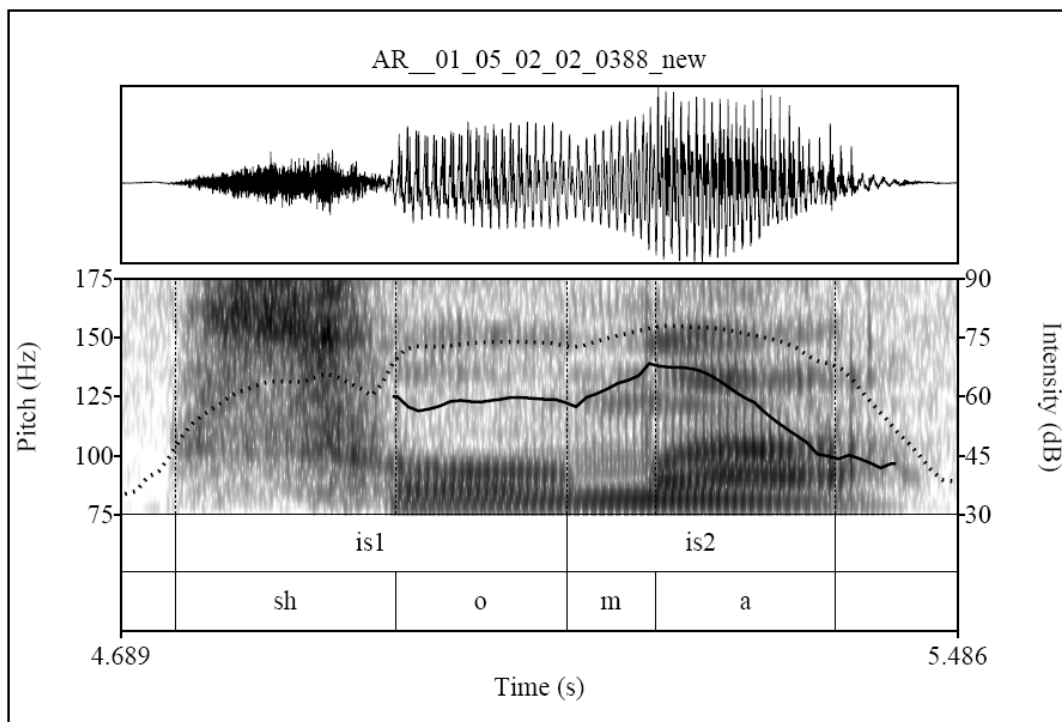
Table 5.15 AR_05 / Adjectives / Isolation / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_05	isolation	-4.4539	11	0.0009725	-13	-20	-6.6

As an example, the pitch trace for the word [s^ho.'ma] *so.ma* ‘new’ is provided in Figure 5.67 below. Pitch slopes slightly upward in σ_1 , at 2.24 Hz/100msec, and

downward in σ_2 at -23 Hz/100msec, for a difference in slope of -25 Hz/100msec. On the other hand, if a listener were attending to the average *pitch* in each syllable, this would *not* prove to be a useful cue for stress, as the increase across syllables is only 1 Hz: pitch over the medial 50% of the vowel was measured as 123 Hz in σ_1 and 124 Hz in σ_2 . The intensity difference across syllables is 2 dB, and the vowel duration difference is 8 msec. These are not large contrasts, leaving pitch slope as the primary cue for stress.

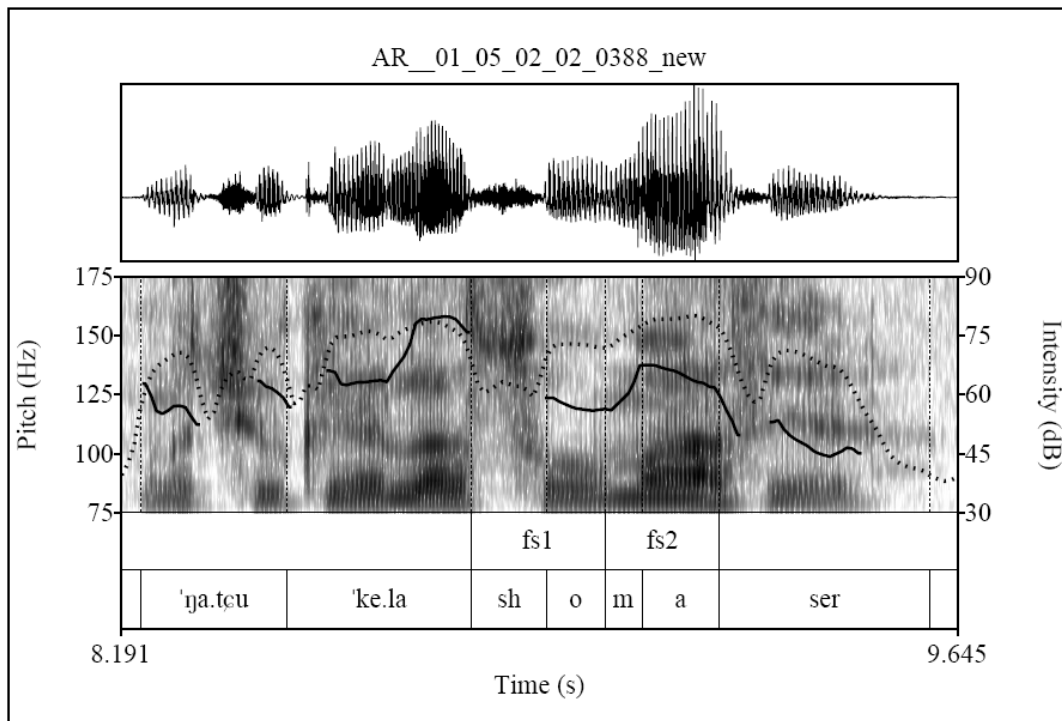
Figure 5.67 AR_05 / Adjectives / Isolation / Pitch trace: [s^ho.'ma] *so.ma* 'new'



This same adjective is fairly representative of the frame forms as well; a pitch trace is shown in Figure 5.68 below. In this setting, pitch slopes are similar on the two syllables: -4.9 Hz/100msec in σ_1 and -8.0 Hz/100msec in σ_2 , for a slope difference of

only -3.0 Hz/100msec. Conversely the difference in the average pitch in each syllable is more significant: 120 Hz in σ_1 , 134 Hz in σ_2 , for a respectable difference of 14 Hz. Here, intensity and vowel duration also contribute to the perception of σ_2 stress: Δ intensity = 7 dB; Δ vowel duration = 35 msec. (The contrast in vowel quality / vowel height across syllables emphasizes the contrast for both of these parameters.)

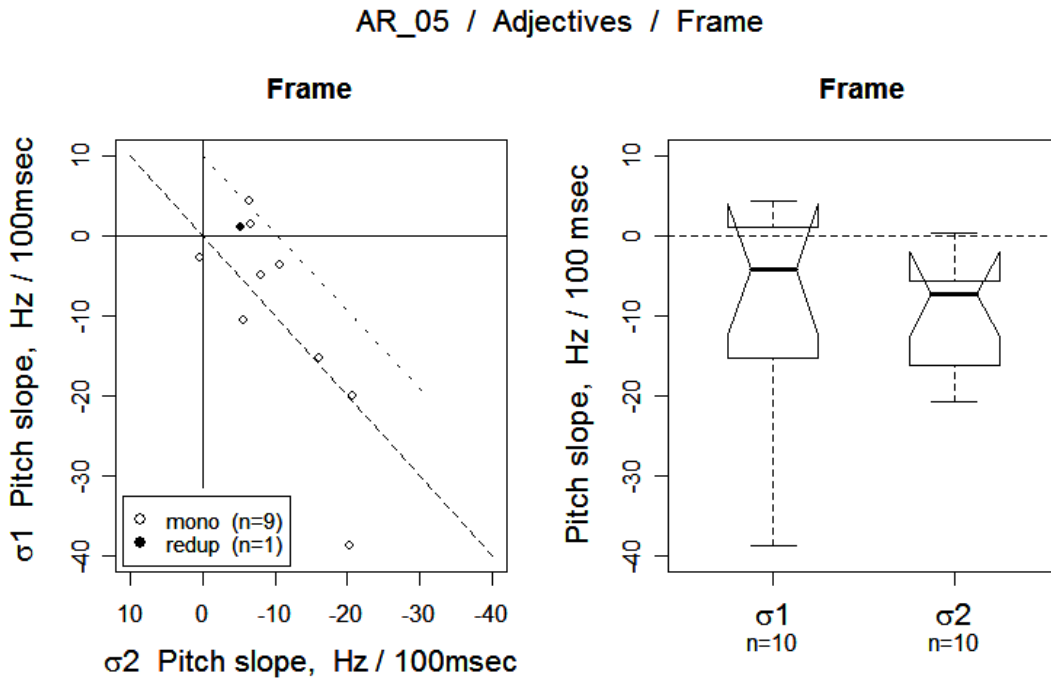
Figure 5.68 AR_05 / Adjectives / Frame / Pitch trace: [s^ho.'ma] *so.ma* 'new'



Pitch slope measurements for all of the frame forms of AR_05 adjectives are plotted in Figure 5.69 below. Points fall close to and to either side of the dashed line. This indicates that the contrast in slope across syllables is very small; pitch slopes “more downward” sometimes in σ_1 and sometimes in σ_2 . (The anomalous form at

coordinates (-20, -39) is again [$\text{ç}^{\text{h}}\text{i}.\text{ko}$] *phyug.po* ‘rich’.) The box-and-whisker plot to the right shows considerable overlap of both the notches and the boxes.

Figure 5.69 AR_05 / Adjectives / Frame: Pitch slope contrasts



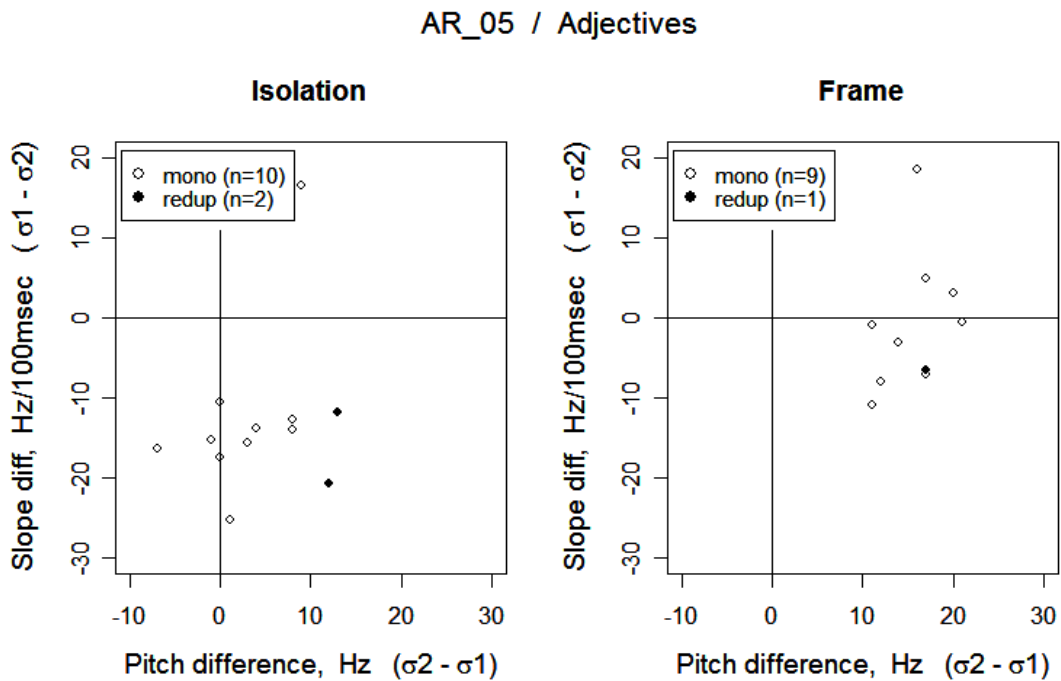
A paired-sample t-test, summarized in Table 5.16 below, confirms that the difference in pitch slope across syllables is not significant, with $p > 0.05$ and a 95% confidence limit which includes zero.

Table 5.16 AR_05 / Adjectives / Frame / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Subset	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_05	frame	-0.3933	9	0.7032	-1.0	-7.1	5.0

The complementary nature of the relationship between pitch and pitch slope in AR_05 adjectives is illustrated in Figure 5.70 below. In the isolation forms, where pitch is not a strong correlate of σ_2 stress, pitch slope is. Conversely, in the frame forms, where pitch slope is not a strong correlate of σ_2 stress, pitch is. Pitch and pitch slope can thus be regarded as two reflexes of fundamental frequency, which is manipulated differently in the two settings in order to convey stress. Of course, no adjectives fall in the upper left quadrant in the graphs below. Those would be cases in which neither pitch nor pitch slope was a reliable correlate of the perceived σ_2 stress.

Figure 5.70 AR_05 / Adjectives: Pitch difference vs. pitch slope difference



In summary, while pitch slope is a consistent correlate of stress for the isolation forms of AR_05 adjectives, it is not a correlate of stress for the frame forms.

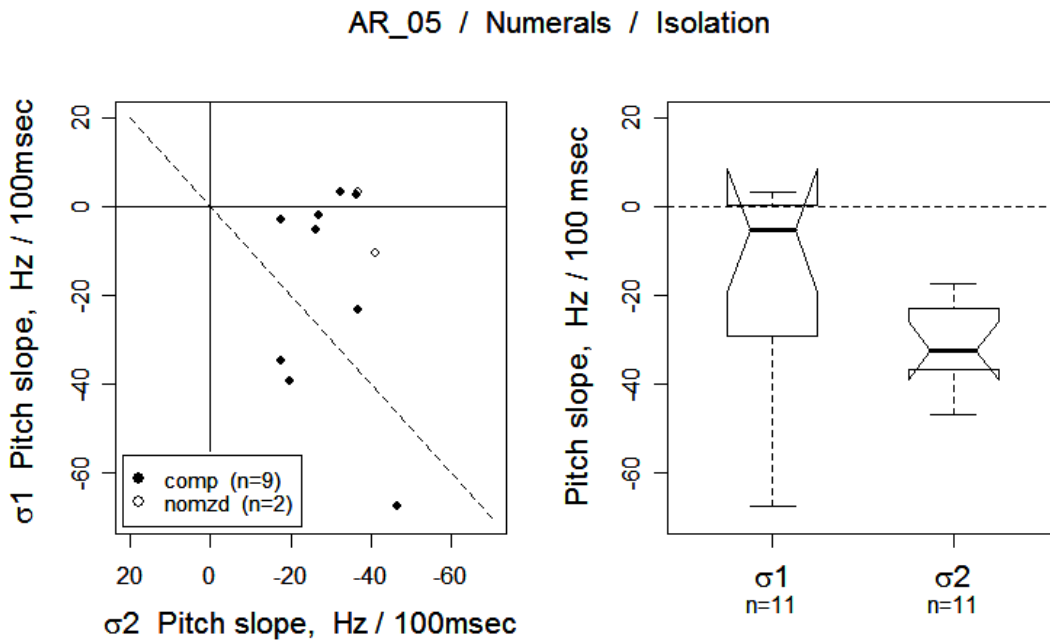
But pitch and pitch slope behave as complementary reflexes of fundamental frequency; in one manifestation or the other, fundamental frequency is manipulated in order to convey stress.

5.3.2.3 AR_05 Numerals

In numerals produced by speaker AR_05, pitch slope does not show a consistent correlation with σ_2 stress.

In Figure 5.71 below, eight of the eleven numerals plot to the right of the dashed line, meaning that pitch slope is “more downward” in σ_2 than in σ_1 . (These include the two nominalized forms: [ˈdʌn.'ba] *bdun.pa* ‘seventh’ – which is partially obscured at (-37, 3.3) – and [ksum.'ba] *gsum.pa* ‘third’.) The other three numerals fall to the left of the dashed line, with slope “more downward” in σ_1 .

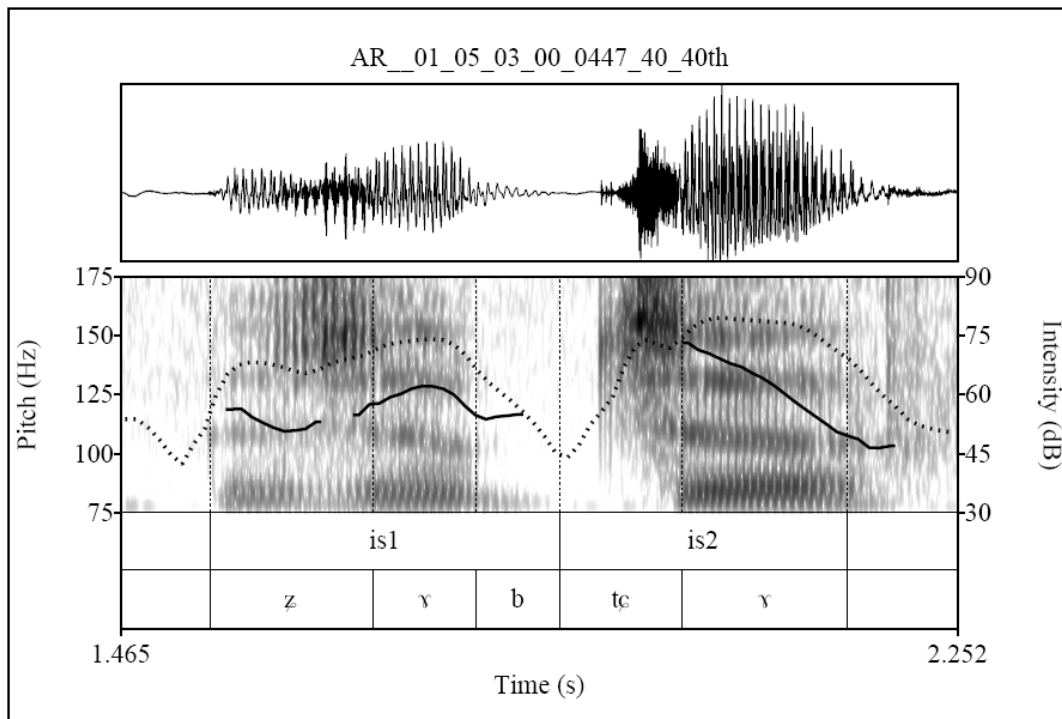
Figure 5.71 AR_05 / Numerals: Pitch slope contrasts



One numeral which falls to the right of the dashed line is [ʒʁb.'tʁʁ] *bzhi.bcu*

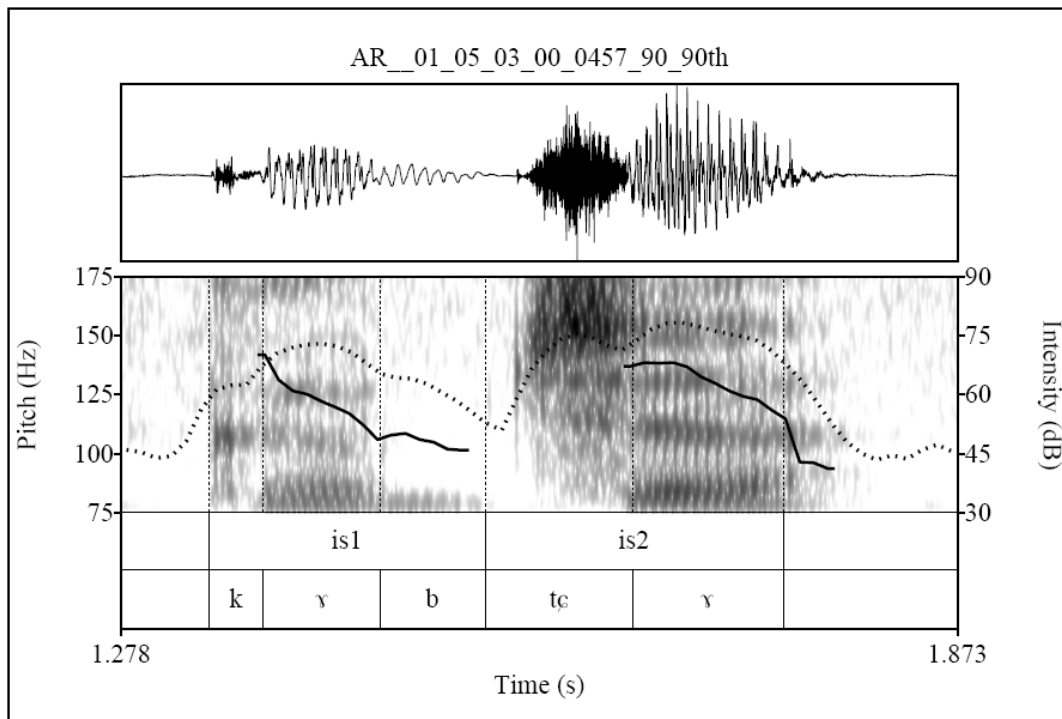
‘forty’; the pitch trace for this word is shown in Figure 5.72 below. Here, pitch is irregular in σ_1 (with a downward slope at the end measured as -5.3 Hz/100msec), and slopes downward steadily and steeply in σ_2 , at -26 Hz/100msec. The slope difference is thus -21 Hz/100msec, which is quite noticeable. Intensity and vowel duration are also both more prominent on σ_2 : Δ intensity = 5 dB, and Δ vowel duration = 58 msec. The difference in pitch across syllables is only 3 Hz.

Figure 5.72 AR_05 / Numerals / Isolation / Pitch trace: [zɤb.'tɕɤ] bzhi.bcu ‘forty’



On the other hand, a numeral that falls to the left of the dashed line in the graph above is [kɤb.'tɕɤ] *dgu.bcu* ‘ninety’, shown in Figure 5.73 below. Pitch slopes downward in σ_1 at -39 Hz/100msec, and downward in σ_2 at -20 Hz/100msec, for a difference of 18 Hz/100msec – i.e., “more downward” in σ_1 than in σ_2 . However, while pitch slope is more prominent in σ_1 , all the other acoustic parameters are more prominent in σ_2 : Δ pitch = 10 Hz; Δ intensity = 5 dB; Δ vowel duration = 24 msec.

Figure 5.73 AR_05 / Numerals / Isolation / Pitch trace: [kʁb.'tɕʁ] *dgu.bcu* ‘ninety’



As expected, a paired-sample t-test shows that the contrast across syllables is *not* statistically significant. As summarized in Table 5.17 below, $p > 0.05$ and the 95% confidence interval spans zero.

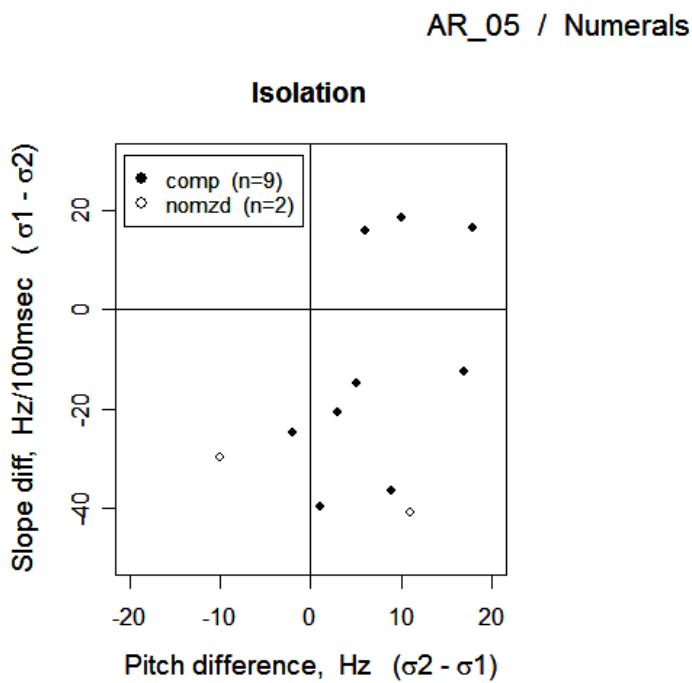
Table 5.17 AR_05 / Numerals / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_05	isolation	-2.0792	10	0.06429	-15	-31	1.1

Finally, the relationship between pitch and pitch slope is illustrated in Figure 5.74 below. Some numerals fall in the lower right quadrant, indicating that both pitch and pitch slope are more prominent in σ_2 , the stressed syllable. For the three

numerals that plot above the x-axis, pitch is a correlate of stress though pitch slope is not; for the two numerals that plot to the left of the y-axis, pitch slope is a correlate of stress though pitch is not. That is, in every case, one or both reflexes of fundamental frequency serve(s) to convey stress.

Figure 5.74 AR_05 / Numerals: Pitch difference vs. pitch slope difference



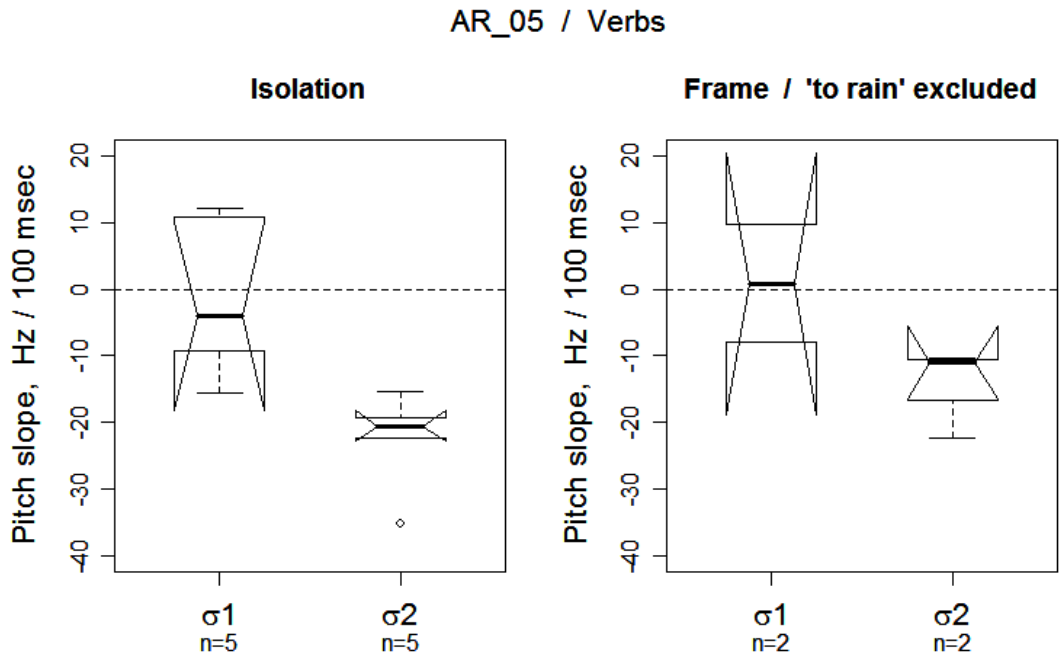
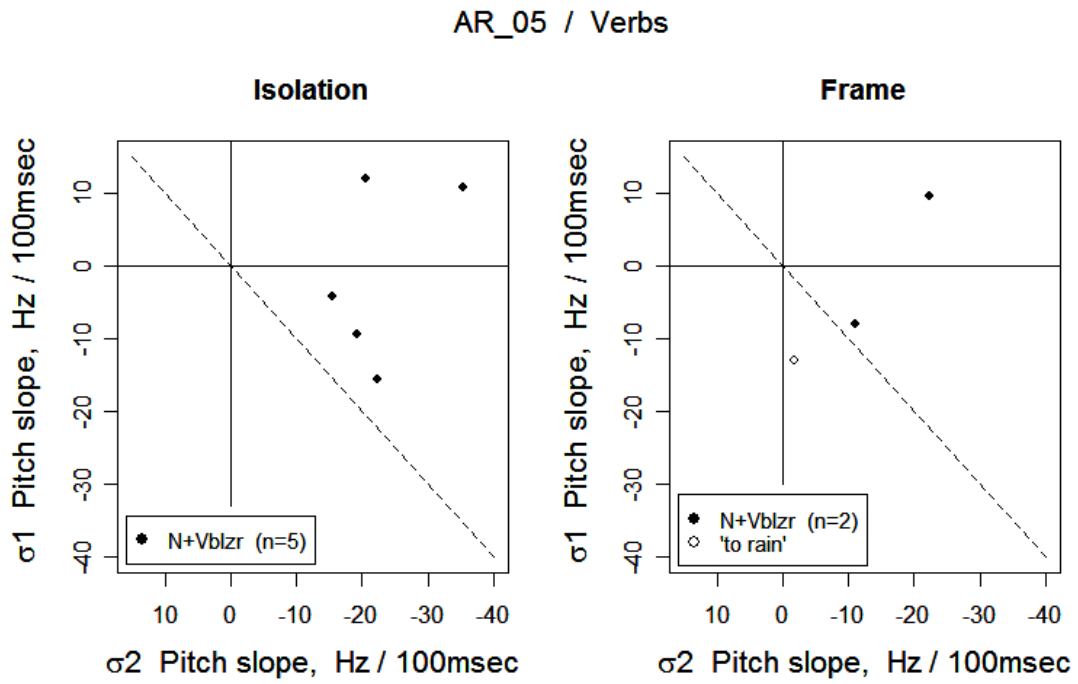
In summary, pitch slope alone is not a consistent cue for σ_2 stress in AR_05 numerals. However, in individual cases, pitch is a cue for stress when pitch slope is not, and pitch slope is a cue for stress when pitch is not. Either way, fundamental frequency is a reliable correlate of stress.

5.3.2.4 AR_05 Verbs

As discussed in section 5.2.2.4, for AR_05 verbs produced in isolation and in the sentence frame, the contrast in pitch across syllables serves as a robust cue for $\sigma 1$ stress, despite the small sample size. Pitch slope patterns are a direct outcome of this contrast in pitch across syllables, so slope is not contextually significant. Since pitch slope is thus epiphenomenal, it does not function as a correlate of stress in verbs. This was also the case for verbs produced by speakers BSh_03, BM_01, and AR_04.

Pitch slope measurements for both isolation and frame forms of AR_05 verbs are plotted in Figure 5.75 below. In isolation, all points plot to the right of the dashed line, meaning that whether pitch in $\sigma 1$ slopes upward or downward, pitch in $\sigma 2$ slopes “more downward”. Of the five frame forms recorded and analyzed, pitch slope measurements could be collected from both syllables of only three of them; one of them is the verb ‘to rain’, which, as discussed previously, was produced with an anomalous stress pattern. For the two remaining frame forms – [‘nda.hɛn] *mda*.’ *phen* ‘to shoot an arrow’ and [‘k^hʌŋ.bʌp] *gangs.babs?* ‘to snow’ – both points also plot to the right of the dashed line. The box-and-whisker plots at the bottom of the figure reflect the fact that, in both settings, pitch slopes vary more widely across different verbs in $\sigma 1$ than in $\sigma 2$.

Figure 5.75 AR_05 / Verbs: Pitch slope contrasts



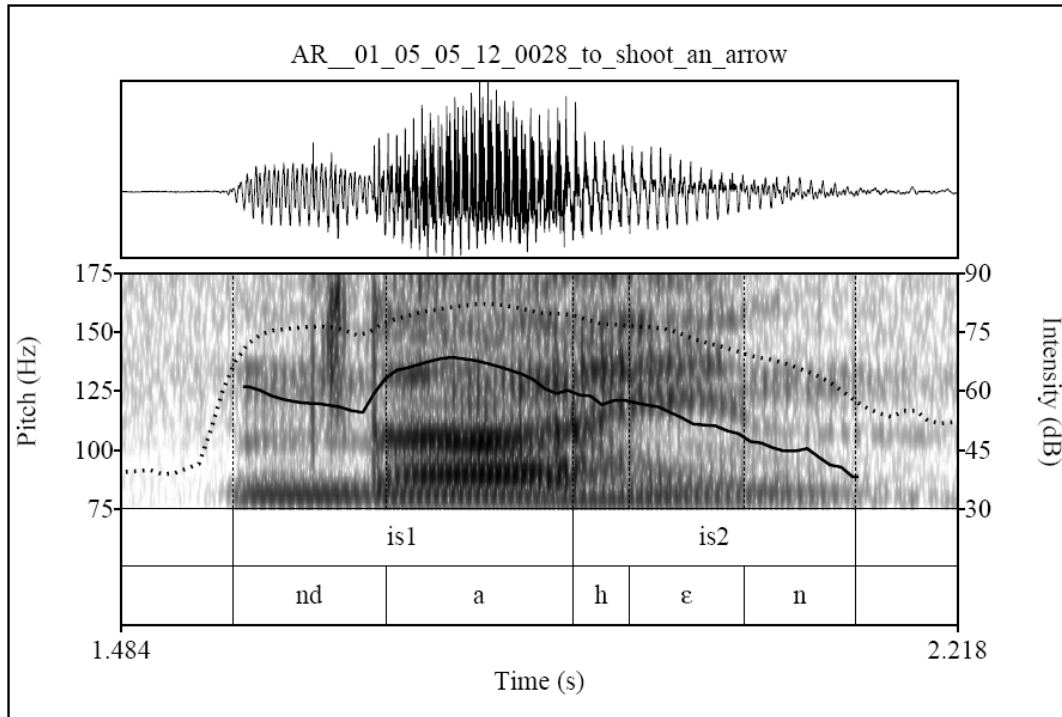
Paired-sample t-tests, summarized in Table 5.18 below, indicate that the contrast in slope across syllables is not statistically significant for verbs in either setting – though only barely so for the isolation forms.

Table 5.18 AR_05 / Verbs / Pitch slope: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (Hz/100msec)	95 % conf. limits	
						lower	upper
AR_05	isolation	-2.7758	4	0.05003	-21	-43	0.005
	frame	-1.2106	1	0.4395	-18	-202	166

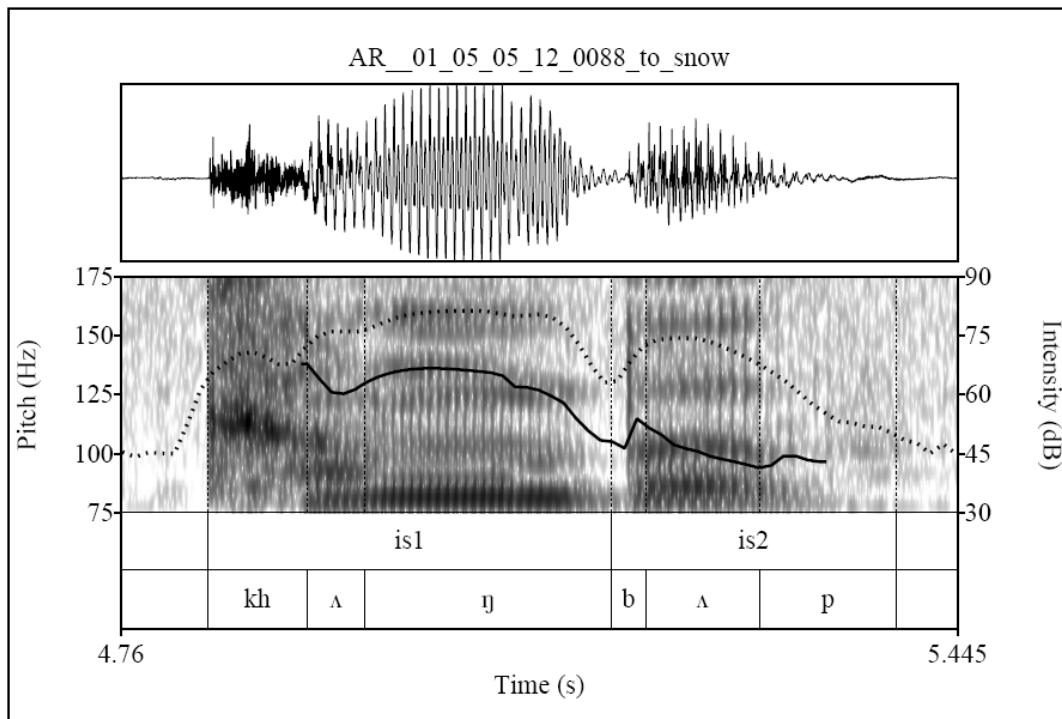
The pitch trace for the verb [ˈnda.hən] *mda* ‘*phen* ‘to shoot an arrow’ produced in isolation is shown in Figure 5.76 below. As demonstrated previously in sections 5.2.1.3 and 5.2.2.4, pitch is a robust correlate of stress in Rebkong Amdo verbs. In this example, pitch is typically higher in σ_1 (137 Hz) than in σ_2 (113 Hz), for a difference of -24 Hz. The decline in pitch from σ_1 to σ_2 defines a steady trajectory, and it is this trajectory that yields the pitch slopes observed in each syllable: -4.1 Hz/100msec on σ_1 , and -15 Hz/100msec on σ_2 , for a slope difference across syllables of -11 Hz/100msec. Since the pitch slope contrast is not contextually significant – it cannot be distinguished from the pitch contrast – it cannot be a correlate of stress. (In this word, both intensity and vowel duration are also more prominent in σ_1 , the stressed syllable, though this may in part be attributable to the contrast in vowel quality across syllables; Δ intensity = -8 dB, Δ vowel duration = -64 msec.)

Figure 5.76 AR_05 / Verbs / Isolation / Pitch trace: ['nda.hɛn] *mda'*.*'phen* 'to shoot an arrow'



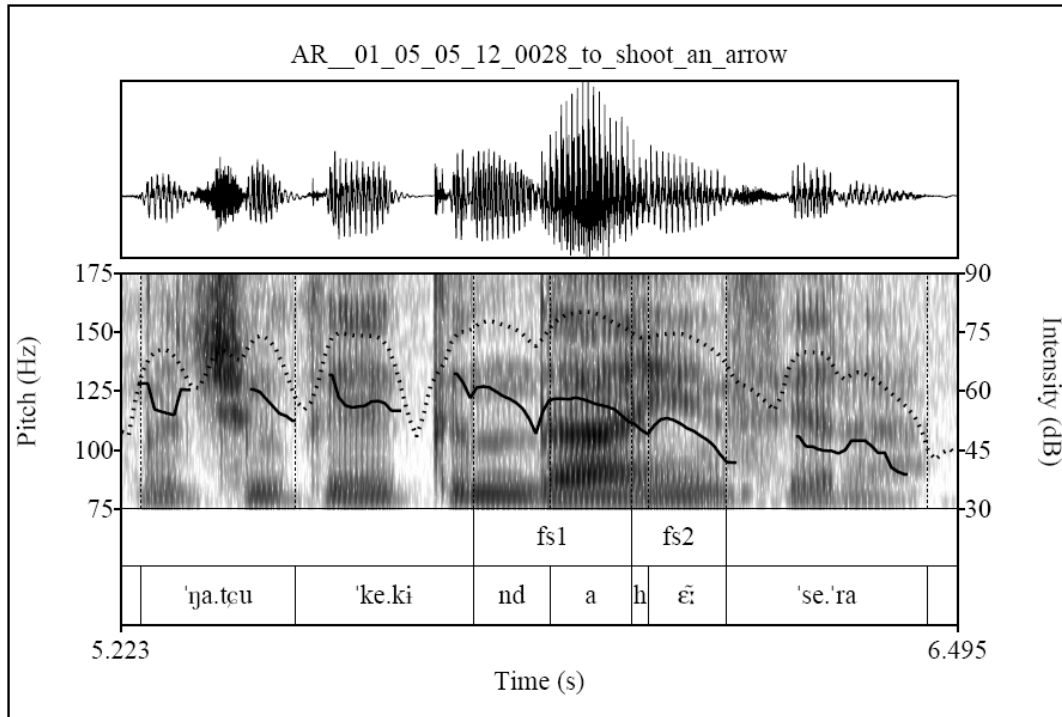
A similar pattern is observed in the verb ['k^hʌŋ.bʌp] *gangs.babs?* 'to snow', illustrated in Figure 5.77 below. Here, the difference in pitch across syllables is -27 Hz. The decline in pitch from σ_1 to σ_2 is reflected in the pitch slopes observed: -9.3 Hz/100msec in σ_1 , and -19 Hz/100msec in σ_2 , for a slope difference of -9.8 Hz/100msec. (Δ intensity = -2 dB, Δ vowel duration = 46 msec.)

Figure 5.77 AR_05 / Verbs / Isolation / Pitch trace: ['k^hʌŋ.bʌp] *gangs.babs?* 'to snow'



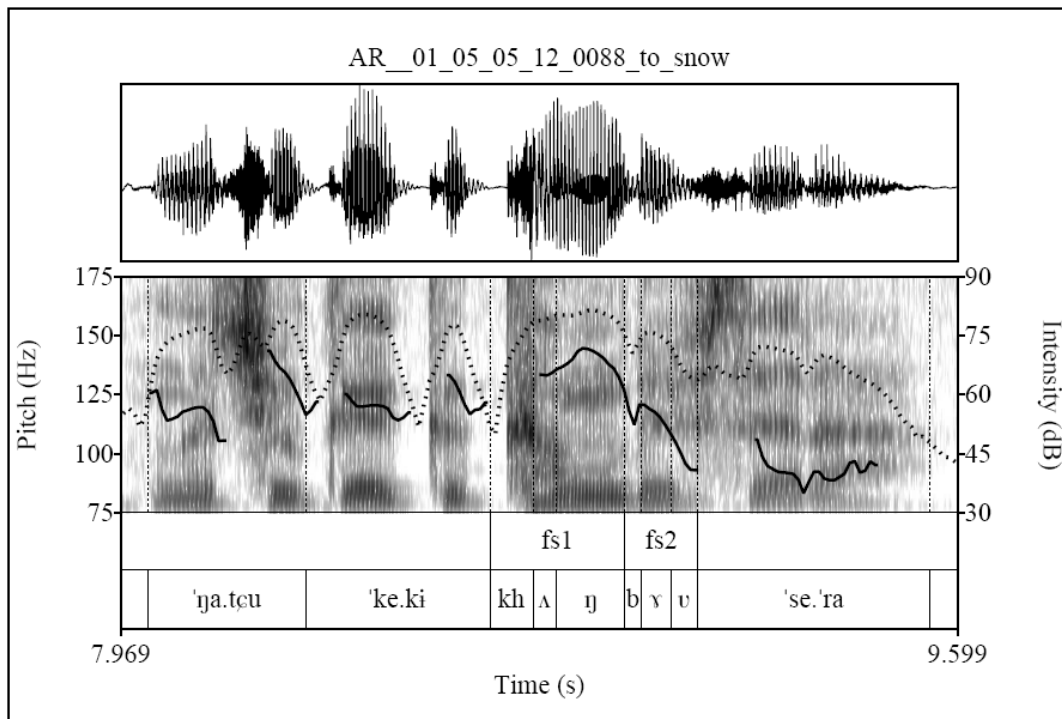
The frame forms for these two words exhibit similar behavior. In ['nda.hẽ:] *nda*. 'phen 'to shoot an arrow' – shown in Figure 5.78 below – pitch declines from 120 Hz in σ_1 to 110 Hz in σ_2 . This decline in pitch across syllables is reflected in the rates of pitch slope: -7.9 Hz/100msec on σ_1 , and -11 Hz/100msec on σ_2 , for a difference of -3.0 Hz/100msec (i.e., slightly more downward in σ_2). Intensity also contributes to the perception of σ_1 stress (Δ intensity = -6 dB); the difference in vowel duration is a trivial -5 msec.

Figure 5.78 AR_05 / Verbs / Frame / Pitch trace: ['nda.hɛn] *mda*.'phen 'to shoot an arrow'



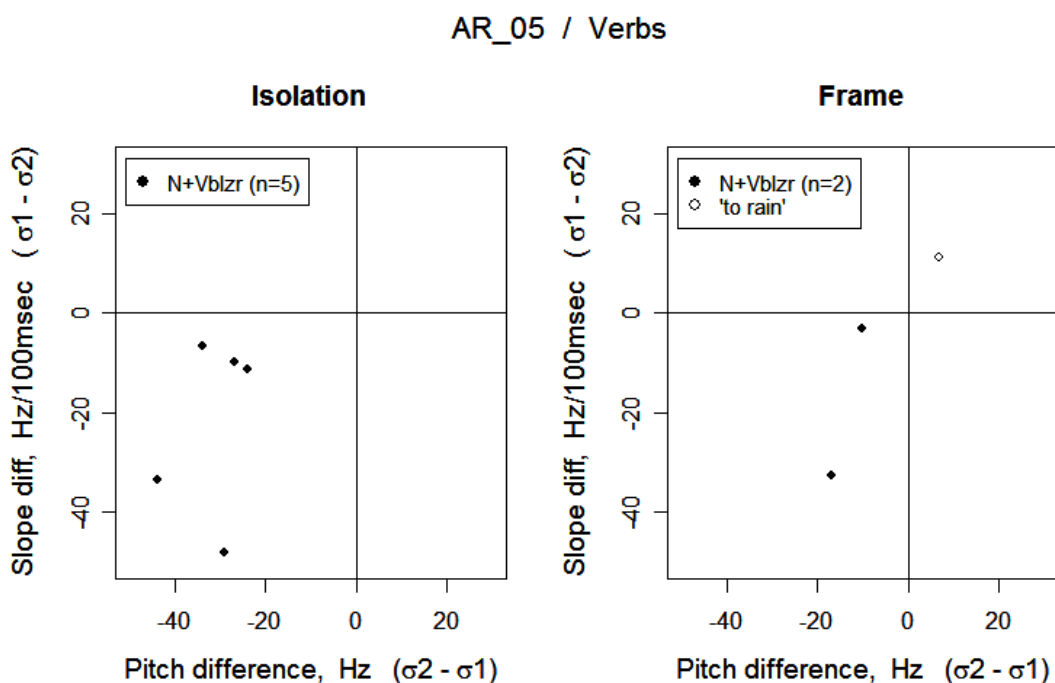
The pitch trace for the frame form of ['k^hʌŋ.bʌp] *gangs.babs?* 'to snow' is shown in Figure 5.79 below. Again, the downward pitch slope in both syllables follows from the decrease in pitch across syllables, which serves as a cue for σ_1 stress: $\Delta\text{pitch} = -17$ Hz; $\Delta\text{pitch slope} = (-22.31 \text{ in } \sigma_2) - (9.7 \text{ in } \sigma_1) = -32$ Hz/100msec. Intensity is also a cue for σ_1 stress ($\Delta\text{intensity} = -4$ dB), while the vowel in σ_2 is actually longer than that in σ_1 ($\Delta\text{vowel duration} = 12$ msec.)

Figure 5.79 AR_05 / Verbs / Frame / Pitch trace: ['k^hʌŋ.bʌp] *gangs.babs?* 'to snow'



The relationship between pitch and pitch slope is illustrated in Figure 5.80 below. Verbs produced in isolation and in the sentence frame (with the exception of the anomalous 'to rain') fall in the lower left quadrant of the graph. The fact that all points plot to the left of the y-axis means that pitch is higher in σ_1 than in σ_2 – i.e., pitch declines from σ_1 to σ_2 . The fact that all points also plot below the x-axis – i.e., with a more downward slope in σ_2 than in σ_1 – is a direct consequence of this pitch pattern: the slope in σ_2 is simply a continuation of the downward trajectory of pitch. Pitch slope is thus an epiphenomenon of pitch; lacking contextual significance, it is not a correlate of the perceived σ_1 stress.

Figure 5.80 AR_05 / Verbs: Pitch difference vs. pitch slope difference



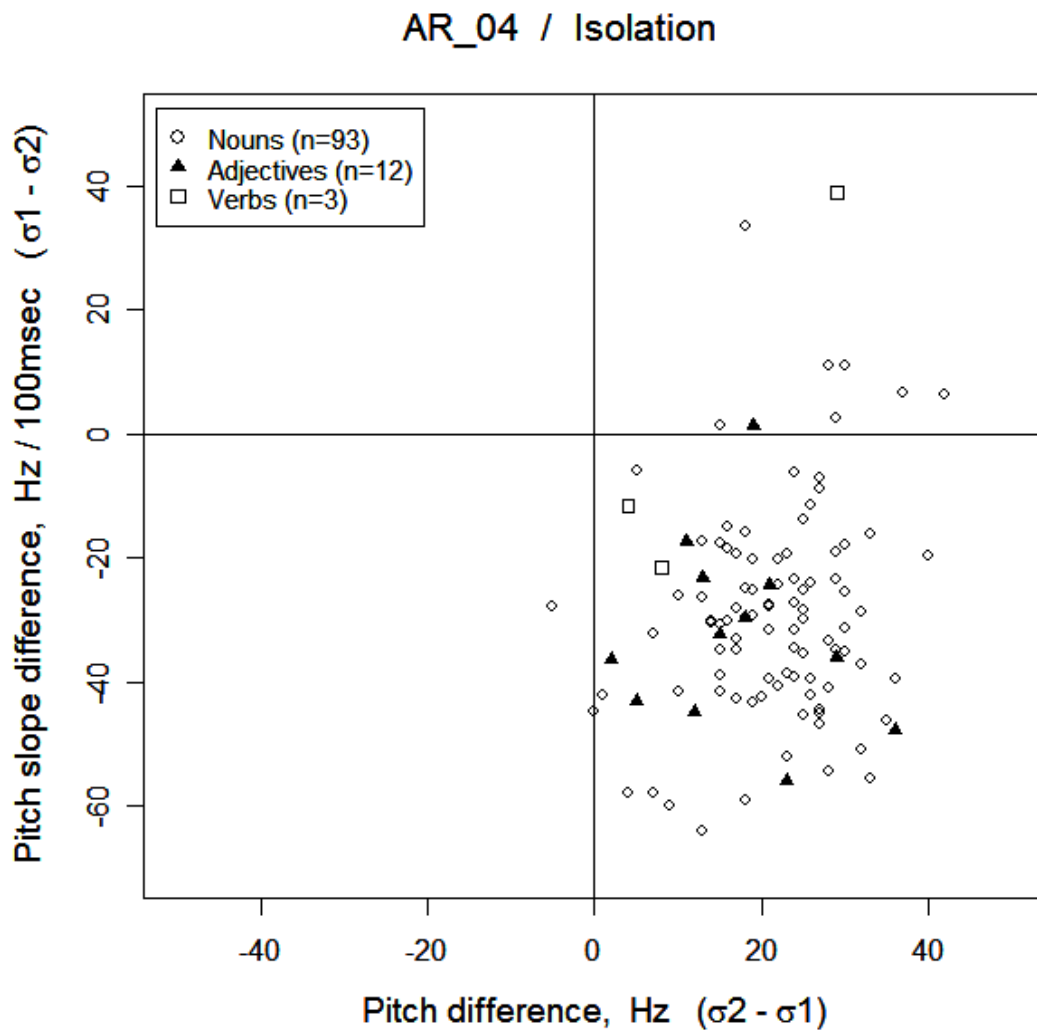
5.3.3 The interaction of pitch and pitch slope

As demonstrated in the sections above, pitch and pitch slope can be regarded as two expressions of a single acoustic resource: fundamental frequency. In Rebkong Amdo, a speaker can use pitch, pitch slope, or both to convey stress. In this section, the relationship between pitch and pitch slope is compared across lexical categories.

In Figure 5.81 below, pitch differences and pitch slope differences across syllables are plotted for isolation forms produced by speaker AR_04, for all lexical categories. As noted previously, this speaker was not comfortable producing verbs in isolation, so these points cannot be considered representative. For nouns and adjectives, most points fall in the lower right quadrant, indicating that pitch and pitch slope reinforce one another as correlates of stress. (Since pitch is a robust correlate of

σ_2 stress regardless of the potential effects of vowel height – as demonstrated in sections 4.2.1.1 and 5.2.1.2 – there was no reason to control for this parameter, so all points are included in the graph below.)

Figure 5.81 AR_04 / Isolation: Pitch difference vs. pitch slope difference



Virtually all of the nouns and adjectives in Figure 5.81 above have a pitch contrast of at least 10 Hz, or a pitch slope contrast of at least -10 Hz/100msec, or both

– that is, they plot away from the graph’s origin. The point closest to the origin is the noun [çu.'ma] *?.ma* ‘tea with tsampa’, at coordinates (5, -5.8).

In some cases, one of these correlates is weak, while the other is strong. These are points falling close to (or above) the x-axis, or close to (or left of) the y-axis.

Crucially, none of the nouns or adjectives fall in the upper-left quadrant of the graph. These would be cases with both pitch and pitch slope more prominent in σ_1 , contradicting the stress pattern.

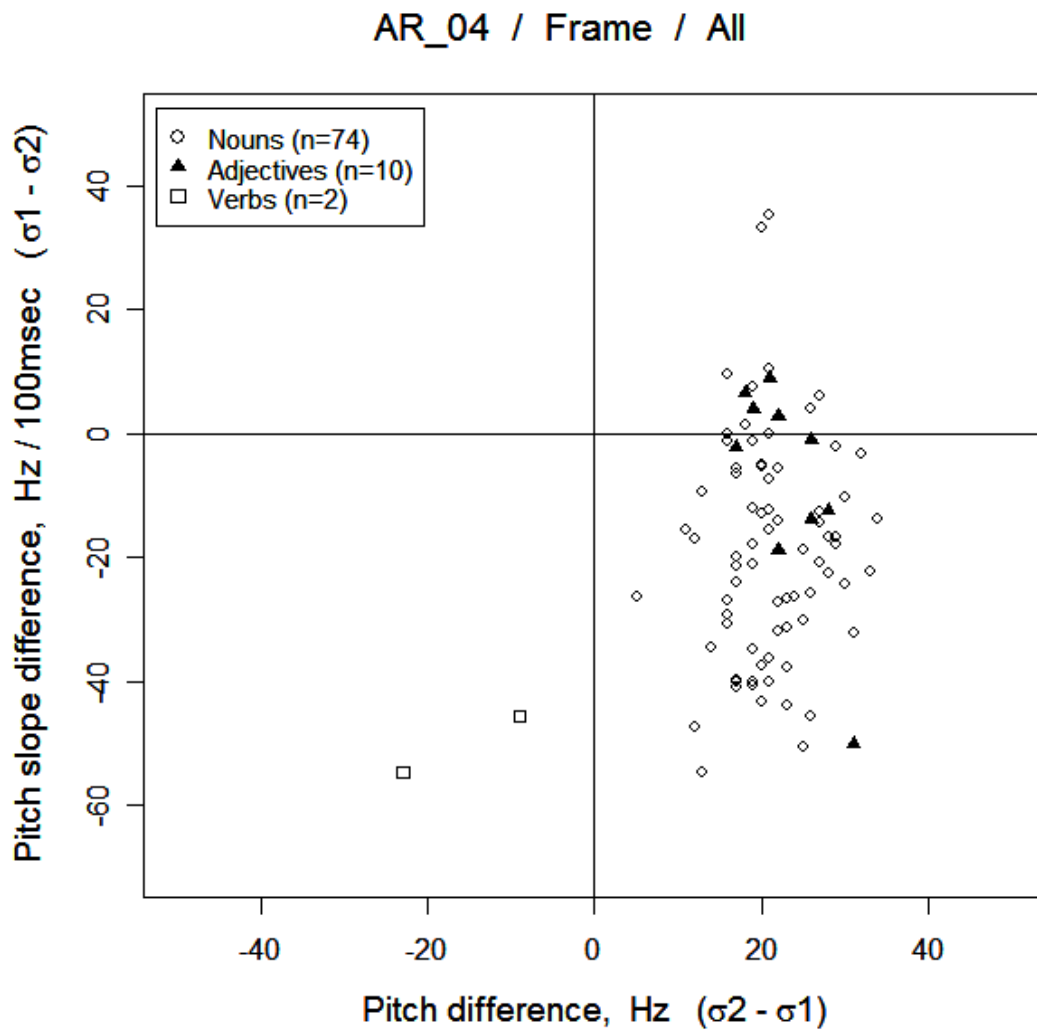
For the AR_04 frame forms, plotted in Figure 5.82 below, the two verbs are clearly distinguished from non-verbs by pitch, but not by pitch slope. (That is, verbs plot left of the y-axis, non-verbs plot right of the y-axis, but all plot below the x-axis.) Again, most nouns and adjectives plot in the lower-right quadrant, indicating that pitch and pitch slope are both more prominent on σ_2 , the stressed syllable. A number of nouns and adjectives have a more prominent pitch slope on σ_1 and thus fall above the x-axis. But in all of these cases, the average pitch is at least 17 Hz higher on σ_2 , which is certainly a distinctive cue for stress.

There are no points close to the origin of the graph: contrasts in pitch are always at least 10 Hz, and/or contrasts in pitch slope are always at least -10 Hz/100msec.

For verbs, the difference in pitch across syllables is negative, meaning that pitch is higher in σ_1 , which is the stressed syllable. Both tokens also plot below the x-axis, indicating a steeper downward slope in σ_2 . As discussed in section 5.3.1.3, this

is completely unrelated to stress, and is a direct result of the drop in pitch across the entire word.

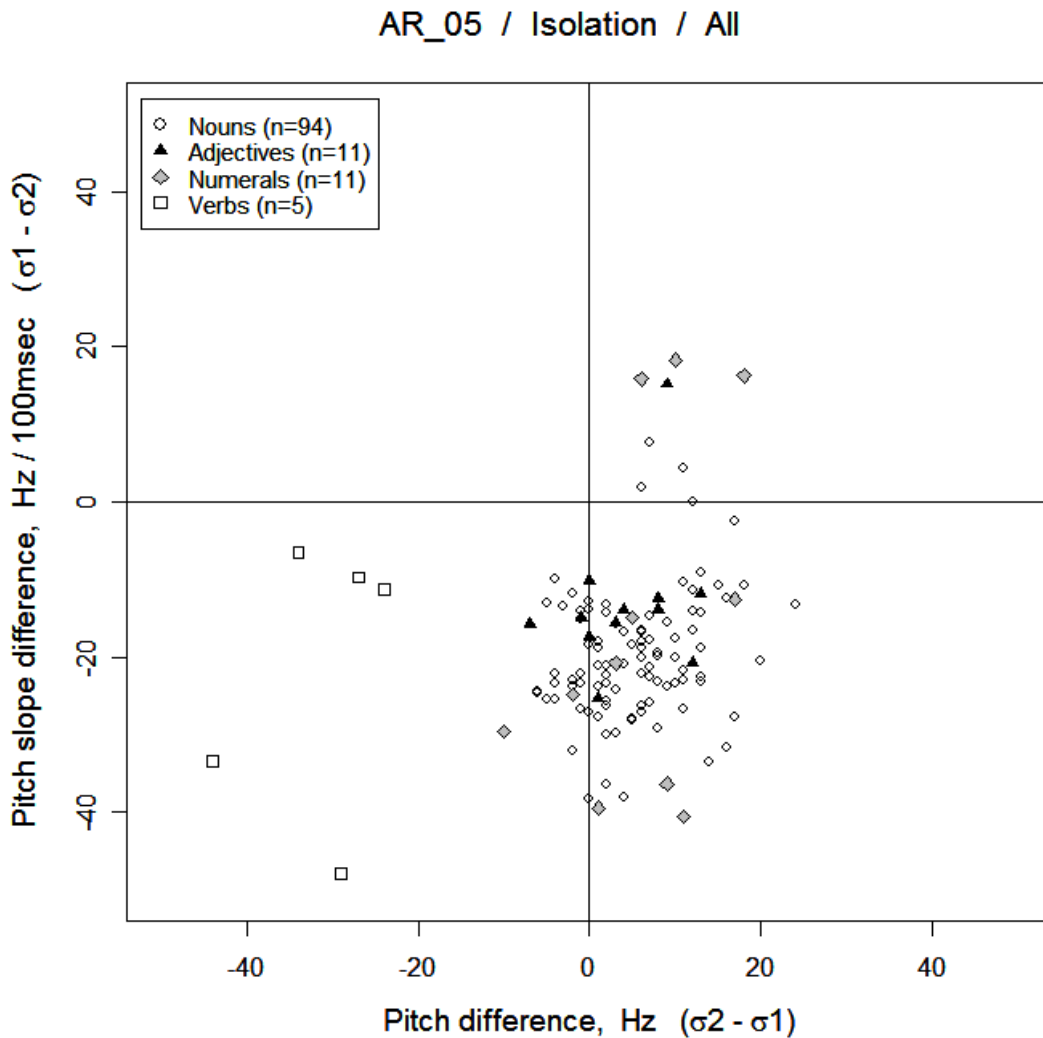
Figure 5.82 AR_04 / Frame: Pitch difference vs. pitch slope difference



Pitch and pitch slope contrasts for AR_05 isolation forms are plotted in Figure 5.83 below. Even before the incidental effects of contrasts in vowel height and syllable closure are controlled for, it is clear that the verbs are again distinguished

from the non-verbs by pitch, though not by pitch slope; again, nouns, adjectives, and numerals generally pattern alike.

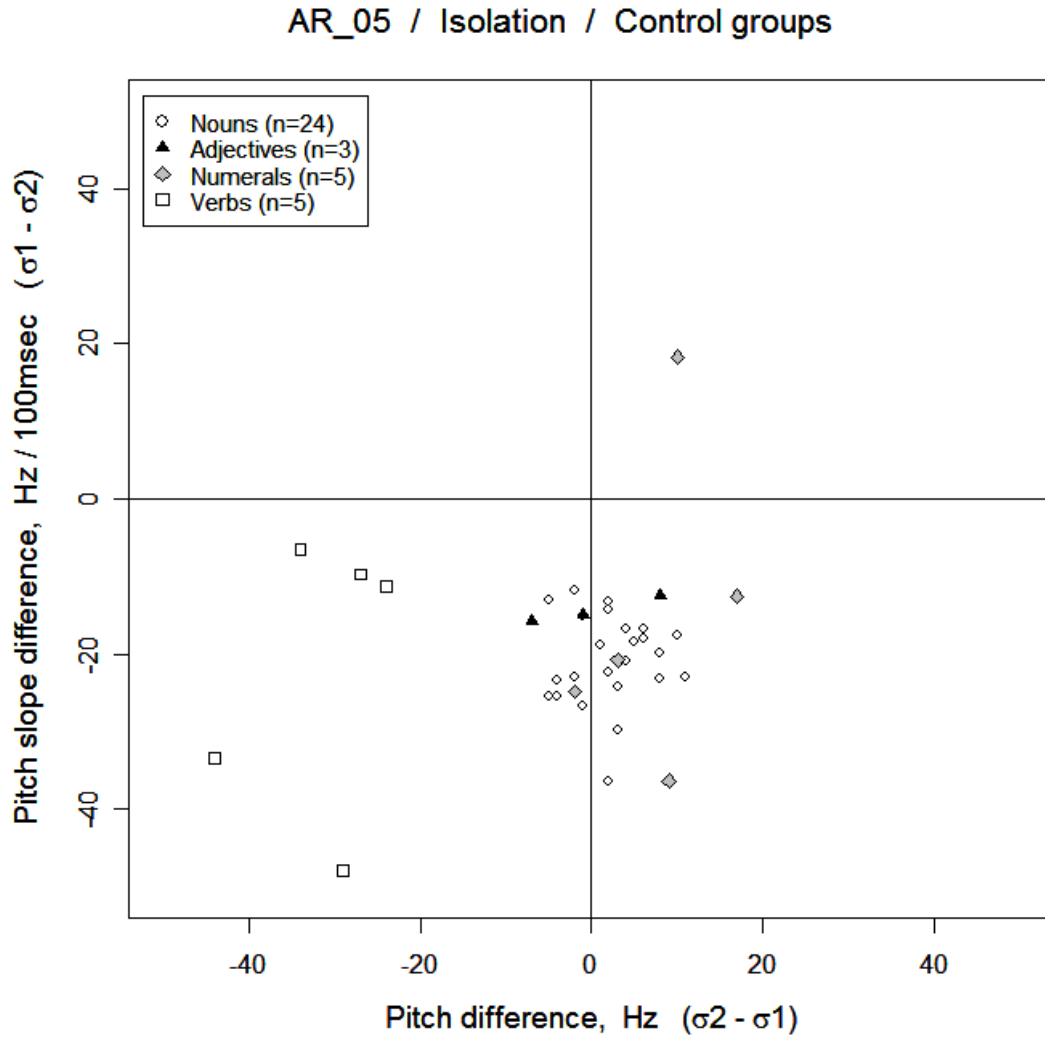
Figure 5.83 AR_05 / Isolation: Pitch difference vs. pitch slope difference



The graph in Figure 5.84 below shows only the appropriate control groups. For non-verbs, the control groups include only those forms in which σ_2 is open, and which also have vowels of the same height in both syllables. For verbs, all tokens are

plotted, since vowel height and σ_2 closure did not have an incidental effect on pitch or pitch slope.

Figure 5.84 AR_05 / Isolation / Control groups: Pitch difference vs. pitch slope difference

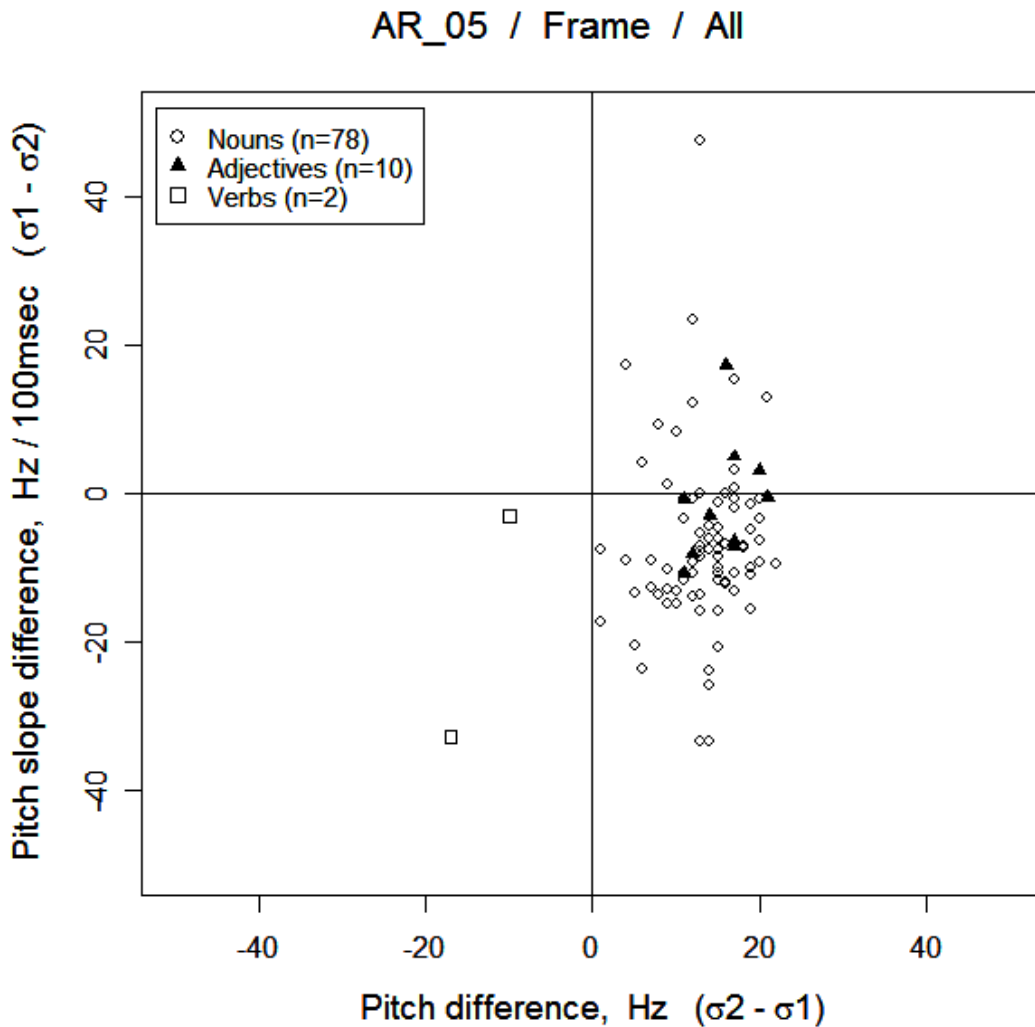


As shown, for the control group of nouns, adjectives, and numerals, pitch is not a consistent cue for stress. These non-verbs plot close to or to either side of the y-axis: sometimes pitch is higher in σ_1 , sometimes it is higher in σ_2 , and the contrast

across syllables is not very dramatic. In all of these cases, pitch slope serves as the reliable correlate of stress, with a contrast in slope of at least -10 Hz/100msec. The only word which plots above the x-axis is the numeral [kʏb.'tɕʏ] *dgu.bcu* 'ninety', discussed in section 5.3.2.3. Crucially, none of the non-verbs plot in the upper left quadrant. These would be cases with both pitch and pitch slope more prominent on σ_1 , contradicting the stress pattern.

Pitch differences and pitch slope differences for the AR_05 frame forms are plotted in Figure 5.85 below. (The anomalous verb 'to rain' has been excluded from this graph.) Once again, verbs are distinguished from non-verbs by pitch, though not by pitch slope.

Figure 5.85 AR_05 / Frame: Pitch difference vs. pitch slope difference

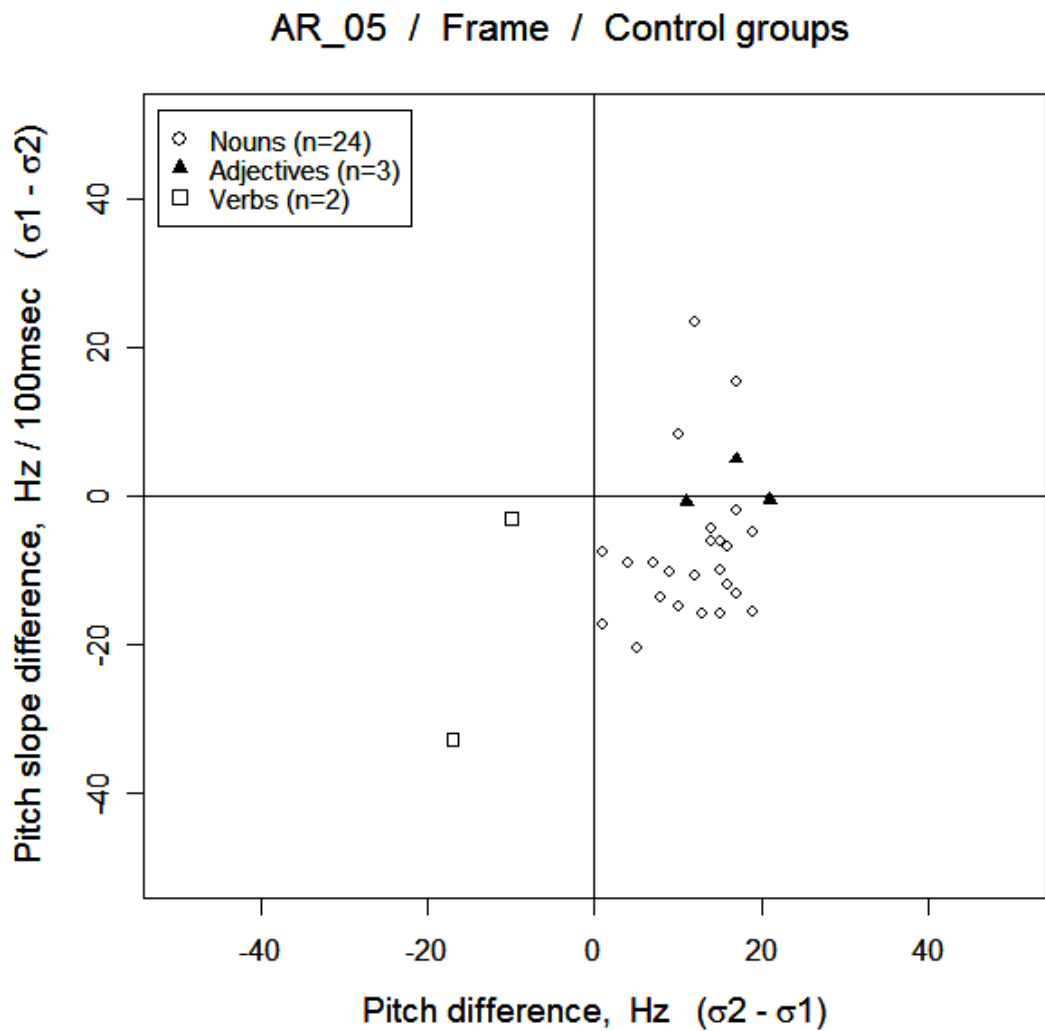


Patterns are more refined in Figure 5.86 below, which shows the control groups for nouns and adjectives – i.e., only those words which have vowels of the same height in both syllables, and in which σ_2 is open. (The full set of two verbs is plotted, since verbs were not affected by vowel height or syllable closure.)

For nouns – with only a few exceptions – both pitch and pitch slope are correlates of stress. For adjectives, pitch is a cue for stress, but pitch slope is not:

points fall close to and to either side of the x-axis. For verbs, pitch is a cue for stress, but pitch slope is not; slope differences are not contextually significant, since they are an epiphenomenon of the pitch contrast.

Figure 5.86 AR_05 / Frame / Control groups: Pitch difference vs. pitch slope difference



In conclusion, fundamental frequency can be regarded as a robust and reliable acoustic correlate of stress in Rebkong Amdo. For non-verbs, a speaker may exercise

flexibility in the manipulation of this acoustic resource, conveying stress sometimes through prominence in pitch, sometimes through prominence in pitch slope, and most often through prominence in both parameters. For verbs, stress is conveyed by pitch; pitch slope patterns are a direct reflection of pitch patterns, so pitch slope never serves as a correlate of stress.

5.4 Intensity

For the two speakers of Rebkong Amdo Tibetan considered here, there are small samples or subsets of data in which intensity does show a correlation with stress. However, I do not generally consider intensity to be a *meaningful* correlate of stress, in the sense that a listener cannot always depend on this acoustic signal to provide a statistically and perceptually significant means of identifying the stressed syllable.

As discussed in section 2.5.2.2, Lehiste (1970) demonstrated that vowels of different heights have different intrinsic intensities – low vowels intrinsically have a higher intensity than high vowels – and she identified 1 dB as the just-noticeable difference. Both of these points are taken into consideration in the analysis of intensity here.

For speaker AR_04, when contrasts in vowel height are controlled for – by focusing on subsets of words which have vowels of the same height in both syllables – nouns show a limited and perceptually weak correlation with σ_2 stress, while adjectives are found to have a slightly higher intensity on σ_1 , contradicting the stress

pattern. The sample of verbs is too small to be conclusive; the patterns here suggest that a larger sample might reveal a correlation between intensity and σ_1 stress.

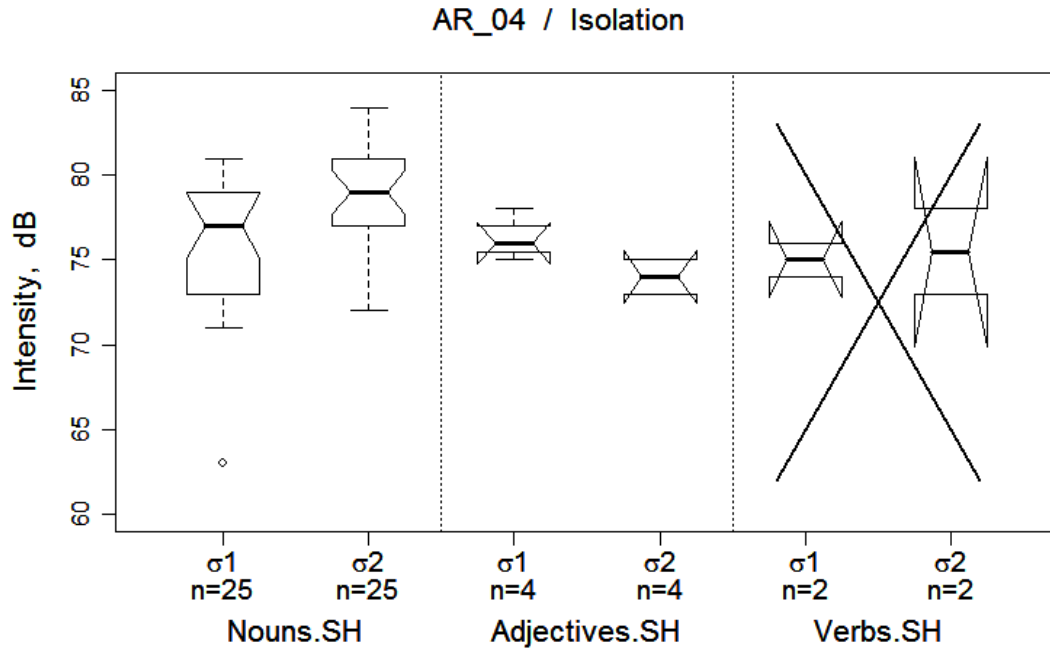
For speaker AR_05, after controlling for contrasts in vowel height, nouns show only a limited and weak correlation with σ_2 stress. This is also the case for adjectives recorded in the sentence frame; for adjectives produced in isolation, the contrast in intensity across syllables is not statistically significant, even for the control group. Numerals (recorded only in isolation) show a correlation between intensity and σ_2 stress, while verbs show a correlation between intensity and σ_1 stress.

Analysis of intensity data for speakers AR_04 and AR_05 is presented in sections 5.4.1 and 5.4.2, below.

5.4.1 Intensity for speaker AR_04

For words produced by speaker AR_04 in isolation, the distribution of intensity measurements in the two syllables of each lexical category is represented in the box-and-whisker plots in Figure 5.87 below. Since the intrinsic variation of intensity as a function of vowel height is such an influential factor in these analyses, these plots are based on the control groups – i.e., those words with vowels of the same height in both syllables. I have drawn an X through the verb plot, since I do not consider this data to be valid; as discussed previously, it proved very difficult for speaker AR_04 to produce verbs in complete isolation, without the normal context of tense, aspect, and evidentiality.

Figure 5.87 AR_04 / Isolation: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)



For the control group of nouns, intensity tends to be higher in σ_2 , the stressed syllable. However, while the intensity contrast across syllables is statistically significant, it may be so small as to barely exceed the difference limen of 1 dB discussed by Lehiste. Thus I consider AR_04 isolation nouns to exhibit a limited and weak correlation with stress – “limited” because the correlation does not pertain to all other vowel height groups, and “weak” because even a statistically significant difference may be barely perceptible.

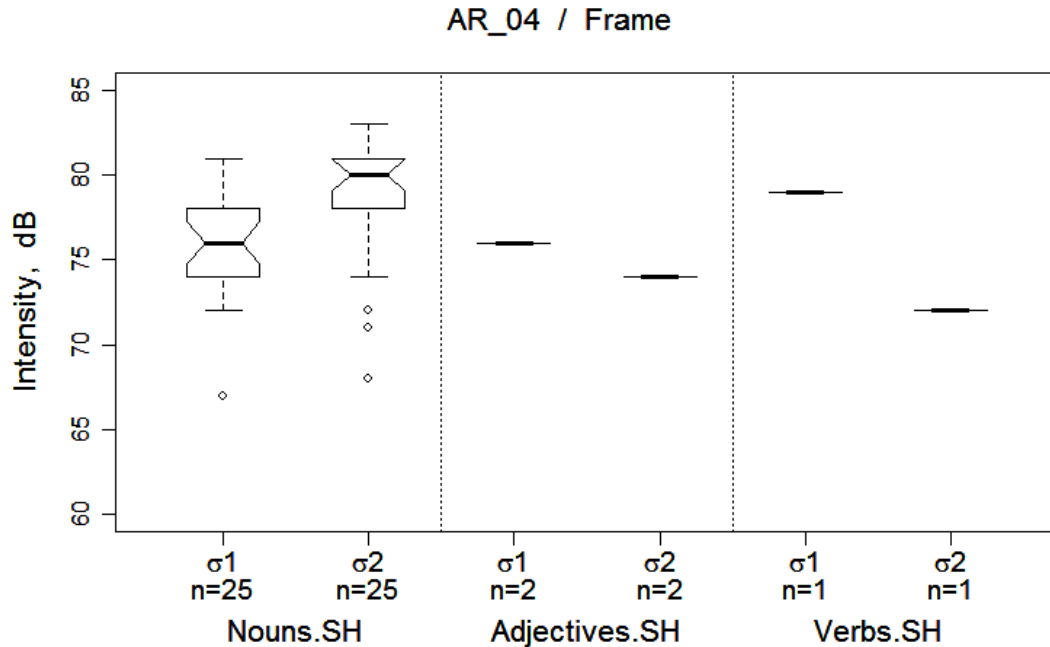
For adjectives in the control group, intensity is higher in σ_1 , the *unstressed* syllable, so it is quite clear that intensity is not a correlate of stress.

The distribution of intensity measurements for AR_04 frame forms are illustrated in Figure 5.88 below. Again, only the subsets on which the final analysis was based are represented.

Nouns and adjectives exhibit the same patterns as they did in isolation: Nouns show only a weak and limited correlation between intensity and σ_2 stress. The control group of adjectives has only two members, and in both of these intensity is higher in σ_1 , contradicting the stress pattern.

Verbs also have a higher intensity in σ_1 , which is the stressed syllable. In one of the two tokens which could be analyzed, this intensity pattern might be attributable to a [lower.higher] vowel height contrast. Thus there was only one control sample showing a correlation with stress.

Figure 5.88 AR_04 / Frame: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)

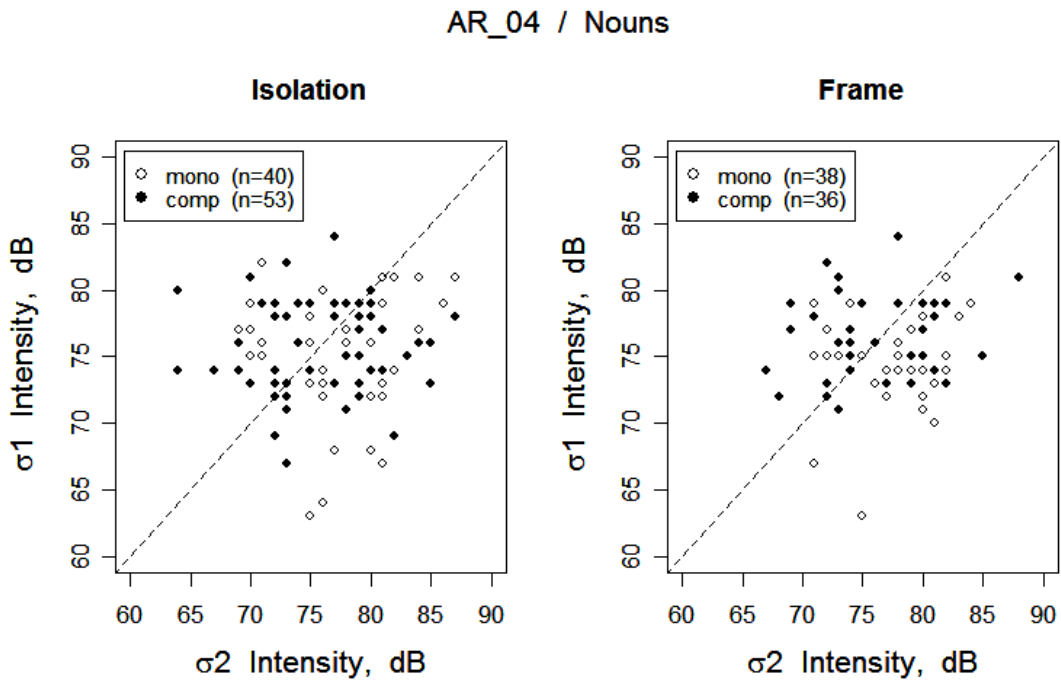


Detailed analyses of intensity data for AR_04 nouns, adjectives, and verbs are presented in sections 5.4.1.1 through 5.4.1.3 below.

5.4.1.1 AR_04 Nouns

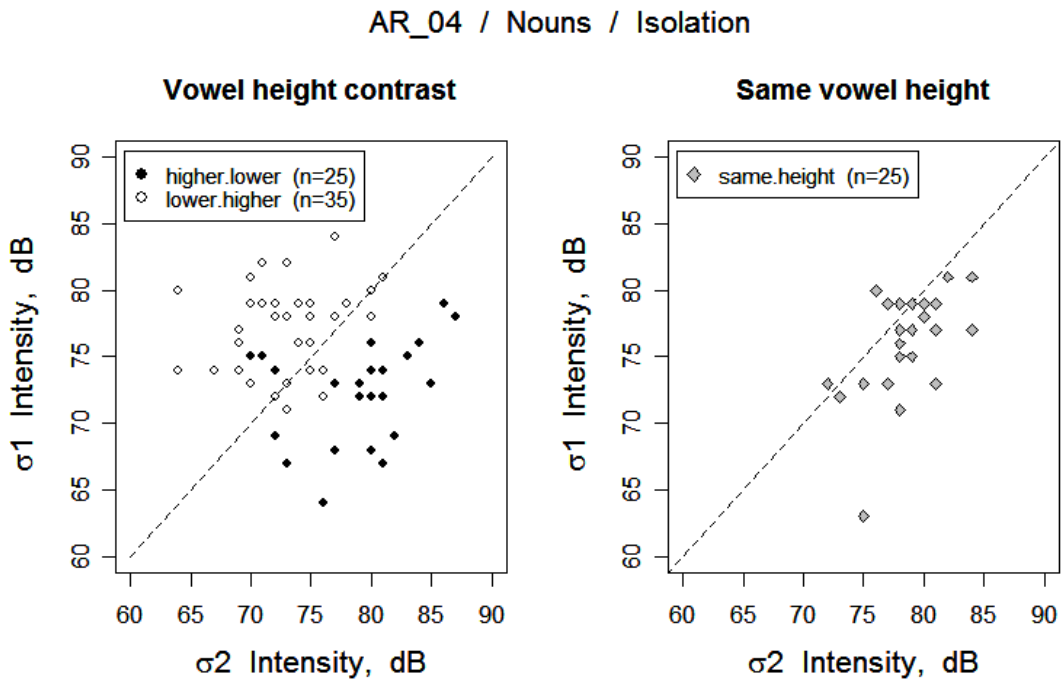
If intensity were a robust correlate of the σ_2 stress perceived on AR_04 nouns, then one would expect to see a consistently higher intensity on σ_2 . However, as shown in Figure 5.89 below, for both isolation and frame forms intensity is higher on σ_1 in about half the words, and higher on σ_2 in about half the words, regardless of morphological structure.

Figure 5.89 AR_04 / Nouns: Intensity



These distributions reflect the interaction of two factors: (a) the intrinsic variation of intensity as a function of vowel height; and (b) the use of intensity to convey stress. In Figure 5.90 below, the isolation forms are plotted in terms of vowel height in the two syllables of a word. (The eight nouns with a diphthong in $\sigma 1$ have been excluded.) In the plot to the left, intensity is usually higher on whichever syllable has a lower vowel, as one would predict following Lehiste (1970): [lower.higher] forms – such as [k^ha.'tu] *kha.thum?* ‘lid, cover’ and [dzo.'mu] *dzo.mo* ‘female dzo’ – fall to the left of the dashed line, while [higher.lower] forms – such as [ʃi.'mo] *sras.mo* ‘girl, daughter’ and [lex.'ka] *las.ka* ‘work’ fall to the right of the dashed line.

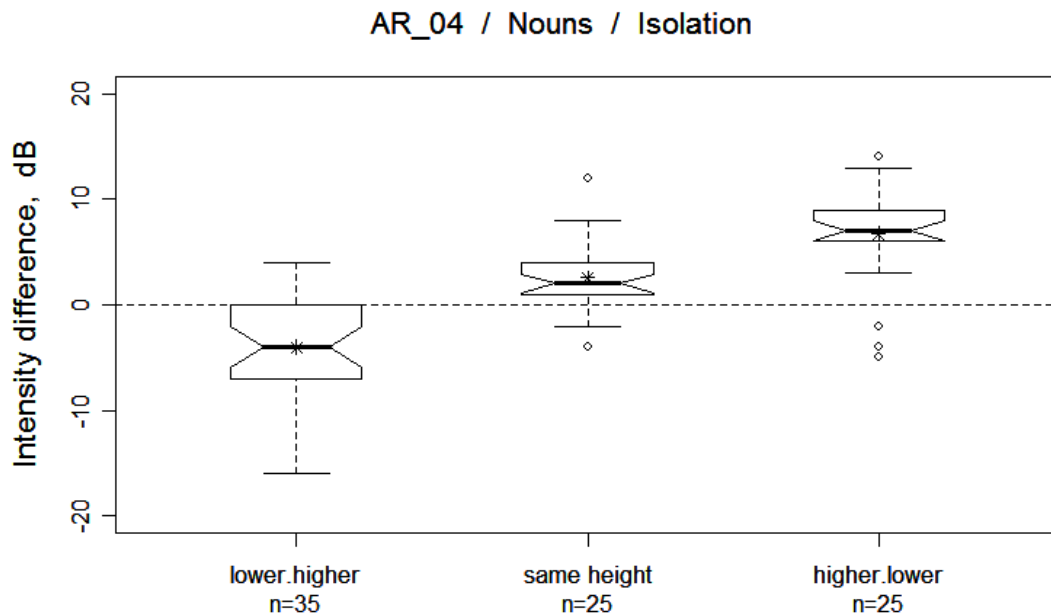
Figure 5.90 AR_04 / Nouns / Isolation: Intensity vs. vowel height contrast



The plot to the right above shows that in the control group – nouns in which the vowels have the same height in both syllables, such as [na.'ma] *mna'.ma* ‘bride’ and [ʔgʁn.'dʁm] *rgun.'brum* ‘grape’ – intensity is most often higher on σ_2 , with most points falling to the right of the dashed line rather than being distributed evenly to either side. This indicates a correspondence between stress and intensity.

The box-and-whisker plot in Figure 5.91 below illustrates the variation in intensity contrast as a function of vowel height contrast. The [same.height] group suggests that the speaker *does* manipulate intensity to convey stress. This correlation is obscured in the [lower.higher] group, and exaggerated in the [higher.lower] group.

Figure 5.91 AR_04 / Nouns / Isolation: Intensity differences vs. vowel height contrasts



Paired-sample t-tests for the isolation forms are summarized in Table 5.19 below. For the [lower.higher] nouns – where intensity corresponds with vowel height and is greater on $\sigma 1$ – the difference across syllables is statistically highly significant ($p \ll 0.05$), and even the minimum probable mean contrast represented by the upper 95% confidence limit (-2.4 dB) is likely to be of perceptual significance. If a listener were focusing on intensity as a cue for stress, s/he would receive acoustic information which contradicts the stress pattern.

Table 5.19 AR_04 / Nouns / Isolation / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	-5.1543	34	1.084e-05	-4.0	-5.6	-2.4
same height	3.7872	24	0.0009006	2.5	1.2	4.0
higher.lower	6.9743	24	3.272e-07	6.7	4.7	8.7

In the [higher.lower] nouns – where intensity is exaggerated on σ_2 because of the vowel height contrast – the difference across syllables is again statistically highly significant ($p \ll 0.05$), averaging 6.7 dB. Even the minimum probable mean difference represented by the lower 95% confidence limit (4.7 dB) would certainly be highly perceptible.

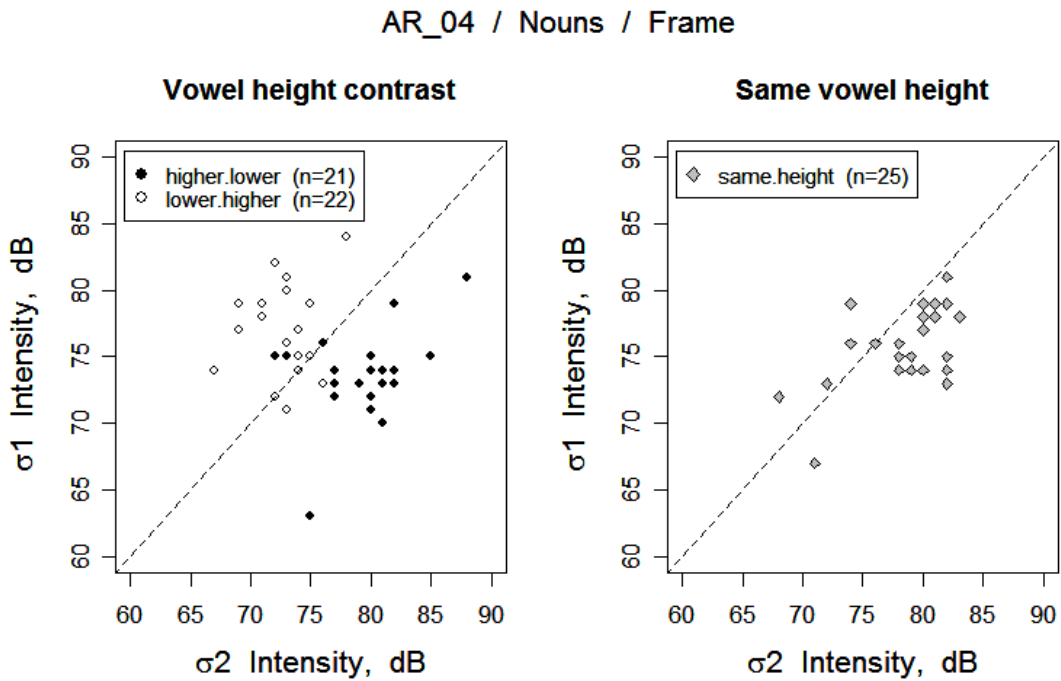
The control group – the subset of nouns which have vowels of the same height in both syllables – provides the most genuine measure of the degree to which speaker AR_04 manipulates intensity to convey stress, since the effects of vowel height contrast are neutralized. Here, the difference in intensity across syllables is statistically significant ($p < 0.05$), with a mean increase of 2.5 dB. However, the minimum probable mean difference indicated by the lower 95% confidence limit (1.2 dB) is barely greater than the just-noticeable difference limen of 1 dB (Lehiste 1970).

For isolation forms of nouns produced by speaker AR_04, then, intensity cannot be considered a consistent and robust cue for stress – an acoustic cue by which the listener can reliably identify the stress pattern. Rather, it is only a limited and weak correlate. In the control group, the contrast across syllables may not always be

highly perceptible. And in the [lower.higher] subset, intensity is higher on σ_1 , the syllable which is *not* stressed.

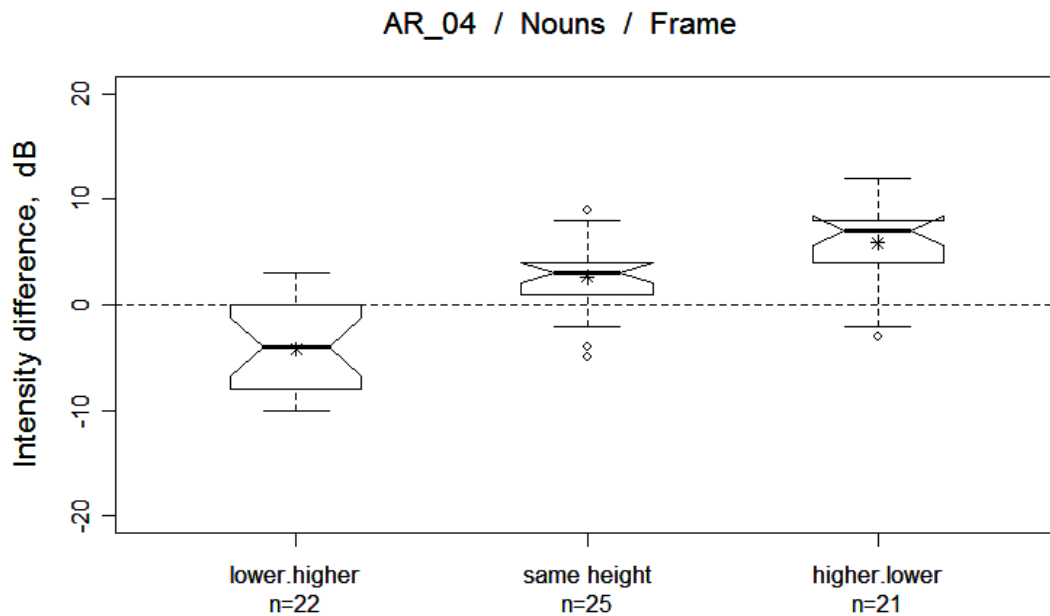
Nouns produced in the sentence frame by this speaker show almost exactly the same patterns. Frame forms are plotted in terms of vowel height contrast in Figure 5.92 below. (The six words with a diphthong in σ_1 were excluded.) The plot on the left shows that intensity is usually higher in whichever syllable has the lower vowel. The plot on the right shows that, when there is no contrast in vowel height across syllables, intensity is usually higher on σ_2 , corresponding with the perceived stress pattern.

Figure 5.92 AR_04 / Nouns / Frame: Intensity vs. vowel height contrast



The trend in intensity variation as a function of vowel height contrast for the frame forms is illustrated in the box-and-whisker plot in Figure 5.93 below. As was the case with the isolation forms, we can see here that the correspondence between intensity and stress is obscured in the [lower.higher] nouns – where the effects of vowel height and stress oppose one another – and exaggerated in the [higher.lower] nouns – where the effects of vowel height and stress reinforce one another.

Figure 5.93 AR_04 / Nouns / Frame: Intensity differences vs. vowel height contrasts



Again, it is the subset with no contrast in vowel height across syllables that provides the crucial information. Here, the increase in intensity is statistically significant ($p < 0.05$), with an average increase of 2.6 dB. However, with a lower 95% confidence limit of 1.3 dB – not much greater than the difference limen of 1.0 dB – the intensity difference may be barely perceptible.

Table 5.20 AR_04 / Nouns / Frame / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	-4.9686	21	6.456e-05	-4.2	-5.9	-2.4
same height	3.9519	24	0.000595	2.6	1.3	4.0
higher.lower	6.788	20	1.336e-06	5.9	4.1	7.7

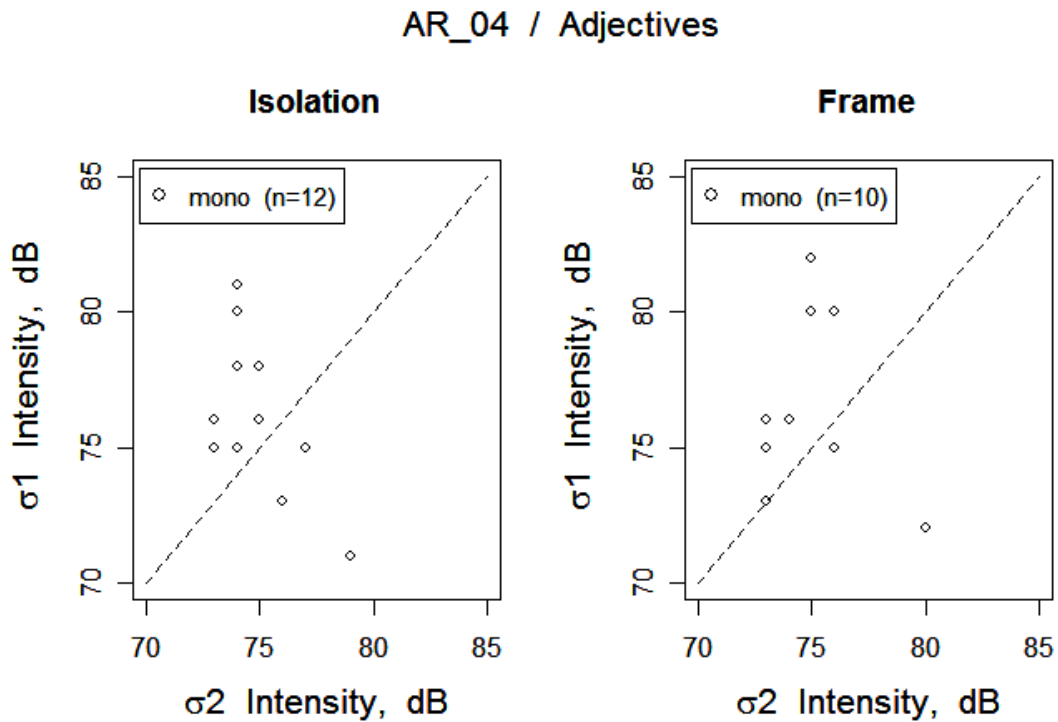
In conclusion, for the AR_04 frame forms – as for the isolation forms – intensity shows a correlation with the perceived σ_2 stress pattern only for a limited, controlled sample, and even there, the intensity contrast is not always highly perceptible.

5.4.1.2 AR_04 Adjectives

Though the sample size is not very large, AR_04 adjectives tend to show a higher intensity on σ_1 , in contradiction to the perceived pattern of σ_2 stress.

Intensity measurements for all isolation and frame forms are plotted in Figure 5.94 below. (Several data points in these plots have identical coordinates, and so overlap.) Most of the adjectives plot to the left of the dashed line, indicating a higher intensity on σ_1 . For all we know at this point, though, the distribution might be entirely attributable to contrasts in vowel height across syllables.

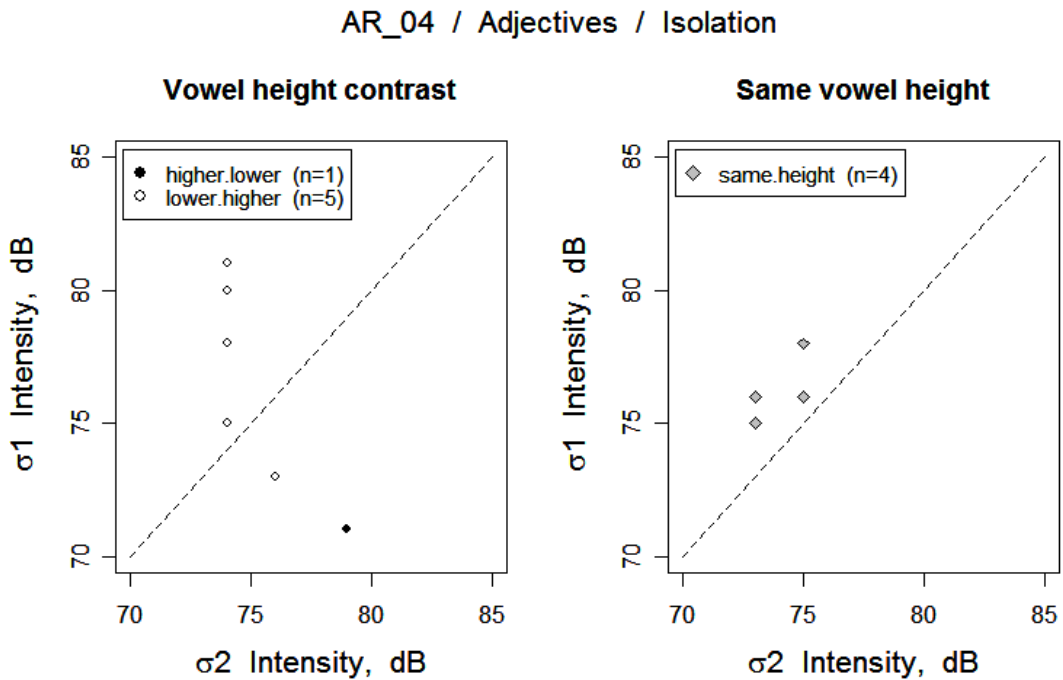
Figure 5.94 AR_04 / Adjectives: Intensity



To test this, the isolation forms are plotted in terms of vowel height in Figure 5.95 below. (The sample size is now even smaller: only ten isolation forms and eight frame forms remain once adjectives with diphthongs are excluded.)

When there is a contrast in vowel height across syllables – as in the graph on the left – intensity seems to behave as one would predict following Lehiste (1970): intensity is usually higher on whichever syllable has the lower vowel. For instance, the [lower.higher] adjective [xkam.'bo] *skam.po* ‘dry, dried’ plots to the left of the dashed line, while the lone [higher.lower] adjective [s^ho.'ma] *so.ma* ‘new’ plots to the right of the dashed line.

Figure 5.95 AR_04 / Adjectives / Isolation: Intensity vs. vowel height contrast

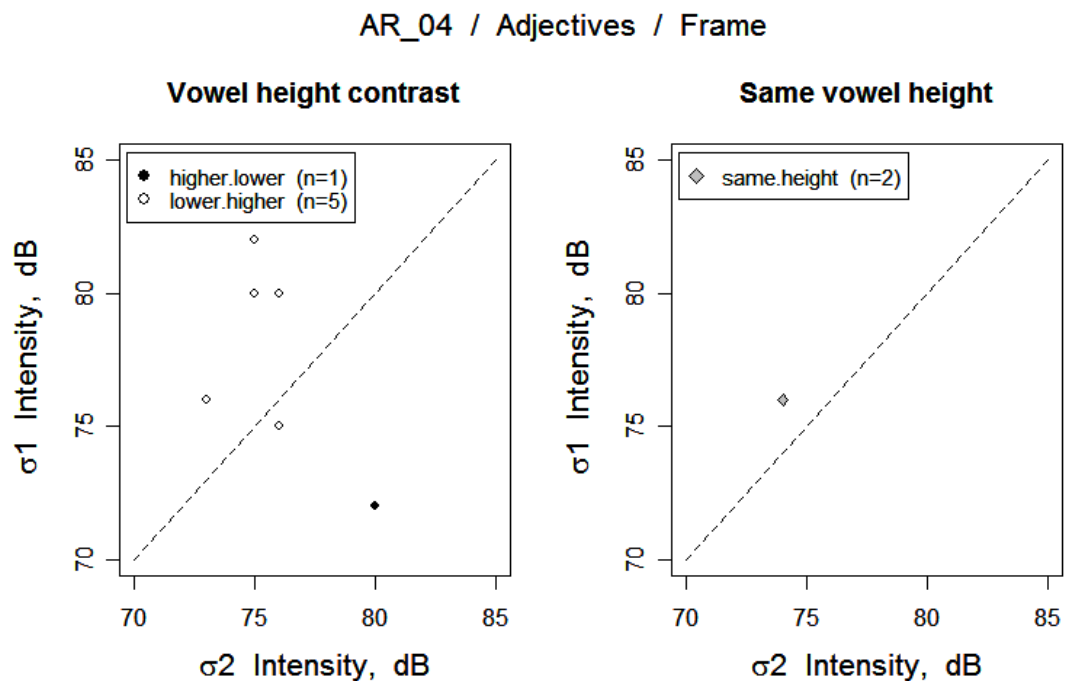


When there is no contrast in vowel height across syllables – as in the graph on the right above – one would predict, following Lehiste, that adjectives would be evenly distributed to either side of the dashed line. And if intensity were a robust correlate of stress, one would predict that all of them would fall to the right of the dashed line. Instead, all of them plot to the left, indicating a higher intensity on $\sigma 1$. These include forms such as [tʰe.'po] *chen.po* ‘big’ and [ʰrʌ.'ŋo] *ring.po* ‘long’.

The frame forms exhibit similar patterns, though the sample is smaller still. In Figure 5.96 below, all but one of the [lower.higher] adjectives have a higher intensity on $\sigma 1$, falling to the left of the dashed line, while the lone [higher.lower] adjective has a higher intensity on $\sigma 2$, and falls to the right. These contrastive cases again indicate

that vowel height has a strong influence on intensity. The two [same.height] adjectives – [tʰe.'po] *chen.po* ‘big’ and [ɕʰɤ.'ko] *phyug.po* ‘rich’ – overlap at the same (σ_2 , σ_1) coordinates, (74, 76), and so appear as a single point in the graph on the right. Of course, two tokens do not comprise a convincing sample, but it is noteworthy all the same that neither of them exhibit a correspondence between intensity and stress.

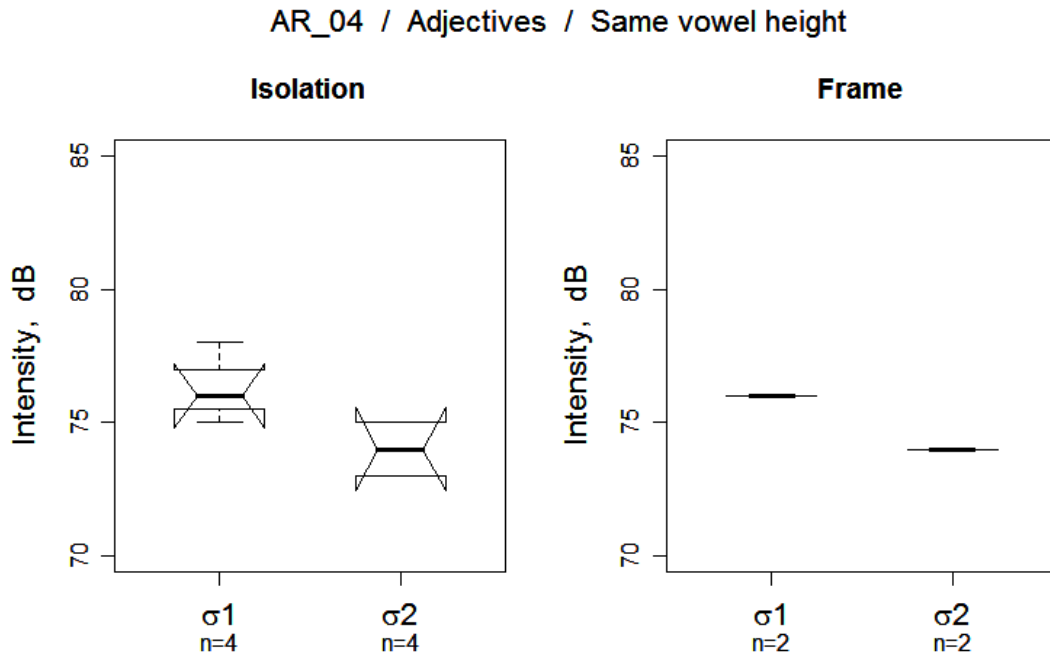
Figure 5.96 AR_04 / Adjectives / Frame: Intensity vs. vowel height contrast



For isolation and frame forms, box-and-whisker plots for the control group – the adjectives with vowels of the same height in both syllables – are shown in Figure 5.97 below. The median intensity is higher on σ_1 than on σ_2 for adjectives in both settings. For the isolation forms, the boxes representing the span of measurements

overlap slightly; for the frame forms there is no box at all because the two adjectives had the same values on both syllables.

Figure 5.97 AR_04 / Adjectives: Intensity with vowels of same height



Paired-sample t-tests for the control groups of isolation and frame forms are summarized in Table 5.21 below. For the isolation forms the difference in intensity is statistically significant, with a p-value < 0.05 ; the mean decrease across syllables is -2.2 dB. The contrast of -0.73 dB indicated by the upper 95% confidence limit is almost certainly not perceptible. For the frame forms, a paired-sample t-test could not be completed since the two adjectives in this group have exactly the same intensity on both syllables.

Table 5.21 AR_04 / Adjectives with same vowel height / Intensity: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
						lower	upper
AR_04	isolation	-4.7001	3	0.01822	-2.2	-3.8	-0.73
	frame	-	1	-	-2.0	-	-

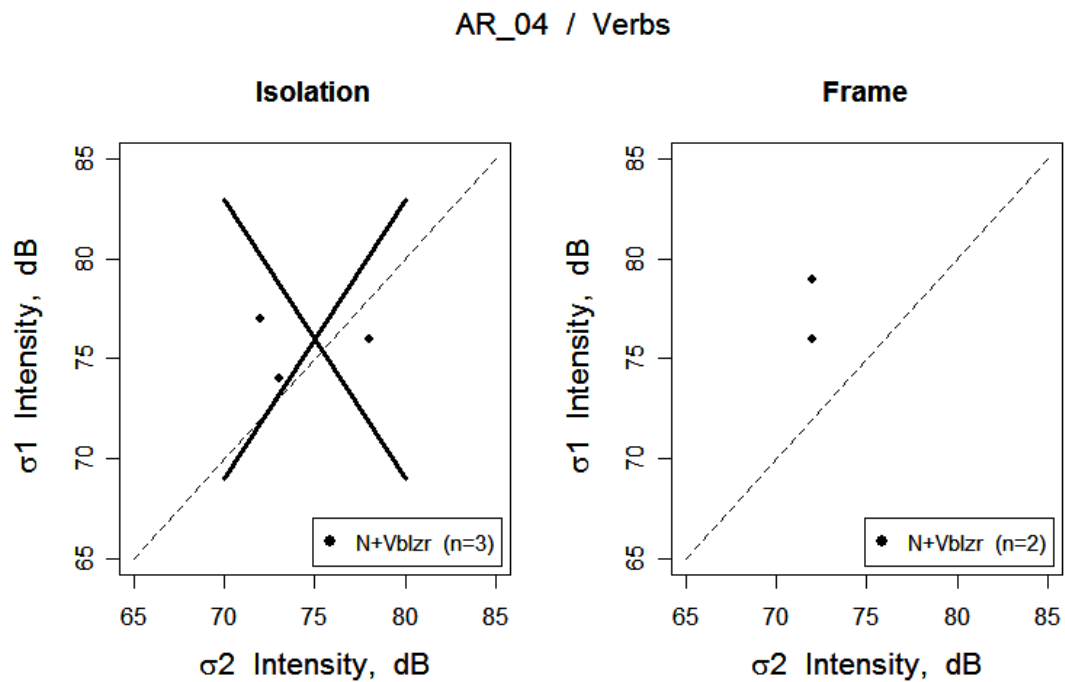
In summary, even though the sample size is quite small, there is certainly no evidence that intensity correlates with σ_2 stress in AR_04 adjectives. On the contrary, it seems that intensity is generally higher on σ_1 , except in the few cases where it is drawn to σ_2 by a contrast in vowel height.

5.4.1.3 AR_04 Verbs

Very little can be said about the behavior of intensity in AR_04 verbs. As noted above, this speaker was not comfortable producing verbs without the context of tense / aspect / evidentiality, with the result that the isolation forms were awkwardly produced and cannot be considered representative. For the frame forms, the sample is too small to control for incidental factors, and so no conclusive findings can be reached.

Intensity measurements for isolation and frame forms are plotted in Figure 5.98 below. The two frame forms which could be segmented and analyzed have a higher intensity on σ_1 , suggesting a correlation with stress.

Figure 5.98 AR_04 / Verbs: Intensity



However, one of the two verbs – [ˈnɒm.biʊ] *gnam.babs* ‘to rain’ – has a [lower.higher] vowel height contrast, and is predicted, on that basis, to have a higher intensity in σ_1 . Thus it is impossible to determine whether the observed intensity contrast is a reflection of stress, or a reflection of vowel height. The one verb with vowels of the same height in both syllables – [ˈt͡ʂɛ̃.d͡ʒɣʊ] *lan.brgyab* ‘to answer’ – also has a higher intensity on σ_1 , the stressed syllable. But one token is hardly sufficient evidence to claim a correlation between intensity and stress.

All that can be said with certainty here is that neither of these examples has a higher intensity on σ_2 , which would contradict the perceived stress pattern. It is

possible that a larger sample might demonstrate that intensity is indeed a cue for stress.

5.4.2 Intensity for speaker AR_05

For speaker AR_05, intensity shows a clear correlation with $\sigma 1$ stress in verbs – both in isolation and in the sentence frame – and with $\sigma 2$ stress in numerals, recorded only in isolation. Intensity shows a limited and weak correlation with the $\sigma 2$ stress perceived on nouns. For adjectives, the sample is too small to be definitive, but the data suggests that intensity may be a correlate of stress for the forms forms, though not for the isolation forms.

Figure 5.99 below shows the distribution of intensity measurements for both syllables of words produced in isolation by speaker AR_05. For nouns and adjectives, since vowel height has such a strong effect on intensity, analysis was based on the control group of words with vowels of the same height in both syllables – so these are the subsets represented here. For verbs, contrasts in vowel height were not relevant to the analysis; since the analysis was thus based on the full set of verbs, that is what is represented here.

For the control group of nouns, the increase in intensity from $\sigma 1$ to $\sigma 2$ is statistically significant, but – as was the case for AR_04 nouns produced in isolation – a statistically significant contrast may be barely greater than the just-noticeable difference. Thus the correlation between intensity and stress exhibited by AR_05 isolation nouns can only be considered limited and weak.

For adjectives, the small sample does not suggest a correlation between intensity and $\sigma 2$ stress. In the box-and-whisker plot below there is almost complete overlap of the measurements in the two syllables.

Like the other non-verbs, numerals are stressed on $\sigma 2$, though they are not shown here for lack of space in the graph. As demonstrated in section 5.4.2.3 below, intensity shows a clear correlation with stress for this group.

AR_05 verbs are stressed on $\sigma 1$, and in all tokens – regardless of vowel height contrast – intensity is also higher on $\sigma 1$.

Figure 5.99 AR_05 / Isolation: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)

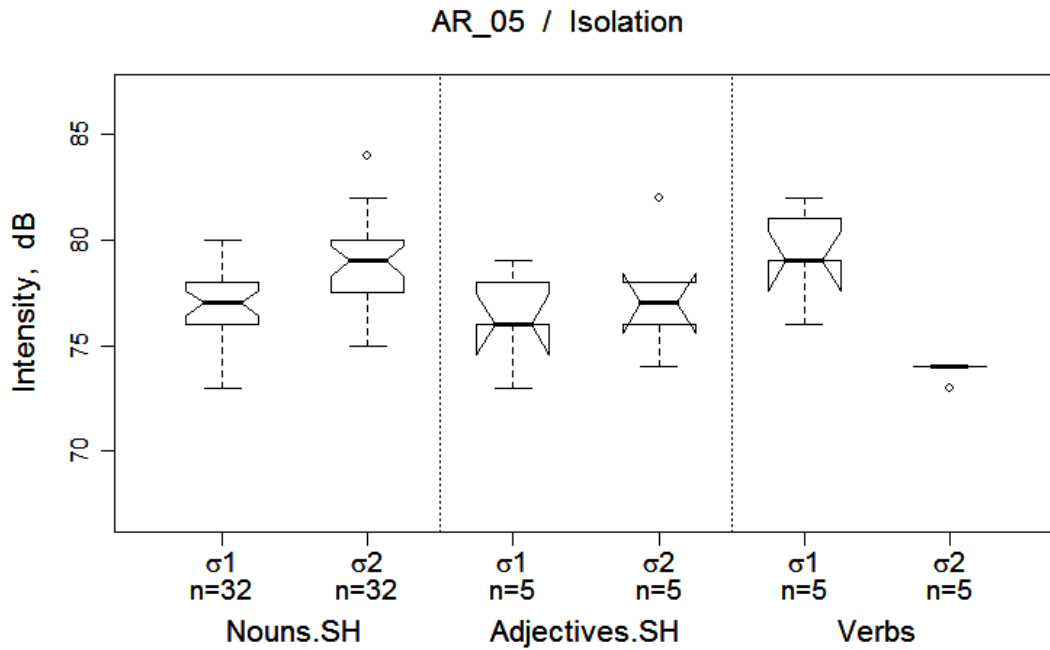
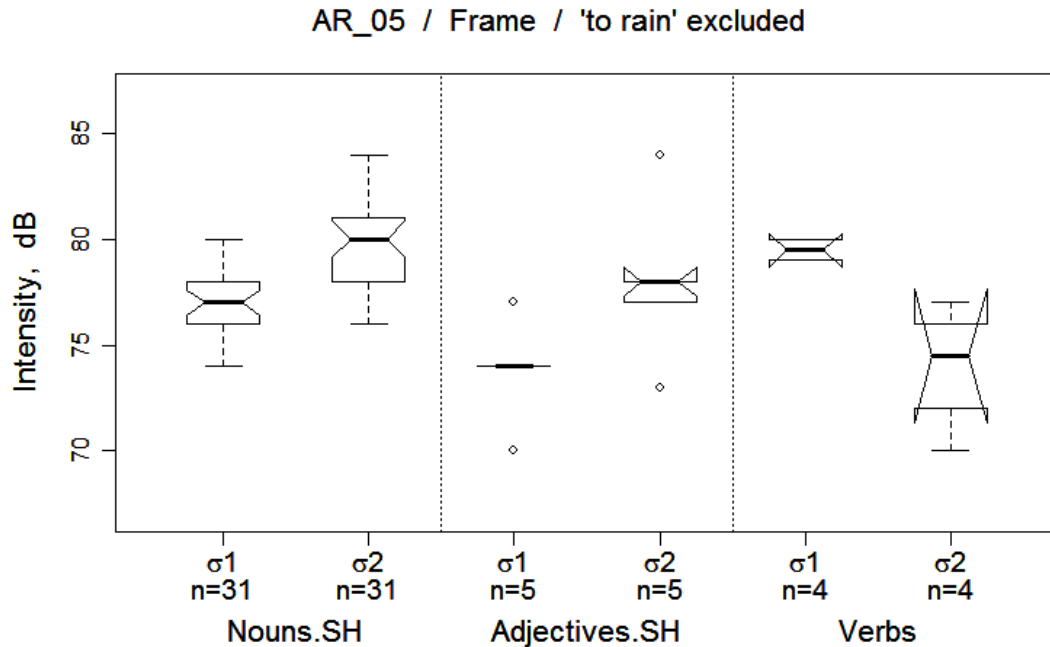


Figure 5.100 below shows the distribution of intensity measurements for AR_05 frame forms. Again, only the subsets on which the final analysis was based are represented.

For both nouns and adjectives, intensity is a correlate of σ_2 stress in the control group, but this does not extend to all height groups: the correlation with stress can be obscured by the intrinsic variation of intensity with vowel height. Intensity can thus be considered to show only a limited and weak correlation with stress for nouns and adjectives produced in the sentence frame.

In verbs produced in the sentence frame, the contrast in intensity across syllables is statistically significant. Given the variability in σ_2 , however, this contrast may be barely above the just-noticeable difference.

Figure 5.100 AR_05 / Frame: Intensity contrasts (“SH” indicates that only words with vowels of the same height in both syllables are plotted)



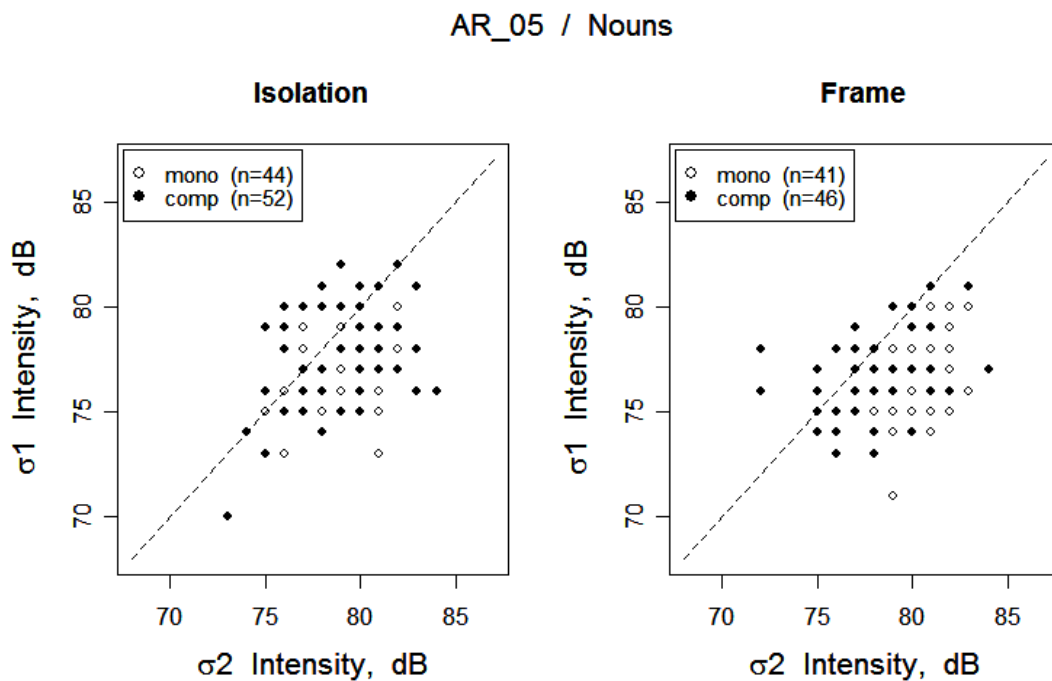
Intensity data for AR_05 nouns, adjectives, numerals and verbs is considered in detail in sections 5.4.2.1 through 5.4.2.4 below.

5.4.2.1 AR_05 Nouns

Intensity measurements for both isolation and frame forms of AR_05 nouns are plotted in Figure 5.101 below. (Note that a number of points plot on top of each other and so are not visible.) As shown, the majority of nouns have a higher intensity on $\sigma 2$ – the stressed syllable – plotting to the right of the dashed line. This is particularly true of the frame forms. Here, only a few compound nouns like [var.'tɕ^hɪl] *bar.chad* ‘obstruction’ and [ra.'luy] *ra.lug* ‘goats and sheep’ plot to the left; these

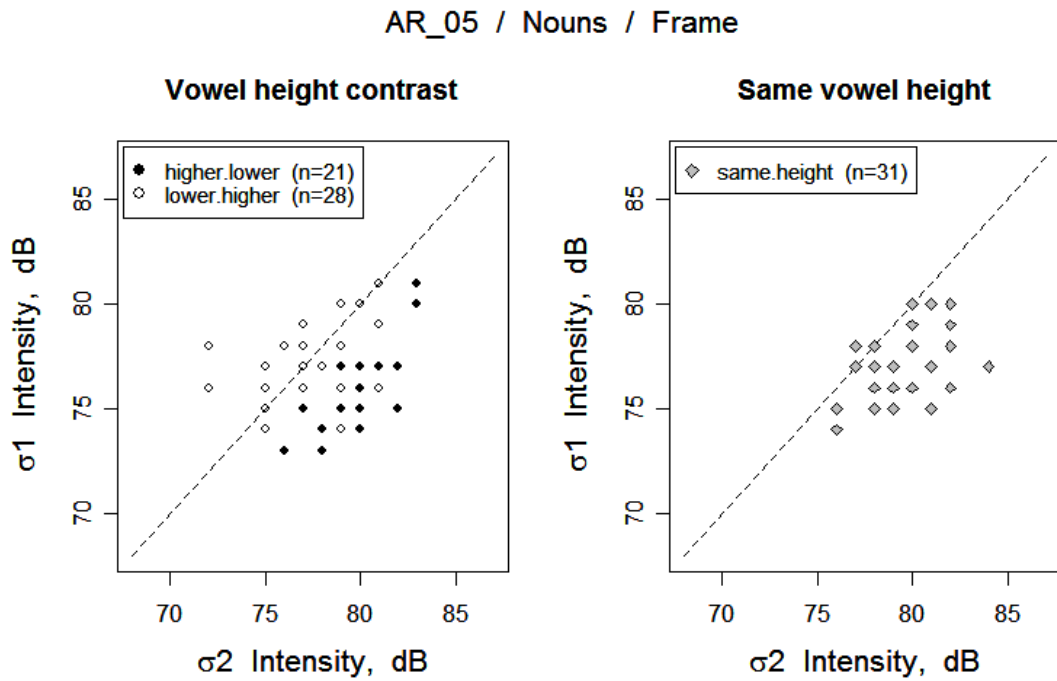
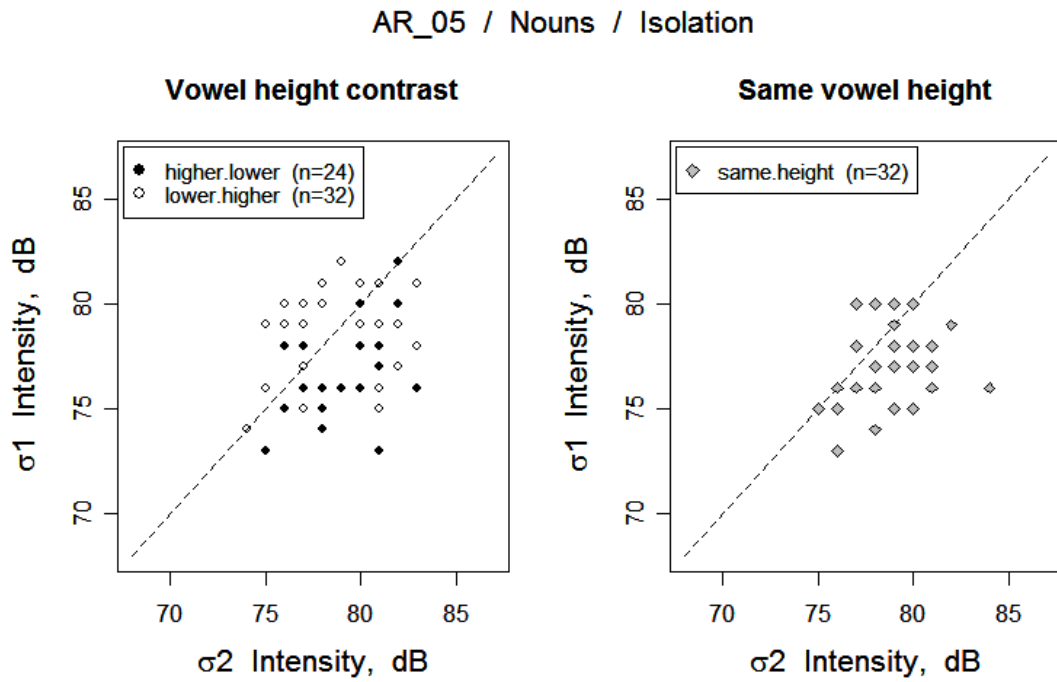
have a [lower.higher] vowel height contrast, suggesting that this factor may play a strong role in the distribution.

Figure 5.101 AR_05 / Nouns: Intensity



Indeed, the importance of vowel height is confirmed by the plots in Figure 5.102 below. For both isolation and frame forms, when there is a contrast in vowel height across syllables – as in the plots on the left – there is a rough tendency for intensity to be higher on whichever syllable has a lower vowel, as one would predict following Lehiste (1970). However, the influence of stress appears to sometimes override this tendency. That is, nouns in the [lower.higher] group might be predicted – on the basis of vowel height alone – to consistently have a higher intensity on $\sigma 1$. Instead, even some of these cases have a considerably higher intensity on $\sigma 2$, the stressed syllable.

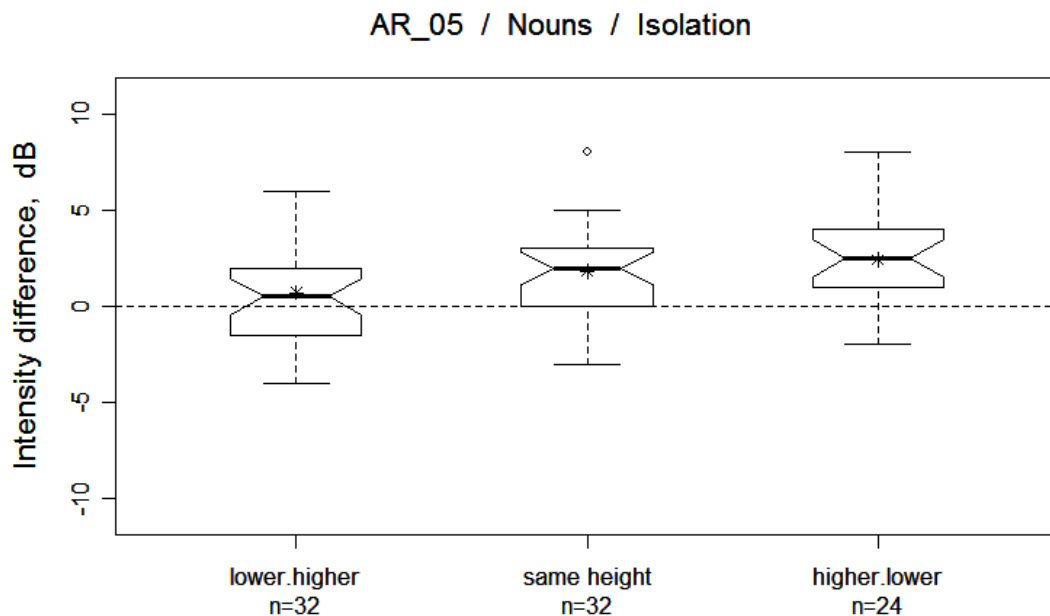
Figure 5.102 AR_05 / Nouns / Isolation: Intensity vs. vowel height contrast



Nouns with vowels of the same height in both syllables are plotted in the graphs to the right above. For both isolation and frame forms, nearly all nouns in these control groups have a higher intensity on σ_2 , indicating a correspondence with the perceived σ_2 stress pattern.

The box-and-whisker plot in Figure 5.103 below illustrates the correspondence between intensity contrast and vowel height contrast for the isolation forms. In the [lower.higher] group, the notch spans the x-axis, indicating that the median difference in intensity across syllables for this subset is not significantly different than zero.

Figure 5.103 AR_05 / Nouns / Isolation: Intensity differences vs. vowel height contrasts



Paired-sample t-tests, summarized in Table 2.8 above, confirm that the contrast for this [lower.higher] group is not statistically significant, with $p > 0.05$ and

a 95% confidence interval that includes zero. For nouns with a [higher.lower] vowel height contrast, the intensity increase from σ_1 to σ_2 is statistically significant ($p \ll 0.05$), with mean differences across syllables of 2.4 dB.

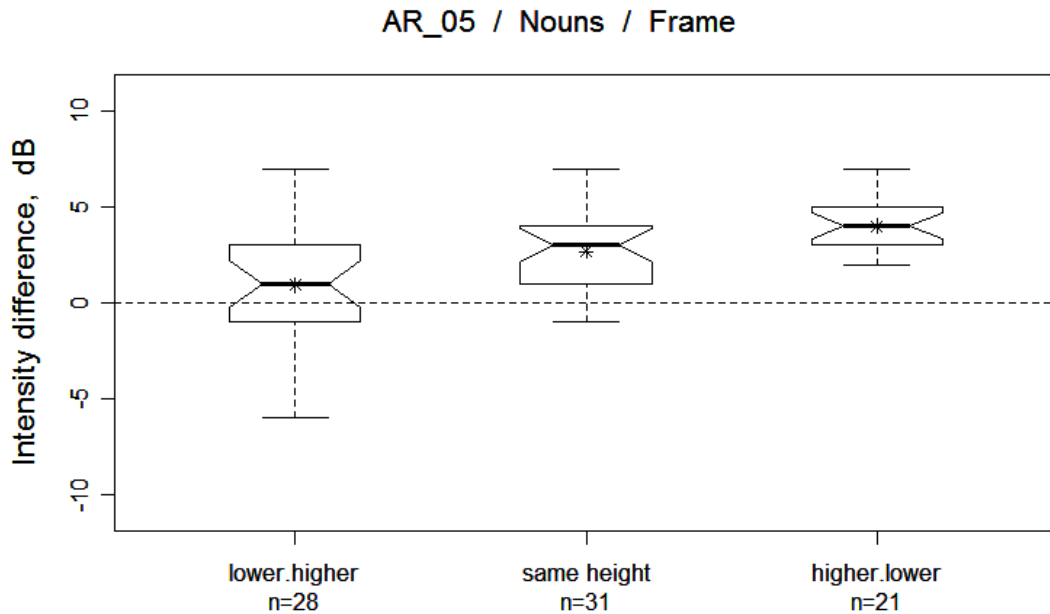
Table 5.22 AR_05 / Nouns / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.442	31	0.1593	0.72	-0.30	1.7
same height	4.529	31	8.243e-05	1.8	1.0	2.6
higher.lower	5.144	23	3.266e-05	2.4	1.4	3.4

For the crucial control group – the [same.height] nouns – the intensity increase is statistically significant ($p \ll 0.05$), but a statistically significant mean increase may be as small as 1 dB (the lower 95% confidence limit), which is barely perceptible (Lehiste 1970). Thus for AR_05 nouns produced in isolation, we can say that only a limited subset of the data shows a correlation between intensity and stress, and the contrast in intensity across syllables may not always be highly perceptible. Thus intensity can only be considered a limited and weak correlate of stress. (This was also the case for nouns produced by speaker AR_04.)

The correspondence between intensity contrast and vowel height for the frame forms is illustrated in the box-and-whisker plot in Figure 5.104 below. Nouns in the [same.height] group show a more distinct increase in intensity across syllables than was the case for the isolation forms.

Figure 5.104 AR_05 / Nouns / Frame: Intensity differences vs. vowel height contrasts



A paired-sample t-test, summarized in Table 5.23 below, confirms that the contrast for the control group is statistically significant, with $p \ll 0.05$ and a mean intensity difference of 2.7 dB. Even the smallest probable mean difference in intensity of 1.9 dB (the lower 95% confidence limit) is somewhat greater than Lehiste’s just-noticeable difference limen of 1.0 dB. For this limited subset, then, I consider intensity to be a meaningful correlate of stress.

Table 5.23 AR_05 / Nouns / Frame / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	1.684	27	0.1037	0.93	-0.20	2.06
same height	7.492	30	2.373e-08	2.7	1.9	3.4
higher.lower	14.4914	20	4.539e-12	4.0	3.4	4.6

For nouns in the [lower.higher] group, the difference in intensity across syllables is not significant: $p > 0.05$ and the 95% confidence interval includes zero. Here, the correspondence between intensity and stress is obscured, since the effects of vowel height and stress oppose one another. For the [higher.lower] nouns, the correspondence between intensity and stress is exaggerated, since the effects of vowel height and stress reinforce one another. The difference in intensity across syllables is highly significant ($p \ll 0.05$), with a mean difference of 4.0 dB and a lower 95% confidence limit of 3.4 dB, which are likely to be highly perceptible.

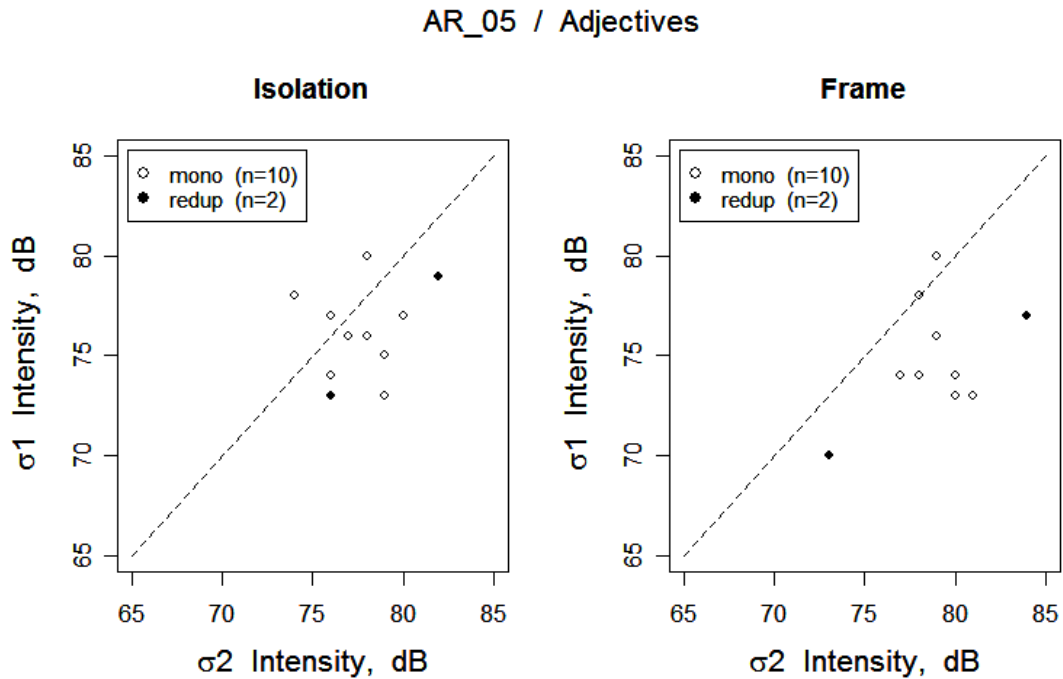
In conclusion, in isolation and in the sentence frame, nouns in the control group show that speaker AR_05 does, indeed, manipulate intensity in correspondence with the σ_2 stress pattern. This correlation is statistically significant, but may be perceptually weak. Furthermore, in nouns in the [lower.higher] group, the correlation between intensity and stress is not strong enough to prevail over the intrinsic effects of vowel height. Thus for AR_05 nouns, intensity can be regarded as only a limited and weak correlate of stress.

5.4.2.2 AR_05 Adjectives

Intensity measurements for AR_05 adjectives are plotted in Figure 5.105 below. The isolation forms fall to either side of the dashed line; sometimes intensity is higher in σ_1 , sometimes it is higher in σ_2 . Most of the frame forms fall to the right of the dashed line, indicating a higher intensity on σ_2 . In both settings, the two

reduplicated forms – [ʒɛv.'lɛv] *leb.leb* ‘flat’ and [tɕʰʊŋ.'tɕʰʊŋ] *chung.chung* ‘small’ – have a higher intensity on σ_2 , suggesting a correlation with stress.

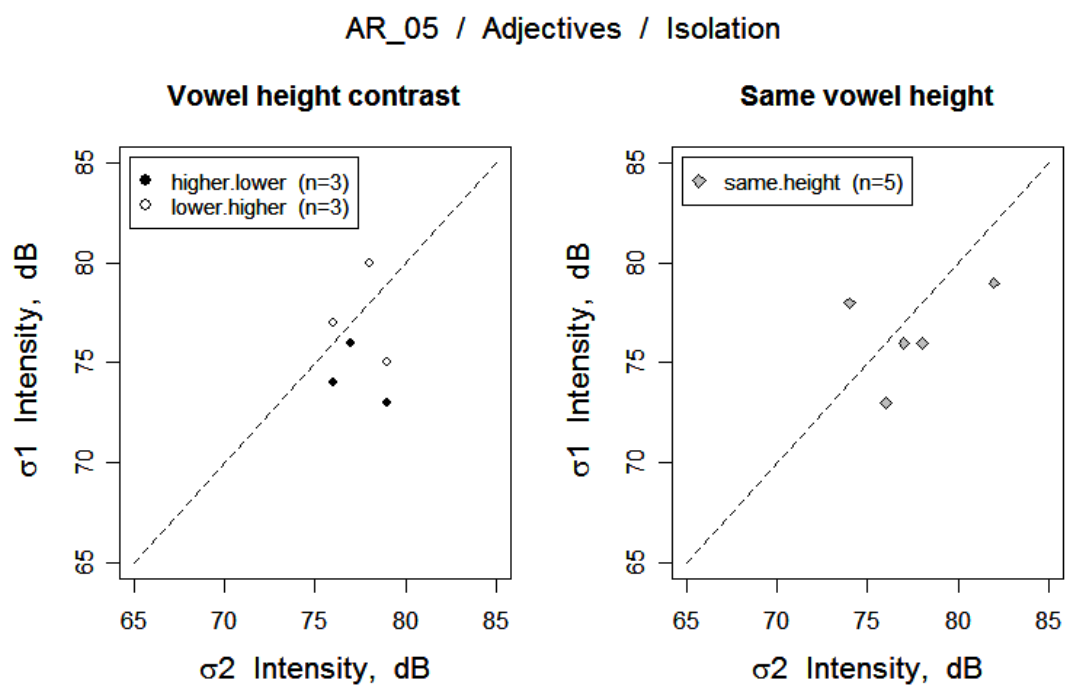
Figure 5.105 AR_05 / Adjectives: Intensity



For isolation forms, the potential effects of vowel height contrast are taken into consideration in the plot in Figure 5.106 below. (Only eleven of the twelve adjectives are plotted, since one contained a diphthong.) Overall, eight tokens have a higher intensity in σ_2 , the stressed syllable. However, the graph on the left side of the figure shows that, for three of these, the intensity pattern may be attributable to the contrast in vowel height – the [higher.lower] adjectives all fall to the right of the dashed line. (The one [lower.higher] adjective with a higher intensity on σ_2 is [tsʰan.'di] *tsha.'di* ‘hot’.)

The graph on the right shows that, when vowel height is the same in both syllables, most adjectives have a higher intensity on σ_2 . Some of the isolation forms plot quite close to the dashed line, though, indicating only a small contrast, and one – [no.'χo] *nag.po* ‘black’ – has a higher intensity on σ_1 .

Figure 5.106 AR_05 / Adjectives / Isolation: Intensity vs. vowel height contrast



Paired-sample t-tests for the isolation forms are summarized in Table 5.24 below. The contrast in intensity across syllables is not statistically significant for any of the vowel height groups; in all cases, $p > 0.05$ and the 95% confidence intervals include zero. For the control group, with such a small sample, the one adjective with a higher intensity on σ_1 has a considerable impact on the statistical analysis. A larger

sample might reveal a correlation between intensity and σ_2 stress. (Though on the other hand, it might not.)

Table 5.24 AR_05 / Adjectives / Isolation: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

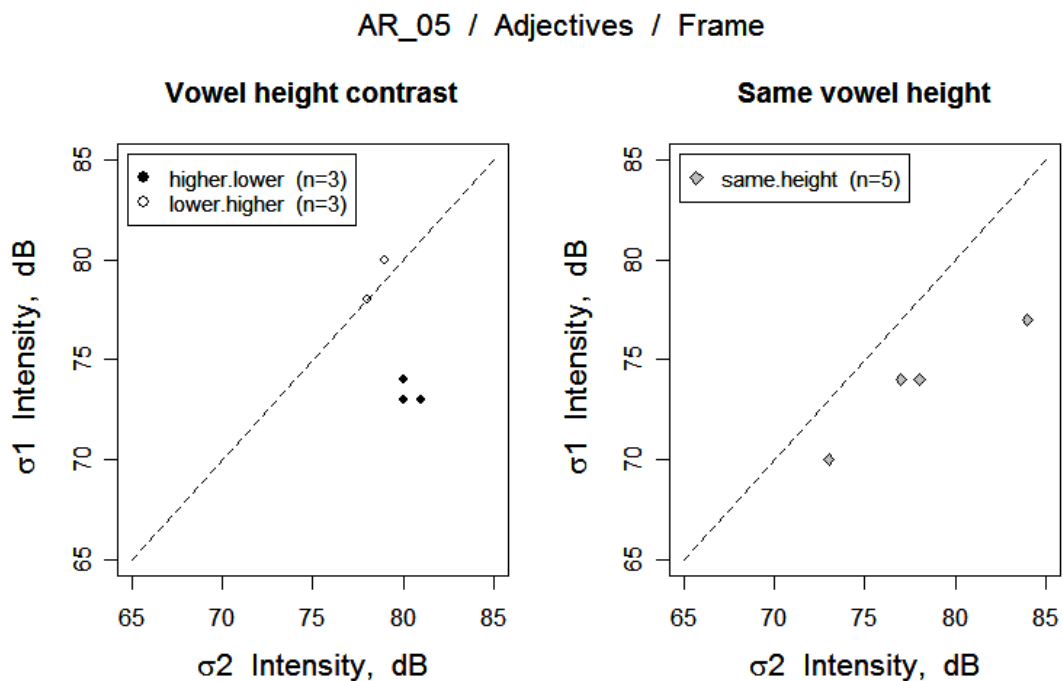
Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	0.1796	2	0.874	0.33	-7.6	8.3
same height	0.767	4	0.4859	1.0	-2.6	4.6
higher.lower	1.964	2	0.1885	3.0	-3.6	9.6

The frame forms of AR_05 adjectives are plotted in terms of vowel height in Figure 5.107 below. Again, as shown in the plot at left, there are three cases in which the higher intensity on σ_2 can be attributed to the intrinsic effects of a [higher.lower] vowel height contrast. Two of the [lower.higher] forms plot on or very close to the dashed line, suggesting that any potential correlations between intensity and stress and between intensity and vowel height neutralize one another. The third [lower.higher] form – [ts^han.'di] *tsha*. 'di' 'hot' – has a higher intensity on σ_2 (as was the case in isolation); this point plots at (σ_2 , σ_1) coordinates (80,74) and is obscured by one of the [higher.lower] forms.

Adjectives in the control group – those with vowels of the same height in both syllables, including [tɕ^he.'po] *chen.po* 'big' and [tɕ^hʊŋ.'tɕ^hʊŋ] *chung.chung* 'small', which overlap at coordinates (78, 74) – all have a higher intensity on σ_2 , plotting to

the right of the dashed line below. These cases suggests a correlation between intensity and $\sigma 2$ stress.

Figure 5.107 AR_05 / Adjectives / Frame: Intensity vs. vowel height contrast



Paired-sample t-tests for the frame forms are summarized in Table 5.25 below. For the control group, even though the sample is small, the test indicates that the difference in intensity across syllables is significant ($p < 0.05$), with a mean increase of 4.2 dB. Even the minimum probable mean increase of 2.2 dB is likely to be clearly perceptible. This subset suggests that intensity is a meaningful correlate of stress for AR_05 nouns produced in the sentence frame.

Table 5.25 AR_05 / Adjectives / Frame: Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Vowels	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	0.7625	2	0.5254	1.6	-7.7	11
same height	5.7155	4	0.004636	4.2	2.2	6.2
higher.lower	12.1244	2	0.006734	7.0	4.5	9.5

In the [lower.higher] nouns, the correlation between intensity and σ_2 stress is obscured by the intrinsic variation of intensity with vowel height. Conversely, in the [higher.lower] nouns it is exaggerated.

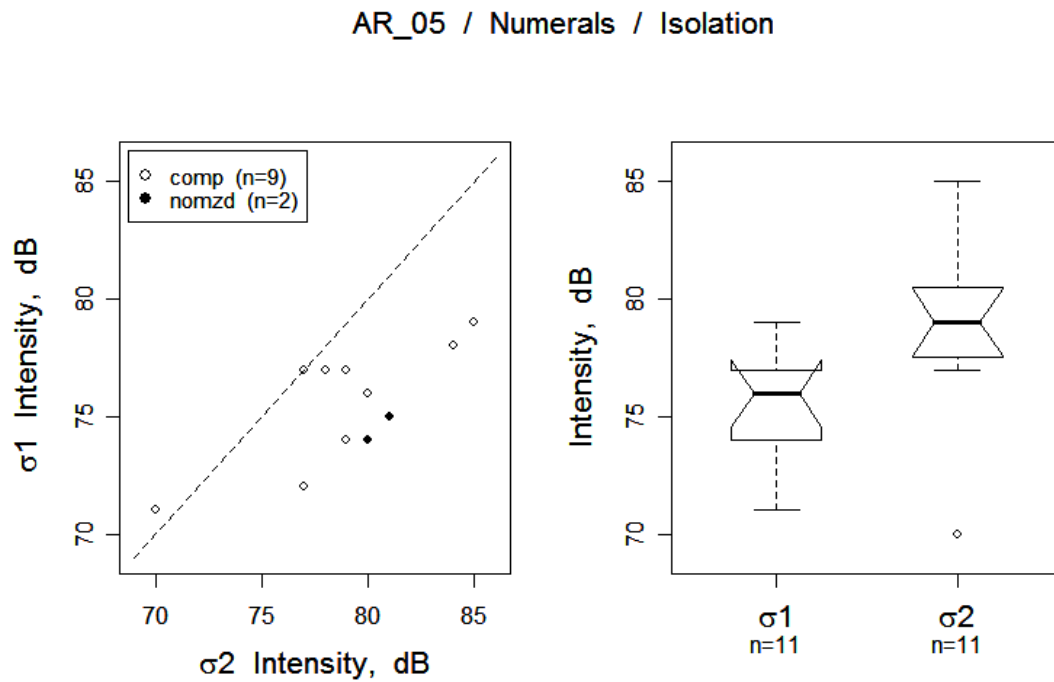
In conclusion, the contrast in intensity across syllables is not statistically significant for the AR_05 isolation forms of adjectives, even in the control group. For the frame forms, the control group suggests that intensity *is* a correlate of stress – at least when it is not overridden by the intrinsic effects of vowel height contrast.

5.4.2.3 AR_05 Numerals

Intensity measurements for numerals are graphed in Figure 5.108 below. As shown, nearly all of the numerals have a higher intensity on σ_2 , and so fall to the right of the dashed line. One exception is [^hdʌn.tʃɛŋ] *bdun.bcu* ‘seventy’, which falls exactly on the dashed line with intensity measured as 77 dB in both syllables. The other is [tʃɛŋ.tʃɪx] *bcu.drug* ‘sixteen’ at (σ_2 , σ_1) coordinates (70,71), in which both vowels are centralized and reduced. (This word has the lowest intensity values on both syllables of all 124 isolation forms analyzed for this speaker. It likely represents

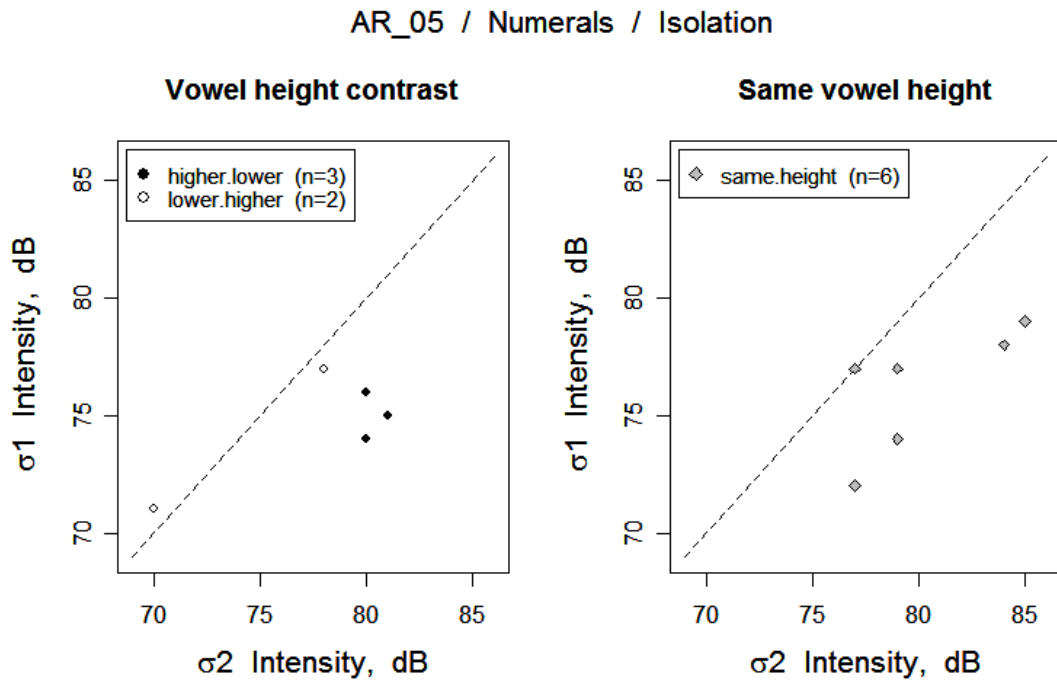
a case of mumbling.) The boxplot to the right in the figure below shows no overlap of either the notches – representing the 95% confidence interval about the median – or the boxes – representing the interquartile range.

Figure 5.108 AR_05 / Numerals: Intensity



In Figure 5.109 below, the numerals are plotted in terms of vowel height. The three [higher.lower] numerals have a higher intensity on σ_2 , which one might attribute to the vowel height contrast. But the [same.height] nouns also mostly have a higher intensity on σ_2 , which must reflect a correlation with stress. In fact, the only numeral with a higher intensity on σ_1 is the anomalous form [tʃɹ.'ʃix] *bcu.drug* ‘sixteen’.

Figure 5.109 AR_05 / Numerals: Intensity vs. vowel height contrast



Paired-sample t-tests comparing intensity across syllables are summarized in Table 5.26 below. For the control group – those numerals with vowels of the same height in both syllables – the intensity difference is statistically significant ($p < 0.05$), averaging 4.0 dB. When all the numerals but the anomalous ‘sixteen’ are considered, the mean intensity difference is 4.1 dB, and the lower 95% confidence limit is 2.5 dB. These findings indicate that intensity is a meaningful correlate of the σ_2 stress pattern perceived on the isolation form of numerals produced by speaker AR_05. (Frame forms were not recorded.)

Table 5.26 AR_05 / Numerals / Intensity by vowel height contrast: Paired-sample t-tests (two-tailed)

Subset	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
					lower	upper
lower.higher	0	1	1	0	-13	13
same height	4	5	0.01032	4.0	1.4	6.6
higher.lower	8	2	0.01527	5.3	2.5	8.2
all	4.5408	10	0.001073	3.6	1.8	5.4
all but '16'	5.6796	9	0.000302	4.1	2.5	5.7

5.4.2.4 AR_05 Verbs

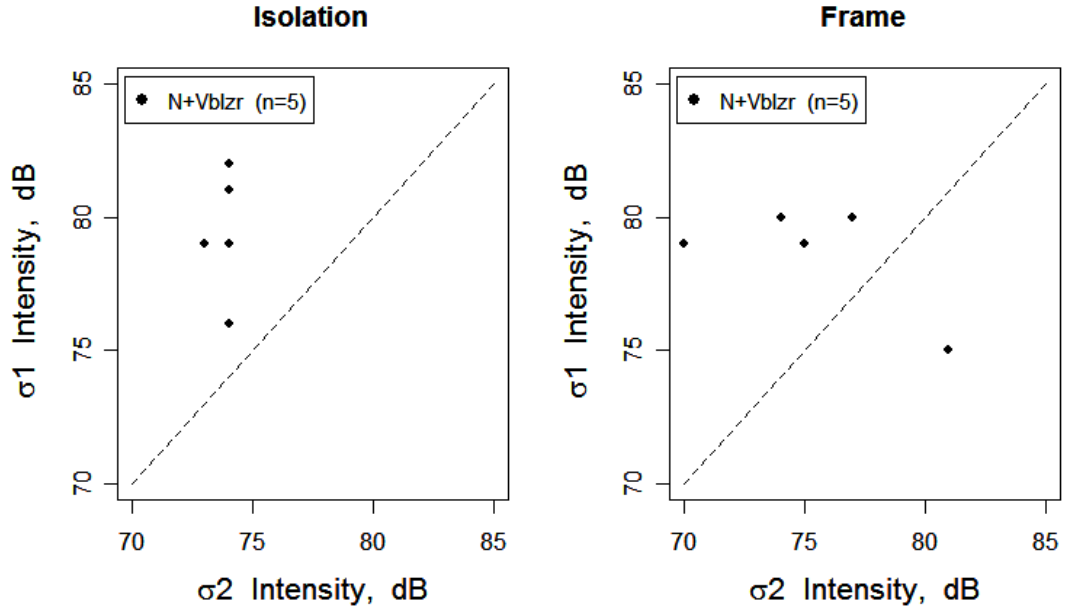
Intensity measurements for isolation and frame forms of verbs produced by speaker AR_05 are plotted in Figure 5.110 below. As shown, all of the verbs in both settings have a higher intensity on σ_1 – the stressed syllable – with the exception of the frame form of [nɒm.'bav] *gnam.babs* ‘to rain’, an anomaly which was discussed previously.

This distribution cannot be attributed to the intrinsic variation of intensity with vowel height, since the sample includes the [lower.higher] verb [nda.hen] *mda* ‘to shoot an arrow’, the [higher.lower] verb [vu.dzəp] *wu.brgyab* ‘to shoot a gun’, and three verbs with vowels of the same height in both syllables, including [k^xɒŋ.bɒp] *gangs.babs* ‘to snow’.

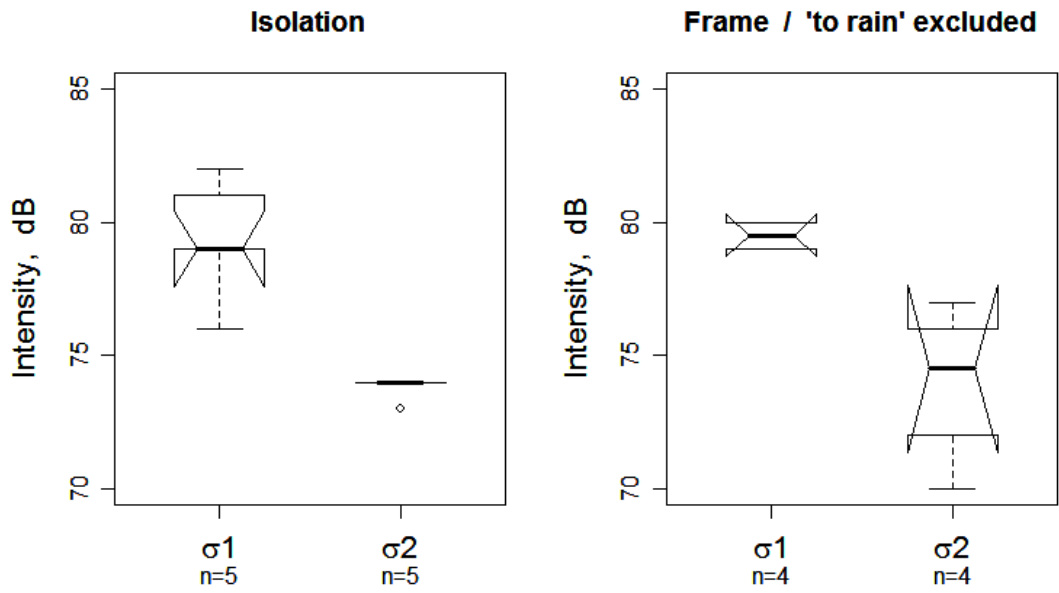
Ruling out this incidental factor, then, we can conclude that intensity is a correlate of the perceived σ_1 stress.

Figure 5.110 AR_05 / Verbs: Intensity

AR_05 / Verbs



AR_05 / Verbs



Paired-sample t-tests, summarized in Table 5.27 below, show that the difference in intensity is statistically significant for both isolation and frame forms: $p < 0.05$, and the mean difference across syllables is ~ -5.5 dB.

Table 5.27 AR_05 / Verbs / Intensity: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (dB)	95 % conf. limits	
						lower	upper
AR_05	isolation	-5.4392	4	0.005546	-5.6	-8.4	-2.7
	frame *	-4.1576	3	0.0253	-5.5	-9.7	-1.3

* The verb ‘to rain’ has been excluded.

5.4.3 The interaction of pitch and intensity

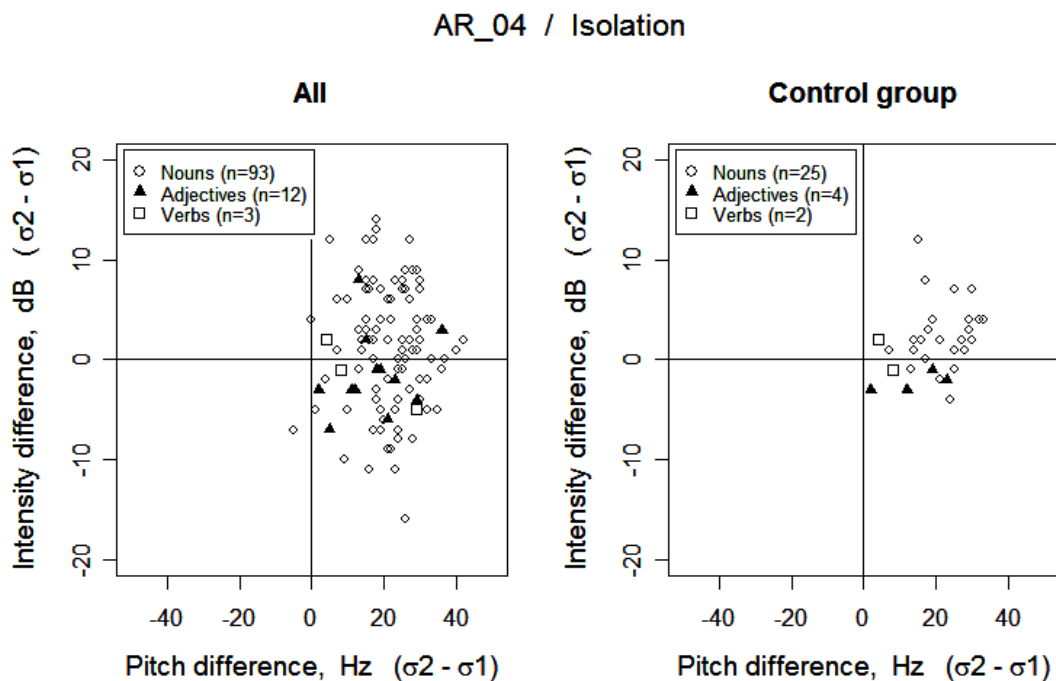
As was the case with Balti, pitch and intensity distinguish non-verbs (nouns, adjectives, and numerals) from verbs in Rebkong Amdo. The only exception to this is the set of isolation forms produced by speaker AR_04, which happens to be where we begin.

For the AR_04 isolation forms, pitch difference and intensity difference are plotted in Figure 5.111 below. The graph on the left includes all of the measurements. As shown, verbs join nouns and adjectives in having a higher pitch on σ_2 ; however, as discussed previously, I do not consider this verb data to be valid. Words from all lexical categories fall both above and below the “Intensity difference = 0” axis.

The graph on the right includes only words which have vowels of the same height in both syllables. Nouns in the control group generally have a higher intensity

on σ_2 , the syllable which is stressed; adjectives in the control group have a slightly higher intensity in σ_1 , the syllable which is not stressed.

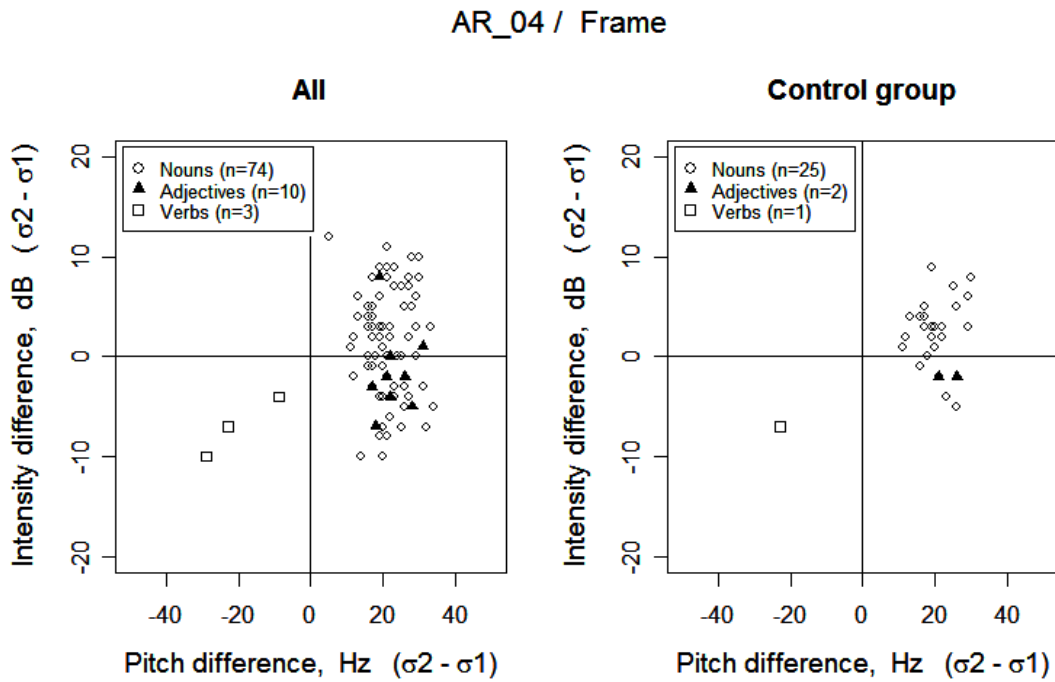
Figure 5.111 AR_04 / Isolation: Pitch difference vs. intensity difference



The frame forms produced by speaker AR_04 are plotted in Figure 5.112 below. Here, we can see that verbs are acoustically distinct from non-verbs, in terms of both pitch and intensity. In the graph on the left, the three verbs are characterized by a higher pitch and higher intensity on σ_1 , the stressed syllable. Nouns and adjectives show the same distribution as in the isolation forms.

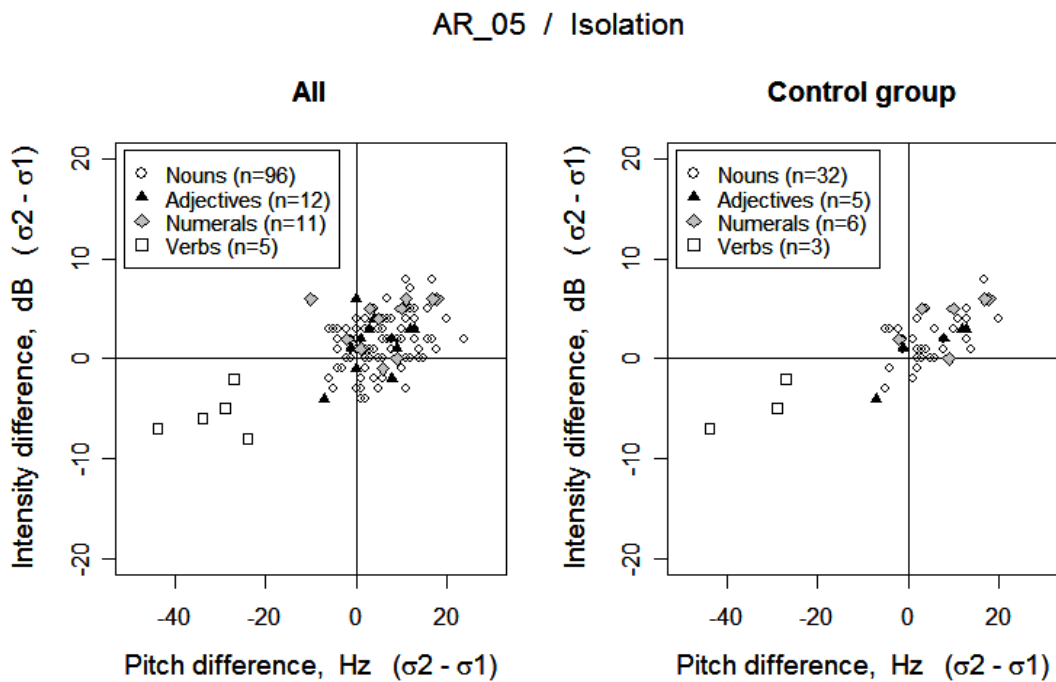
Interestingly, adjectives seem to exhibit the pitch contrasts typical of nouns, and the intensity contrasts typical of verbs.

Figure 5.112 AR_04 / Frame: Pitch difference vs. intensity difference



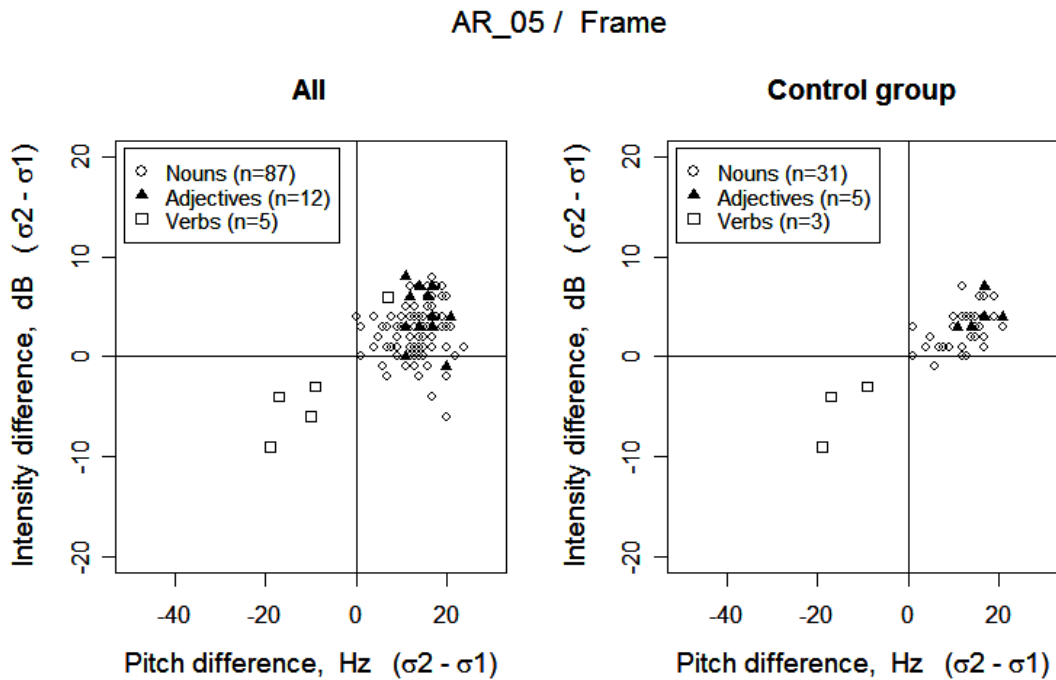
For the isolation forms produced by speaker AR_05, verbs are clearly distinct from nouns, adjectives, and numerals, as shown in Figure 5.113 below. For the verbs, once again, both pitch and intensity correspond with the σ_1 stress pattern. For the non-verbs, pitch is higher on σ_2 in most – but not all – cases; as demonstrated in section 5.3, when pitch is not a reliable cue for σ_2 stress, pitch slope is. In the control group, shown in the graph at right, most of the non-verbs have a higher intensity on σ_2 .

Figure 5.113 AR_05 / Isolation: Pitch difference vs. intensity difference



The AR_05 frame forms once again confirm the acoustic distinction between verbs and non-verbs. In the graph on the left in Figure 5.114 below, the one verb which plots among the nouns is the anomalous [$n\lambda m. bau$] *gnams.babs* ‘to rain’ (rain+fall), discussed previously. In all but one of the adjectives, intensity is higher on σ_2 , the stressed syllable. For nouns, it is only in the control group that most forms have a higher intensity on σ_2 .

Figure 5.114 AR_05 / Frame: Pitch difference vs. intensity difference



What the graphs here demonstrate is that – except in the AR_04 isolation forms – pitch and intensity together clearly differentiate between verbs and non-verbs. Anomalous tokens excluded, it seems that a listener would be very unlikely to mistake a verb for a non-verb: even the points from each group which have the least separation between them on the graph occupy distinct acoustic spaces.

5.5 Vowel duration

As discussed in section 2.3.2.1, there are a number of factors which may incidentally cause vowels to be lengthened, thus obscuring or overriding a potential correlation between vowel duration and stress. Specifically: (a) vowels are often lengthened in utterance-final position; (b) vowels are often longer in open syllables

than in closed syllables⁴⁶; and (c) low vowels are often longer than high vowels. In order to control for (a), isolation forms of non-verbs must be excluded from analysis since, in these cases, the stressed syllable is in utterance-final position. Factors (b) and (c) can be controlled for by focusing on the subset of forms in which both syllables have the same vowel in both syllables, and have the same type of syllable closure (i.e., [open.open] or [closed.closed]).

Analyses of vowel duration measurements for speakers AR_04 and AR_05 are presented in sections 5.5.1 and 5.5.2 below. They provide no clear evidence of a meaningful correlation between vowel duration and stress, for words of any lexical category produced by either speaker.

5.5.1 Vowel duration for speaker AR_04

Analysis of vowel duration data for speaker AR_04 is not very fruitful. As noted above, the isolation forms of non-verbs cannot be seriously considered at all, since it is not possible to determine whether a longer vowel in σ_2 is attributable to stress or simply to the syllable's utterance-final position. For the frame forms of non-verbs, once incidental factors have been controlled for, the samples which remain are generally too small to yield a definitive conclusion.

The isolation forms of verbs must also be excluded from consideration, since – as noted previously – the speaker was not comfortable producing verbs without

⁴⁶ Vowels are also usually longer when they precede a voiced coda than when they precede a voiceless coda, but this factor is never relevant in this analysis.

some larger context of tense / aspect / evidentiality. The sample of frame forms of verbs is again too small to be helpful.

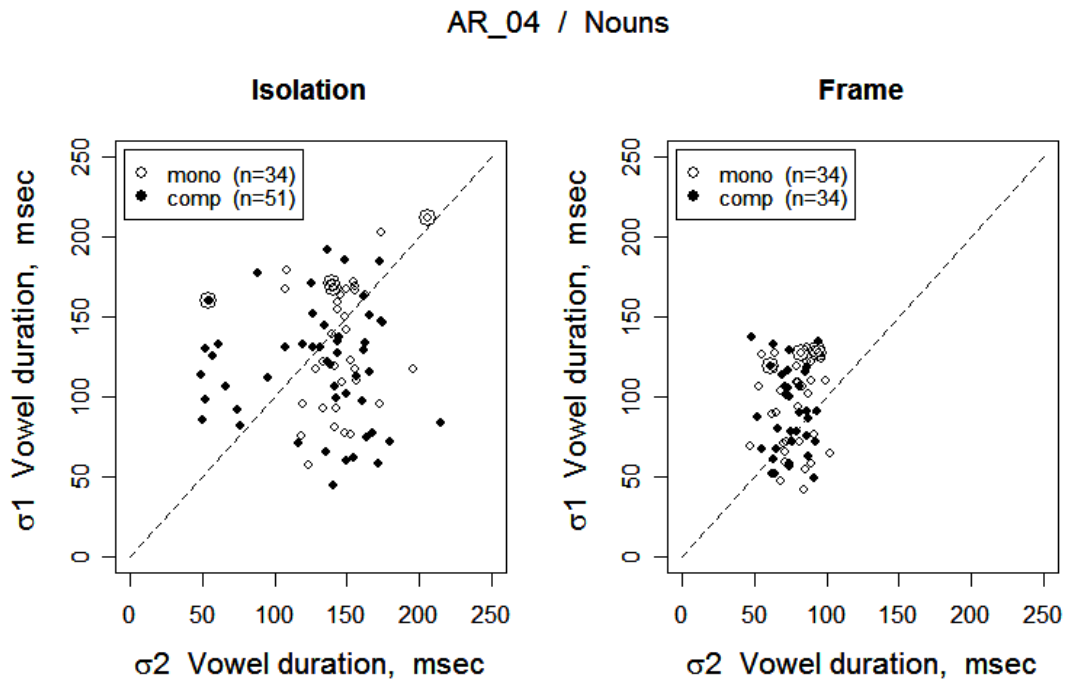
Vowel duration data for nouns, adjectives, and verbs produced by speaker AR_04 is presented briefly in sections 5.5.1.1 through 5.5.1.3 below.

5.5.1.1 AR_04 Nouns

If vowel duration were a correlate of σ_2 stress in Rebkong Amdo nouns, then we would expect vowels in σ_2 to be longer than vowels in σ_1 . But this is not what we observe.

Vowel duration measurements for AR_04 nouns are plotted in Figure 5.115 below. I consider these to be “gross” distribution patterns, because incidental variation due to vowel quality, syllable closure type, and position in the utterance has not been factored out at this point. Nonetheless, these plots provide a general overview and a useful starting point. (Forms with diphthongs – which occur only in σ_1 – have been excluded here; compensatory lengthening – which also occurs only in σ_1 – is indicated by a circle around a data point.)

Figure 5.115 AR_04 / Nouns: Gross vowel duration (circles indicate compensatory lengthening on σ_1 ; forms with diphthongs excluded)



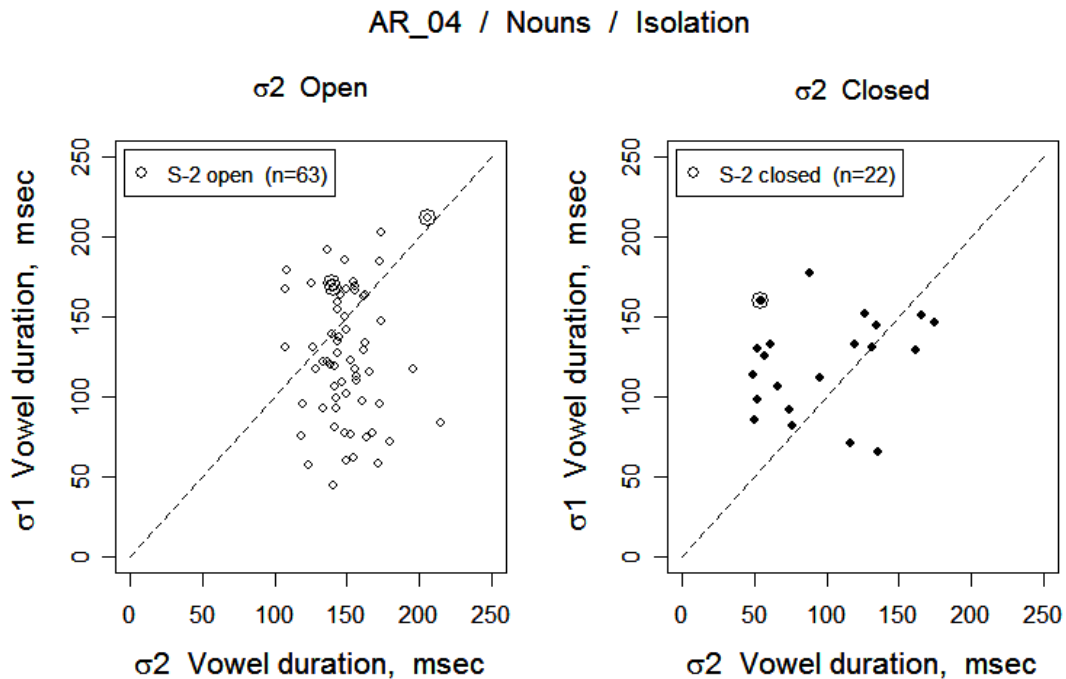
As shown in the graph on the right above, the frame forms – embedded in the carrier sentence ‘In our language, we *X* say’ [ʔa.tʃu ʔe.ki *X* ʔe.ra] – define a fairly tight cluster, a stubby vertical band with greater variation in σ_1 than in σ_2 : duration ranges from ~40 msec to ~140 msec in σ_1 , and from ~50 msec to ~100 msec in σ_2 . Among the isolation forms – shown in the graph on the left – there are some compound nouns which plot within this same range, but otherwise there is considerably more variation and the σ_2 vowels are much longer.

For the isolation forms, it is not possible to distinguish any potential lengthening associated with stress from any potential lengthening associated with σ_2 ’s utterance-final position. This effect is compounded by the fact that σ_2 is so often

open (with the vowel's end point determined by the decline in intensity in the Praat script, as described in section 2.2.2), since monomorphemic nouns generally end in a variant of [-pa] or [-ma]. As noted previously, vowels in open syllables are often longer than vowels in closed syllables. Indeed, both of these factors affect vowel duration here.

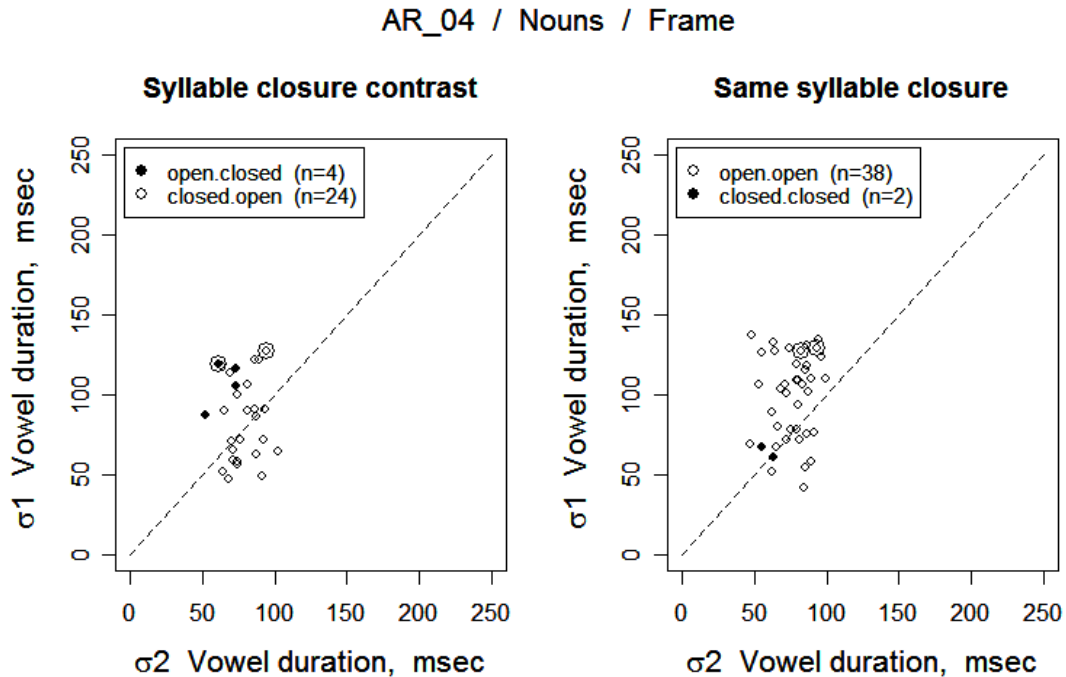
To illustrate, the isolation forms are plotted in terms of whether σ_2 is open or closed in Figure 5.116 below. As shown in the graph on the left, when σ_2 is open, its vowel is usually longer than ~120 msec. We cannot know whether this reflects a correlation with stress or simply the absence of conditions which might pose a limit on the vowel's continuation. The graph on the right shows that even some nouns with σ_2 closed have a vowel longer than 120 msec; these may be cases where the vowel is lengthened by virtue of being in utterance-final position, even though the syllable is closed.

Figure 5.116 AR_04 / Nouns / Isolation: Vowel duration vs. $\sigma 2$ closure (circles indicate compensatory lengthening on $\sigma 1$; forms with diphthongs excluded)



AR_04 frame forms exhibit greater cohesiveness and consistency than the isolation forms, as illustrated in Figure 5.117 below. Nouns with different syllable templates, whether they are monomorphemic – e.g., [na.'ma] *mna* 'ma' 'bride'; [ʔxkar.'ma] *skar.ma* 'star' – or compound – e.g., [xta.'ra] *rta.ra* 'stable, horse pen'; [dɔk.'s^ha] '*brog.sa* 'nomad area'; [xtsa.'t^hʊŋ] *rtswa.thang* 'pasture, grazing area'; [h^əməɣ.'tɰix] *dmag.khrug* 'war' – all plot within the same area in the two graphs below.

Figure 5.117 AR_04 / Nouns / Frame: Vowel duration vs. syllable closure contrast (circles indicate compensatory lengthening on σ_1)



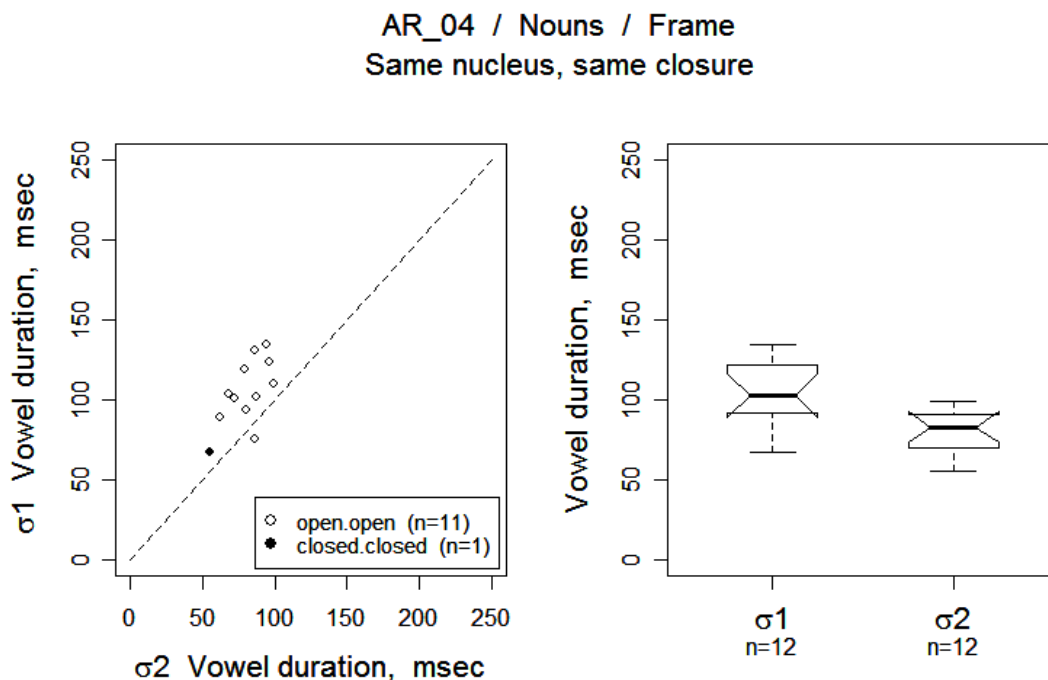
This cohesiveness is not surprising: within the carrier sentence, the target word – whether it has an open σ_2 or a closed σ_2 – is always followed by ['se.'ra] ‘say’. The duration of the σ_2 vowel is thus constrained in the sentence frame by three factors: (a) even in open final syllables, the vowel cannot be casually prolonged since the speaker must move on to complete the utterance; (b) whether open or closed, σ_2 of the target word is not in utterance final position, so its vowel is not subject to the lengthening that one might predict on that basis; and (c) after repeating disyllabic words within the context of the same carrier sentence over and over again, the speaker may have developed a sort of rhythm in producing this fixed number of syllables, further reducing variation.

When there is a contrast in syllable closure – as in the graph on the left in Figure 5.117 above – the handful of [open.closed] nouns all have a longer vowel in σ_1 , as one might predict, and so plot to the left of the dashed line. The majority of the [closed.open] nouns fall to the right of the dashed line, with a longer vowel in σ_2 .

Nouns with the same type of syllable closure – i.e., the [open.open] and [closed.closed] subsets shown in the graph on the right in Figure 5.117 above – mostly plot to the left of the dashed line, with a longer vowel in σ_1 . What dominates the pattern here is that σ_1 vowels show greater variation than do σ_2 vowels.

Finally, to determine whether there is a correlation between stress and vowel duration, this subset of [same.closure] nouns can be further limited to include only those words which have the same vowel in both syllables, thus controlling for the intrinsic variation of vowel duration which is associated with vowel quality (Lehiste 1970). Nouns from this control group are plotted in Figure 5.118 below. The one [closed.closed] noun in this subset is [gʁn.'dʁm] *rgun.'brum* ‘grape’; the set of [open.open] nouns includes [kʒa.'pa] *klad.pa/glad.pa* ‘brain’ and [rʁ.'lʁ] *ril.bu* ‘round pill’. If there were a correlation between vowel duration and stress, then we would expect the nouns in this control group, in particular, to fall to the right of the dashed line. However, as shown, nearly all of them fall instead to the left of the dashed line, indicating a longer vowel in σ_1 , contradicting the perceived stress pattern.

Figure 5.118 AR_04 / Nouns / Frame / Control group: Vowel duration contrasts (circles indicate compensatory lengthening on σ_1)



A paired-sample t-test, summarized in Table 5.28 below, confirms that the contrast in vowel duration across syllables for this control group is statistically significant, with $p < 0.05$ and the vowel in σ_1 , on average, 24 msec longer than the vowel in σ_2 .

Table 5.28 AR_04 / Nouns / Frame / Vowel duration / Control group: Results of paired-sample t-tests (two-tailed)

Speaker	Setting	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
						lower	upper
AR_04	frame	-5.1035	11	0.0003422	-24	-34	-14

In conclusion, then, isolation forms of AR_04 nouns cannot be evaluated for a correlation between vowel duration and stress, because incidental factors cannot be

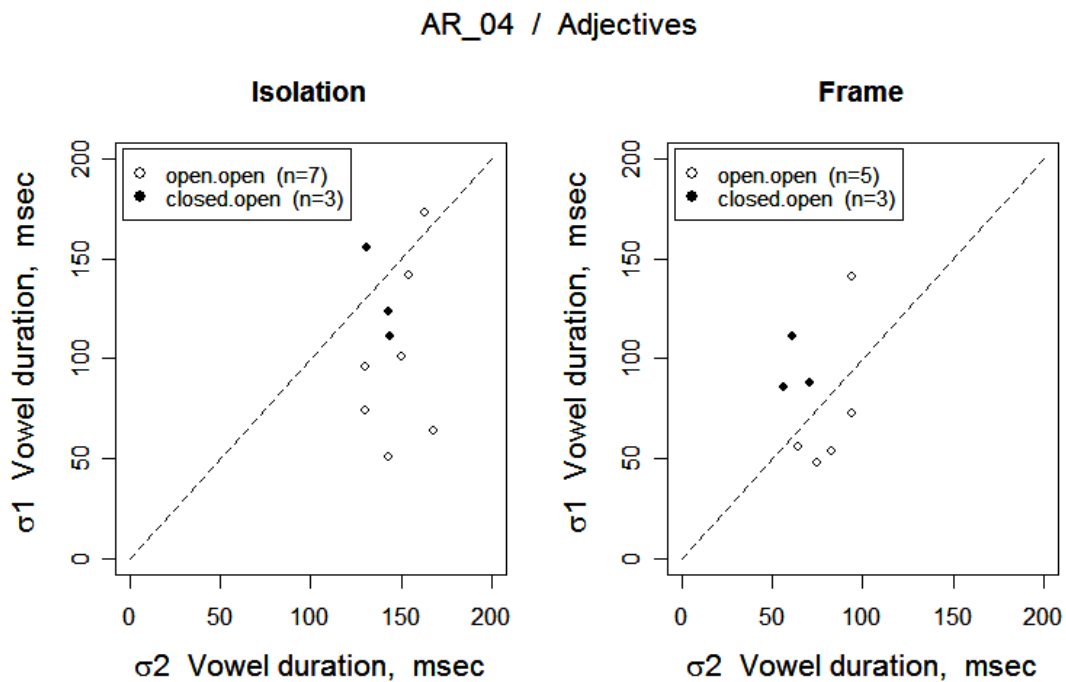
controlled for. The control group of frame forms indicates that the vowel in σ_1 is longer than the vowel in σ_2 , in opposition to the stress pattern.

While a native speaker of Rebkong Amdo may not notice that the vowel in σ_1 is sometimes longer than that in σ_2 , contradicting the stress pattern, this is something that a native speaker of English might subconsciously notice – and it might even present a source of confusion regarding the locus of stress, since vowel duration plays a role in conveying stress in English.

5.5.1.2 AR_04 Adjectives

As was the case with nouns, AR_04 adjectives produced in isolation are not useful in assessing the potential correlation between vowel duration and stress. As illustrated in Figure 5.119 below, the vowels in σ_2 are somewhat longer in the isolation forms than in the frame forms, but it is not possible to determine whether this lengthening is associated with stress, or with σ_2 's structure or utterance-final position.

Figure 5.119 AR_04 / Adjectives: Gross vowel duration vs. syllable closure



As it happens, the frame forms are not useful in assessing vowel duration either. All of the adjectives here are monomorphemic, and so have an open σ_2 . We might expect that adjectives with σ_1 closed would have a relatively longer vowel in σ_2 , but this is not the case; as illustrated in the plot on the right above, the three [closed.open] adjectives – [zar.'mo] *gzar.po* ‘steep’, [xkam.'bo] *skam.po* ‘dry, dried’, and [ts^han.'di] *tsha.'di* ‘hot’ – fall instead to the left of the dashed line. This may be due to the contrast in vowel height across syllables: lower vowels are intrinsically longer (Lehiste 1970). None of the [open.open] adjectives have [a] in σ_1 : e.g., [s^he.'ru] *ser.po* ‘yellow’ and [s^ho.'ma] *so.ma* ‘new’.

The sample of frame forms is simply too small to include a subset which has the same vowel in both syllables (and the same syllable closure). Without such a control group, it is not possible to assess the potential correlation between vowel duration and stress.

All that can be concluded here is that vowel duration is more constrained in the context of the sentence frame than when produced in isolation.

5.5.1.3 AR_04 Verbs

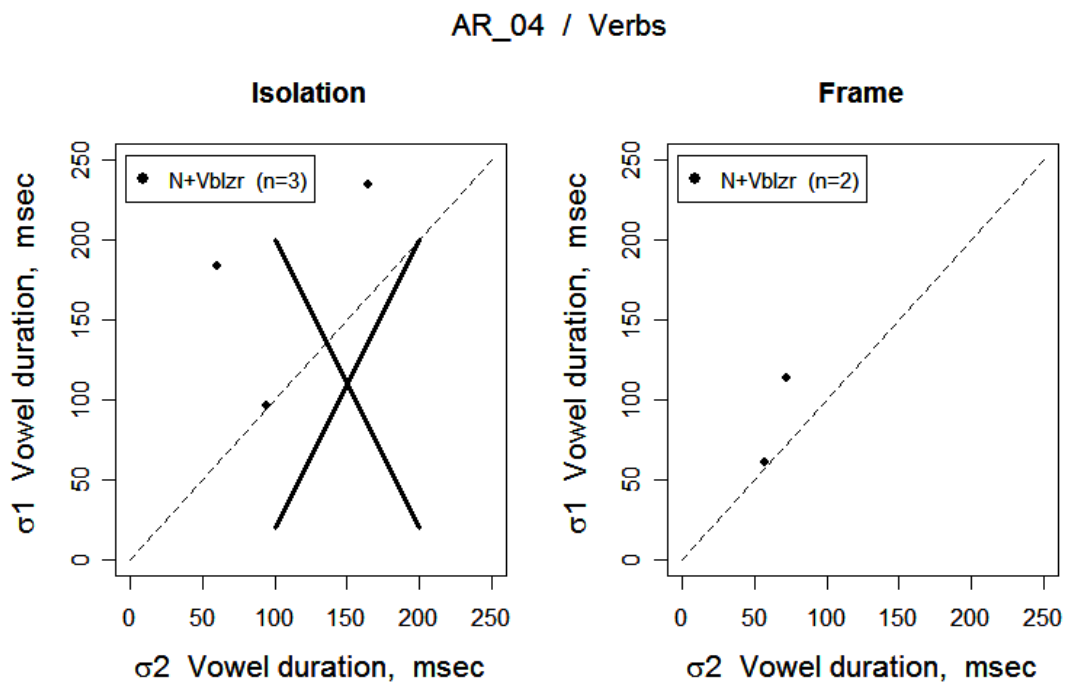
In contrast to nouns, adjectives, and numerals, verbs are stressed on σ_1 . This means that the stressed syllable can never be in utterance-final position – even when produced in isolation. Thus, in theory, we can consider whether vowels are longer in the stressed syllable of isolation forms without worrying about the incidental effects of utterance-final lengthening.

In practice, however, in the case of speaker AR_04, the isolation forms cannot be considered anyway: as discussed previously, this speaker was not comfortable producing a verb with no context, resulting in odd acoustic contrasts. Thus meaningful information can only be garnered from the frame forms.

If vowel duration were a correlate of stress in Rebkong Amdo verbs, we would expect to see longer vowels in σ_1 , particularly when both syllables have the same vowel and the same type of syllable closure. Unfortunately, the small sample of verbs recorded did not include any such control cases, and with so few measurements, it is impossible to draw any firm conclusions about a possible relationship between stress and duration.

Vowel duration measurements for the two frame forms recorded from speaker AR_04 are plotted in Figure 5.120 below.

Figure 5.120 AR_04 / Verbs: Gross vowel duration contrasts



Vowel duration differences and ratios are shown for these verbs in Table 4.32 above. When both syllables are closed with voiced codas, as in (a), there is almost no contrast in duration. For (b) the vowel is much longer in $\sigma 1$ than in $\sigma 2$, just as one would predict based on the [open.closed] syllable template.

Table 5.29 AR_04 / Verbs / Frame / [closed.open] / Same nucleus: Vowel duration ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_2/σ_1)
to rain	'nɿm.biʊ	<i>snam.bab</i>	-4	0.63
to give an answer	'ʒẽ.tɕɿʊ	<i>lan.brgyab</i>	-42	0.93

Without a larger sample – which can be controlled for the effects of incidental factors – nothing further can be said about vowel duration in AR_04 verbs.

5.5.2 Vowel duration for speaker AR_05

As was the case with speaker AR_04, analysis of vowel duration data for speaker AR_05 is only minimally informative. Again, the isolation forms of non-verbs must be excluded from consideration, since it is not possible to control for the potential lengthening effect of σ_2 's utterance-final position. And for the frame forms of non-verbs, once incidental factors have been controlled for, the subsets which remain do not show evidence of a meaningful correlation with stress.

For the verbs, too, the samples of both isolation and frame forms are too small to yield any conclusive findings.

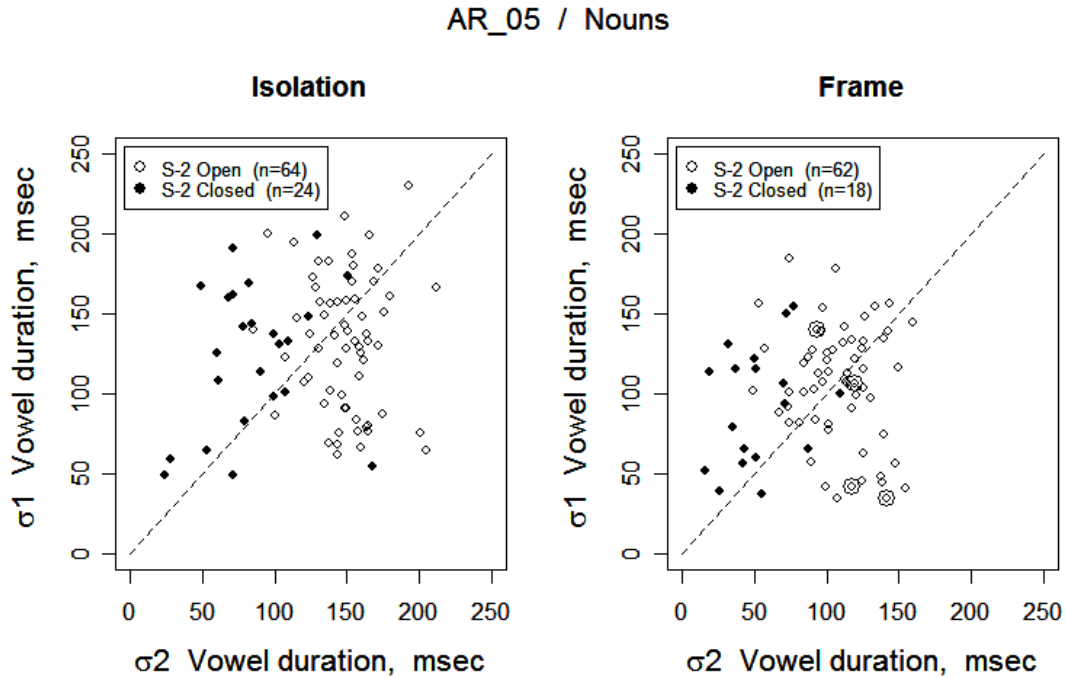
Vowel duration data for nouns, adjectives, numerals, and verbs produced by speaker AR_05 is presented briefly in sections 5.5.2.1 through 5.5.2.4 below.

5.5.2.1 AR_05 Nouns

Gross vowel duration measurements for nouns produced by speaker AR_05 are plotted in Figure 5.121 below. For both isolation and frame forms, the effect of σ_2

closure is readily apparent: the vowel in σ_2 is longer when σ_2 is open than when it is closed. That is, in both plots, the open circles fall to the right of the filled circles.

Figure 5.121 AR_05 / Nouns: Gross vowel duration contrasts (circles indicate compensatory lengthening on σ_1)



We can also see that the vowel in σ_2 is longer when this syllable is in utterance-final position than when it is not: the entire cluster of points representing the isolation forms is slightly further to the right than the cluster of points representing the frame forms. Since it is impossible to separate out this effect of utterance-final lengthening, there is nothing to be learned from further consideration of the isolation forms.

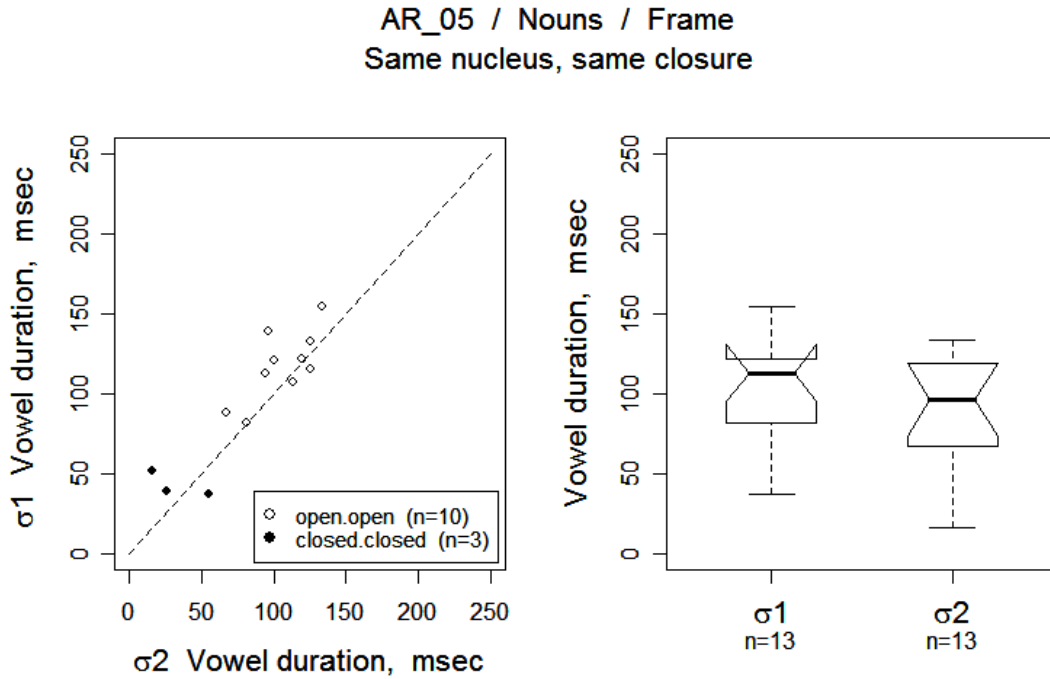
In the frame forms, the syllable of interest – σ_2 – is never in utterance-final position, since it is followed by the remainder of the carrier sentence [ʔa.tʃu ʔkɛ.ki X

'se.'ra] 'In our language, we *X* say'. Thus the potential effect of utterance-final lengthening is not a concern.

The potential effects of syllable closure contrast and vowel quality contrast on duration can also be eliminated, by considering a subset in which both syllables have the same closure, and also have the same vowel. Vowel duration measurements for the thirteen nouns in this control group are plotted in Figure 5.122 below. If vowel duration served as an acoustic correlate of stress, then we would expect these nouns to have a longer vowel in σ_2 , and plot to the right of the dashed line. Instead, most nouns fall to the left of the line, and others plot fairly close to it.

As one would expect, nouns in which both syllables are closed – e.g., [xkʌŋ.'lʌm] *rkang.lam* 'footpath, trail'; [tʃʌʔ.'tʰʌʔ] *lcags.thag* 'iron chain' – have shorter vowels in both syllables and plot closer to the graph's origin, while nouns in which both syllables are open – e.g., [do.'χo] *bdag.po* 'boss, head man'; [ta.'ra] *rta.ra* 'stable, horse pen' – have longer vowels in both syllables and plot further from the graph's origin.

Figure 5.122 AR_05 / Nouns / Frame / Control group: Vowel duration contrasts



A paired-sample t-test for this control group, summarized in Table 5.30 below, indicates that the difference in vowel duration across syllables is statistically significant, with $p < 0.05$ and a mean difference of -12 msec (i.e., longer in σ_1 , the syllable which is *not* stressed). This mean difference is barely greater than the just-noticeable difference, and the upper 95% confidence limit is only 0.77 msec. That is, in repeated sampling of nouns produced in the sentence frame, a mean duration difference of < 1 msec would be statistically significant; such a contrast would not be perceptible, much less perceptually significant.

Table 5.30 AR_05 / Nouns / Frame / Vowel duration / Control group: Results of paired-sample t-tests (two-tailed)

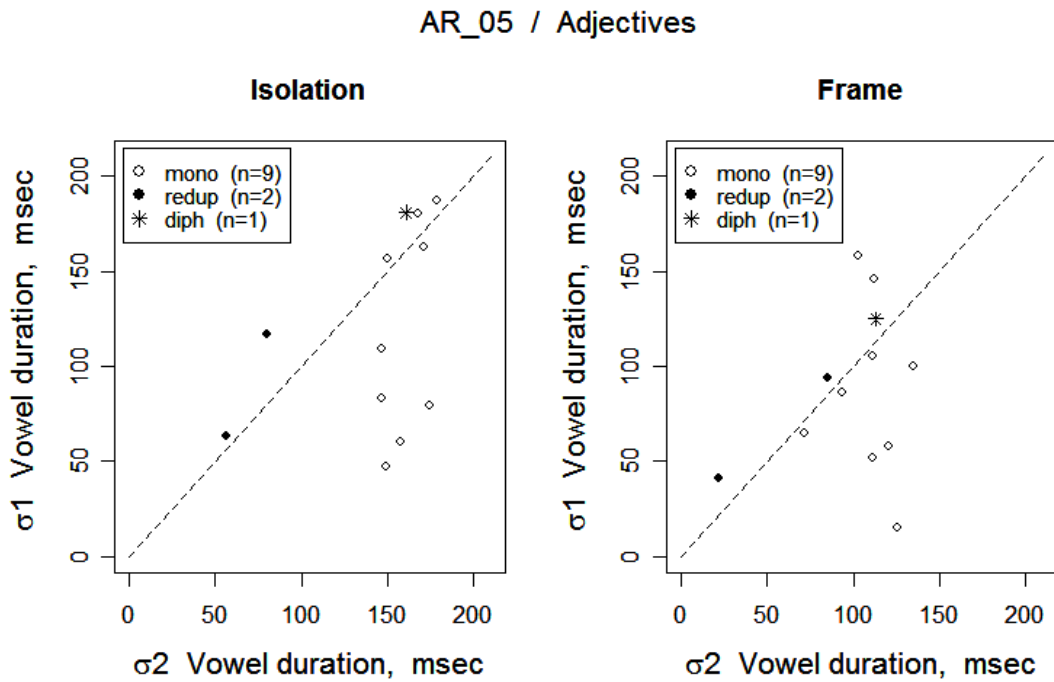
Speaker	Setting	t	DF	p-value	Mean diff (msec)	95 % conf. limits	
						lower	upper
AR_05	frame	-2.3356	12	0.03768	-12	-22	-0.77

In conclusion, when AR_05 nouns are controlled for the potential effects of position in the utterance, syllable closure type, and vowel quality, there is no evidence to suggest a meaningful correlation between vowel duration and stress. On the contrary, there is a limited and weak tendency for the vowel to be longer in σ_1 , the syllable which is *not* stressed.

5.5.2.2 AR_05 Adjectives

Among the AR_05 adjectives, there are only two words that are controlled for both vowel quality and syllable closure: these are the two reduplicated forms [lev.'lev] *leb.leb* ‘flat’ and [tʰʊŋ.'tʰʊŋ] *chung.chung* ‘small’. As shown in Figure 5.123 below, in isolation and in the sentence frame these two adjectives have a longer vowel in σ_1 – the syllable which is *not* stressed – and so plot to the left of the dashed line.

Figure 5.123 AR_05 / Adjectives: Gross vowel duration contrasts



None of the other adjectives are helpful in determining whether or not vowel duration is a correlate of σ_2 stress. In the isolation forms, lengthening may occur in σ_2 simply because it is in utterance-final position; this cannot be distinguished from any potential lengthening which is related to stress. For both isolation and frame forms, vowel duration in σ_2 may also be affected by a contrast in syllable closure (vowels in open syllables tend to be longer than those in closed syllables) or contrasts in vowel quality (lower vowels tend to be longer than higher vowels).

The only samples controlled for all of these incidental factors are the reduplicated forms in the sentence frame. Both of these plot close to but just left of the dashed line; the difference in vowel duration across syllables is -9 msec in one

case, and -19 msec in the other case. This certainly does not suggest a correlation between vowel duration and σ_2 stress.

The only pattern that is clear here is that, for the monomorphemic adjectives – like [kɔim.'bo] *dkon.po* ‘rare’ and [xtsʌŋ.'ma] *gtsang.ma* ‘clean’ – there is less variation in the duration of vowels in σ_2 (which all have similar structure and content) than in the duration of vowels σ_1 . There is no evidence of a correspondence between vowel duration and stress.

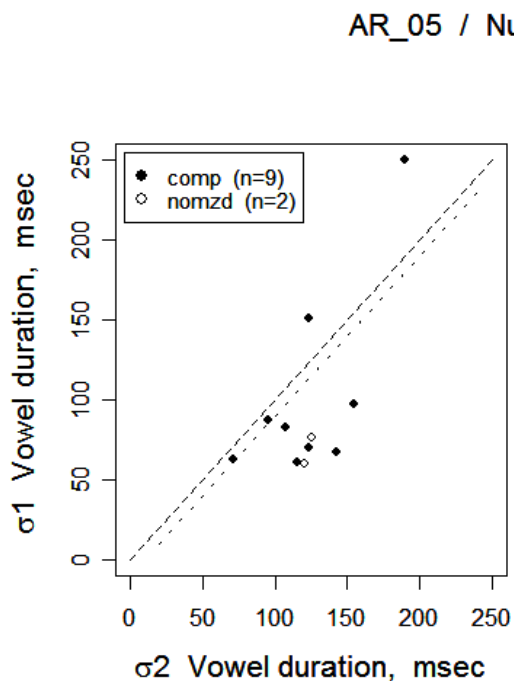
5.5.2.3 AR_05 Numerals

Because the numerals were only recorded in isolation from speaker AR_05, it is not possible to determine whether or not there is a correlation between vowel duration and σ_2 stress. There is no way to control for the potential intrinsic lengthening of a vowel in the final syllable of an utterance. Vowel quality contrasts and syllable closure contrasts may also have an intrinsic effect.

Duration measurements for all numerals which could be reliably segmented and analyzed are illustrated in Figure 5.124 below. The one token which is controlled for both vowel quality and syllable closure – the compound form [tɕʌb.'dʌŋ] *bcu.bdun* ‘seventeen’ – has a longer vowel in σ_2 , plotting at coordinates (71, 63), between the dashed “0” line and the dotted “+10 msec” reference line. All of the other numerals have a contrast across syllables in either vowel quality, or syllable closure, or both. Thus even if the numerals had been recorded in the sentence frame, it would

still not be possible to control for all potential incidental factors and focus on identifying a correlation with stress.

Figure 5.124 AR_05 / Numerals / Isolation: Gross vowel duration contrasts



5.5.2.4 AR_05 Verbs

If vowel duration were a correlate of stress for Rebkong Amdo verbs, then we would expect vowels to be longer in $\sigma 1$ – since this is the locus of stress for verbs – than in $\sigma 2$.

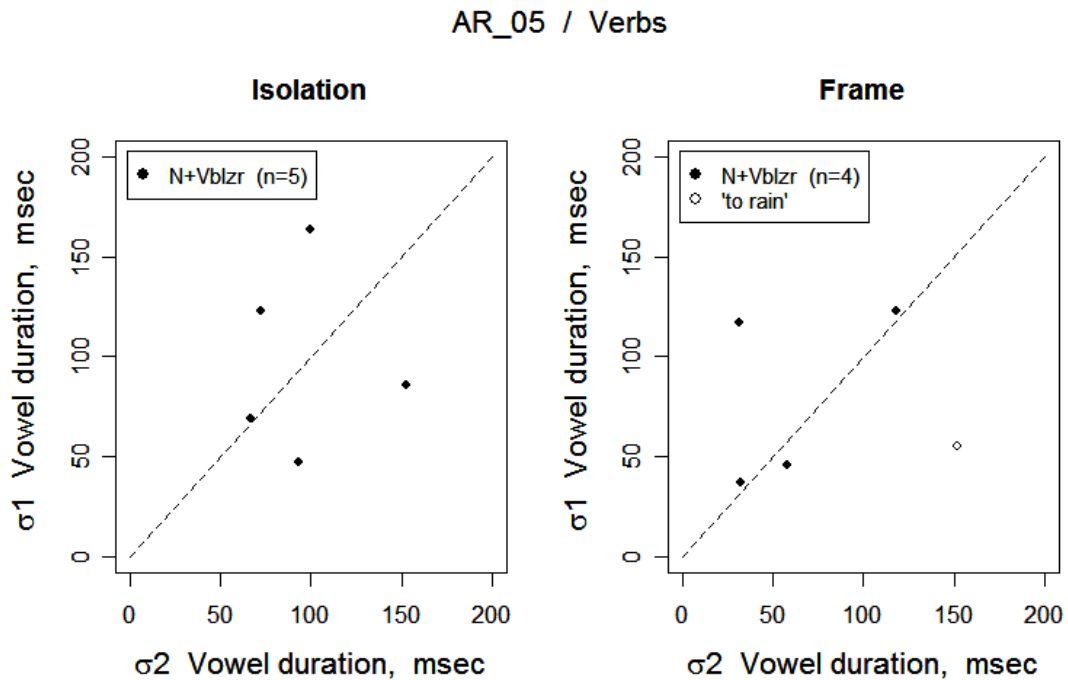
Tautologically, $\sigma 1$ can never be the last syllable of a disyllabic word. This means that the isolation forms of verbs can provide meaningful data – which was not the case with the non-verbs – since here there is no danger that vowel lengthening

associated with σ_1 stress could be confused with vowel lengthening associated with utterance-final position.

Like the non-verbs, though, the sample of verbs available is unfortunately too small to permit a meaningful assessment of a potential correlation between vowel duration and stress.

Vowel duration measurements for AR_05 verbs produced in isolation and in the sentence frame are plotted in Figure 5.125 below. For the isolation forms, points fall to either side of the dashed line; for the frame forms, most points fall close to or left of the line, with the exception of 'to rain'. (As discussed previously, the frame form of [nɒm.bau] *gnam.babs* 'to rain' was produced with an anomalous stress pattern.)

Figure 5.125 AR_05 / Verbs: Gross vowel duration contrasts



Vowel duration differences and ratios for the five isolation forms are shown in Table 5.31 below. The two weather verbs [$k^h\Lambda\eta.b\Lambda p$] *gangs.babs* ‘to snow’ and [$h^{\text{ho}}n\Lambda m.b\Lambda p$] *gnam.babs* ‘to rain’ are controlled in terms of syllable closure and vowel quality, and both have a longer vowel in σ_2 , the *unstressed* syllable. But this may be idiosyncratic, and no other patterns are evident. The sample is really too small to yield conclusive findings.

Table 5.31 AR_05 / Verbs / Isolation: Vowel duration differences and ratios

Gloss	IPA	WT	Dur diff (msec)	V / V: ratio (σ_2/σ_1)
to give an answer	'lɛn.tɤv	<i>lan.brgyab</i>	-51	0.59
to shoot an arrow	'nda.hɛn	<i>mda'.phen</i>	-64	0.61
to shoot a gun	'vu.dzəp	<i>wu.brgyab</i>	-2	0.97
to rain	$h^{\text{ho}}n\Lambda m.b\Lambda p$	<i>gnam.babs</i>	67	1.78
to snow	$k^h\Lambda\eta.b\Lambda p$	<i>gangs.babs</i>	46	1.98
Mean			-0.80	1.18

Duration contrasts for the frame forms seem to be just as arbitrary, given the measurements summarized in Table 5.32 below. The sample is simply too small to identify a subset of verbs controlled for syllable closure and vowel quality. It would be helpful to consider a larger sample in future.

Table 5.32 AR_05 / Verbs / Frame: Vowel duration differences and ratios

Gloss	IPA	WT	Dur diff (msec)	V / V': ratio (σ_2/σ_1)
to answer	'lɛn.tɤv	<i>lan.brgyab</i>	-86	0.26
to shoot a gun	'ʊu.dʒɪp	<i>wu.brgyab</i>	-5	0.86
to shoot an arrow	'nda.hẽ	<i>mda'.phen</i>	-5	0.96
to snow	'k ^x ʌŋ.bɤv	<i>gangs.babs</i>	12	1.26
Mean			-21	0.84

Note: 'to rain' has been excluded

In conclusion, too few verbs were recorded from speaker AR_05 to begin to determine whether or not there is a correlation between vowel duration and stress.

5.6 Recapitulation

Table 5.33 below summarizes the contributions made by pitch, pitch slope, intensity, and vowel duration to the stress patterns perceived in Balti.

For non-verbs, σ_2 stress is conveyed by the two reflexes of fundamental frequency: pitch and pitch slope. Intensity and vowel duration do not correlate with stress, and, in fact, are sometimes more prominent on σ_1 , contrary to the stress pattern.

For verbs, σ_1 stress is conveyed by both pitch and intensity.

Table 5.33 Acoustic correlates of stress in Rebkong Amdo

	Speaker	Setting	Pitch	Pitch slope	Intensity	Vowel duration
Nouns (σ_2)	AR_04	isol	robust	robust	limited, weak	NA
		frame	robust	robust	limited, weak	no – I
	AR_05	isol	no	robust	limited, weak	NA
		frame	yes	weak	limited	no – I
Adjs (σ_2)	AR_04	isol	robust	robust	no – I	NA
		frame	robust	no	no – I	IS
	AR_05	isol	no	robust	no – SS	NA
		frame	robust	no	yes – SS	no – SS
Nums (σ_2)	AR_05	isol	limited	no	yes	NA
Verbs (σ_1)	AR_04	isol	NA	NA	NA	NA
		frame	possible – SS	NA	possible – SS	IS
	AR_05	isol	robust	NA	robust	no – SS
		frame	yes	NA	yes	no – SS

NA This parameter could not be analyzed.

I In fact, the Inverse relationship is observed: intensity is higher on σ_1 , the syllable which is *not* stressed.

IS Insufficient Sample. Too small to analyze conclusively.

SS Small Sample. The conclusion presented in the table should be considered tentative, as it was based on a very limited data set.

possible – The sample size is small, but provides no evidence which would contradict a potential correlation with stress.

6. Summary, implications, and future research

I begin this chapter, in section 6.1, with a summary of the analyses of stress patterns and stress correlates presented in Chapters 4 and 5. Based on these results, in section 6.2 I then offer a reconstruction of stress in Proto-Tibetan. Section 6.3 provides background information on the tone patterns observed in the Innovative dialects of Tibetan, and a review of previous research on Tibetan tonogenesis. This material lays the groundwork for the hypothesis I present in section 6.4 regarding the role of stress in Tibetan tonogenesis.

In section 6.5 I consider the implications of this analysis in terms of the direction of tone split in Tibetan. Finally, in section 6.6 I briefly mention several topics that will be interesting to investigate in the future.

6.1 Stress in Balti and Rebkong Amdo Tibetan

The core contributions of this study are descriptive and analytical: I identify the stress patterns and the acoustic correlates of stress in Balti Tibetan and Rebkong Amdo Tibetan, as summarized in sections 6.1.1 and 6.1.2 below.

6.1.1 Stress patterns

Disyllabic words in Balti and Rebkong Amdo exhibit the same stress patterns, distinguishing non-verbs from verbs. Non-verbs – nouns, adjectives, and numerals – are stressed on σ_2 . Verbs are stressed on σ_1 .

As noted in section 3.1.1, it is fairly unusual for a language to exhibit distinct stress patterns in different lexical categories. For Zhongu Tibetan, (Sun 2003) reports the same stress patterns and stress contrasts as are described here, and English has noun-verb minimal pairs like 'rebel - re'bel, 'contrast - con'trast, and 'insight - in'cite. This also occurs in Dumi Rai, a Kiranti language spoken in Nepal; as described by van Driem (1993: 58): “verbs and deverbatives are always stressed on the root. Nouns and other parts of speech are stressed on the first syllable.” Other languages exhibiting this phenomenon could probably be identified.

6.1.2 Stress correlates

In Chapters 5 and 6 I presented my analyses of the acoustic correlates of stress in Balti and Rebkong Amdo. My focus was on identifying *meaningful* acoustic correlates – i.e., contrasts across syllables which are statistically, perceptually, and contextually significant, as described in Chapter 2. Results are summarized in sections 6.1.2.1 and 6.1.2.2 below, for non-verbs and verbs respectively.

6.1.2.1 Stress correlates in non-verbs (σ_2 stress)

For nouns and adjectives, the fundamental frequency-related parameters pitch and pitch slope are the most consistent and robust cues for σ_2 stress. Intensity plays a more limited role. For numerals, the Balti speaker uses only pitch to convey stress, and the Rebkong Amdo speaker uses only intensity to convey stress.

Table 5.33 below provides a more detailed summary of the behavior of acoustic parameters in nouns. For Balti, pitch is the only robust and consistent

acoustic correlate of σ_2 stress. Intensity is a weak correlate of stress only in a limited subset of the nouns produced by speaker BM_01. For Rebkong Amdo, both pitch and pitch slope serve to convey stress in nouns – sometimes reinforcing one another, and sometimes complementing one another, as discussed in section 4.3.3. Intensity shows only a limited – and usually weak – correlation with stress.

Table 6.1 Meaningful acoustic correlates of stress for nouns

Dialect	Speaker	Setting	<i>n</i>	Pitch	Pitch slope	Intensity	Vowel duration
Balti	BSh_03	isol	77	robust	no	no	NA
		frame	64	robust	no	no	no – SS
	BM_01	isol	72	yes	no	limited, weak	NA
Amdo	AR_04	isol	93	robust	robust	limited, weak	NA
		frame	74	robust	robust	limited, weak	no – I
	AR_05	isol	96	no	robust	limited, weak	NA
		frame	87	yes	weak	limited	no – I
	Shared correlates				F0 – yes	limited	no

NA This parameter could not be analyzed.

I In fact, the Inverse correlation is observed.

SS Small Sample. The conclusion presented in the table should be considered tentative, as it was based on a very limited data set.

In the bottom row of the table above, I identify fundamental frequency as the only acoustic correlate of stress which is common to both dialects. As discussed in section 2.2.2, for purposes of this study, I identify “pitch” (average F0 across the medial 50% of a vowel) and “pitch slope” (the change in F0 over the span of the vowel) as acoustic correlates of stress. These are both measurements of F0, and can be thought of as distinct reflexes of F0. That is there are two ways in which speakers manipulate fundamental frequency in order to convey stress. Balti speakers only

express F0 as “pitch”. Rebkong Amdo speakers sometimes manipulate F0 to create contrasts in “pitch”, and sometimes manipulate F0 to create contrasts in “pitch slope”.

Table 6.2 below provides a more detailed account of the behavior of acoustic parameters in adjectives. For Balti, pitch is the only cue for σ_2 stress. For Rebkong Amdo, both pitch and pitch slope serve to convey stress – sometimes reinforcing one another, and sometimes complementing one another, as was the case for nouns. For speaker AR_04, intensity shows an inverse correlation with stress: intensity is higher on σ_1 , the syllable which is *not* stressed. Again, the fundamental frequency-related parameters are the only shared correlates.

Table 6.2 Meaningful acoustic correlates of stress for adjectives

Dialect	Speaker	Setting	<i>n</i>	Pitch	Pitch slope	Intensity	Vowel duration
Balti	BSh_03	isol	8	robust	no	no – SS	NA
		frame	7	robust	no	no – SS	no – SS
	BM_01	isol	26	yes	no	no	NA
Amdo	AR_04	isol	12	robust	robust	no – I	NA
		frame	10	robust	no	no – I	IS
	AR_05	isol	12	no	robust	no – SS	NA
		frame	12	robust	no	yes – SS	no – SS
Shared correlates				F0 – yes		no	no

NA This parameter could not be analyzed.

I In fact, the Inverse correlation is observed.

IS Insufficient Sample. Too small to analyze conclusively.

SS Small Sample. The conclusion presented in the table should be considered tentative, as it was based on a very limited data set.

The numerals recorded for this study exhibit a high degree of acoustic uniformity. Unlike the nouns and adjectives, they constitute a unique and restricted

semantic set. Furthermore, they were recorded in sequence within a very short span of time. (Also, they were recorded only in isolation, which means that vowel duration could not be assessed.) As shown in Table 6.3 below, speaker BSh_03 used only pitch to convey σ_2 stress, while speaker AR_05 used only intensity. There are no stress correlates common to the two dialects.

Table 6.3 Meaningful acoustic correlates of stress for numerals

Dialect	Speaker	Setting	<i>n</i>	Pitch	Pitch slope	Intensity	Vowel duration
Balti	BSh_03	isol	9	robust	no	no	NA
Amdo	AR_05	isol	11	limited	no	yes	NA
Shared correlates					?	?	?

NA This parameter could not be analyzed.

6.1.2.2 Stress correlates in verbs (σ_1 stress)

For verbs, both pitch and intensity are consistent acoustic correlates of σ_1 stress for both speakers in both dialects, as summarized in Table 6.4 below. Pitch slope is never a meaningful acoustic correlate of stress. (It lacks perceptual significance in verbs, as discussed in sections 4.3 and 5.3.) Vowel duration seems to contribute to the perception of σ_1 stress in Balti verbs, but in Rebkong Amdo the sample was too small to permit a conclusive analysis.

Table 6.4 Meaningful acoustic correlates of stress for verbs

Dialect	Speaker	Setting	<i>n</i>	Pitch	Pitch slope	Intensity	Vowel duration
Balti	BSh_03	isol	35	robust	no	robust	limited
		frame	31	robust	no	robust	limited
	BM_01	citation	8	robust	no	robust	yes – SS
		N+Vblzr	6	robust	no	limited	yes – SS
Amdo	AR_04	isol	3	NA	NA	NA	NA
		frame	2	possible – SS	NA	possible – SS	IS
	AR_05	isol	5	robust	NA	robust	no – SS
		frame	5	yes	NA	yes	no – SS
Shared correlates				yes	no	yes	?

NA This parameter could not be analyzed.

IS Insufficient Sample. Too small to analyze conclusively.

SS Small Sample. The conclusion presented in the table should be considered tentative, as it was based on a very limited data set.

6.2 Reconstructing stress in Proto-Tibetan

As discussed in section 1.2, the distribution of the conservative Archaic dialects – like Balti and Rebkong Amdo – at the western and eastern edges of the Tibetan language area is consistent with the history of the Tibetan empire. At its peak in the 8th century, garrisons and settlements were established in these areas. When the empire collapsed late in the 9th century, these Tibetan-speaking communities were left stranded at the periphery, isolated politically and geographically from each other and from the central plateau.

Previous research (discussed in section 1.4) has shown that, by comparing the various Tibetan dialects to each other and to Written Tibetan, Proto-Tibetan can be

reconstructed as lacking contrastive tone but rich in complex consonant clusters.

These characteristics are preserved in the Archaic dialects.

I take a similar approach in reconstructing stress patterns and stress correlates for Proto-Tibetan.

6.2.1 Stress patterns

Since nouns, adjectives, and numerals in both Balti and Rebkong Amdo are stressed on σ_2 , I reconstruct a pattern of σ_2 stress for Proto-Tibetan nouns, adjectives, and numerals. Likewise, since verbs in both Balti and Rebkong Amdo are stressed on σ_1 , I reconstruct a pattern of σ_1 stress for Proto-Tibetan verbs.

6.2.2 Stress correlates

The acoustic correlates of stress in Proto-Tibetan can likewise be reconstructed by comparing the acoustic correlates of stress in Balti and Rebkong Amdo. This is the methodology that I refer to as “historical comparative acoustics”.

The bottom rows of Table 5.33 through Table 6.4 showed the acoustic correlates of stress which are common to both speakers of both dialects. Ruling out borrowing and coincidence as sources of these shared features – as discussed previously in section 1.4 – we can conclude that they were inherited from a common parent. That is, we can reconstruct these features for Proto-Tibetan.

I thus reconstruct fundamental frequency as a robust and consistent acoustic correlate of σ_2 stress in Proto-Tibetan nouns and adjectives. (See Table 5.33 and Table 6.2.) And I reconstruct both pitch and intensity as acoustic correlates of σ_1

stress in Proto-Tibetan verbs. Vowel duration may also play a role here. (See Table 6.4.)

The stress correlates for Proto-Tibetan numerals cannot be reconstructed, since the Balti and Rebkong Amdo speakers did not utilize acoustic resources in the same way for this lexical category. (See Table 6.3.)

6.3 Background on Tibetan tone and tonogenesis

As mentioned briefly in section 1.4, previous researchers have not only reconstructed sounds and words for Proto-Tibetan, but have also developed several hypotheses to explain how tone arose as a contrastive feature in the modern spoken Innovative dialects.

The term “tonogenesis” was coined by Matisoff (1970) to refer to both the development of tonal contrasts in a previously non-tonal language, and to the increase in the number of tonal contrasts in an already-tonal language. In the case of Tibetan, since Proto-Tibetan has been reconstructed as non-tonal and since the modern spoken Innovative dialects are tonal, it is apparent that tonogenesis has occurred.

One of the main theoretical contributions of this dissertation is my consideration of the role that stress has played in this process, a factor which has been almost entirely overlooked in the past.

In section 6.3.1 below I describe the tone patterns which are commonly observed in the tonal Innovative dialects, and which must be accounted for by a theory of Tibetan tonogenesis. In section 6.3.2 I provide an overview of previous

hypotheses of Tibetan tonogenesis, and in section 6.4 I present my own view, which incorporates the role played by stress.

6.3.1 Tone patterns in the Innovative dialects of Tibetan

As discussed in section 1.2, the Innovative dialects are, by definition, those which make contrastive use of tone, and which have simplex syllable onsets and reduced (or no) syllable codas. They are spoken across the vast expanse of the Tibetan plateau, physically separating Archaic dialects of the western and eastern peripheries, such as Balti and Rebkong Amdo.

In sections 6.3.1.1 and 6.3.1.2 below I provide an overview of the tone patterns observed in monosyllabic and disyllabic words, respectively. I begin each section with specific examples from Tokpe Gola Tibetan, an Innovative dialect spoken in the Taplejung district of northeastern Nepal on which I have conducted considerable field research. The tone patterns observed in Tokpe Gola are typical of those reported for other Tibetan dialects, as I demonstrate with a brief summary of previous literature.

6.3.1.1 Monosyllabic words

In Tokpe Gola, monosyllabic words fall into one of two phonological tone categories: high (H) or low (L). High-toned words are historically associated with voiceless unaspirated or voiceless aspirated onset consonants, and modal phonation; in certain phonological environments – i.e., with certain WT onsets and codas – they exhibit a sharp falling tone and a laryngealized onset. Low-toned words are

associated with voiced unaspirated or voiceless aspirated onset consonants⁴⁷, and often occur with breathy phonation; in certain environments they exhibit a gentle rising tone, and some voiced stop onsets are prenasalized. In the examples in Table 6.5 below, “..” indicates breathy phonation and low tone, while the diacritics on the words for ‘tiger’ and ‘copper’ indicate high [checked] falling and low rising tones, respectively.

Table 6.5 Tokpe Gola Tibetan: Tone patterns in monosyllabic non-verbs

Tone	IPA	WT	Gloss
H tone:	pū	<i>spu</i>	hair, fur
	p ^h ū	<i>phu</i>	upper part of a valley
L tone:	ḅu	<i>sbug</i>	inside
	^m ḅu	<i>'bu</i>	insect
H tone:	tāk	<i>stag</i>	tiger
	t ^h āk	<i>khrag</i>	blood
L tone:	t ^h ak	<i>brag</i>	boulder
	ḍak	<i>bdag</i>	1 SG.HUM †
H tone:	lō	<i>glo</i>	cough
L tone:	ḷo	<i>lo</i>	year
H tone:	sāŋ	<i>sang</i>	tomorrow
L tone:	sāŋ	<i>zangs</i>	copper

† 1st person singular humilific

As discussed in section 1.3, there is a direct correlation between these tone patterns and the Written Tibetan spelling. For instance, the onset consonant of [ḅu]

⁴⁷ Tibetan dialects differ in terms of where tone splits occur. In Tokpe Gola, voiceless aspirated onsets can occur with either low or high tone – e.g., [p^hu] *phu* ‘upper part of a valley’ vs. [p^ha^hk] *brag* ‘boulder’ in Table 6.5 above. In Kyirong (Huber 2005), voiceless plain onsets can occur with either low or high tone. Such dialectal differences can be correlated systematically with WT spelling.

sbug ‘inside’ is predictably voiced because the root letter *b* has an *s* superscript.

Likewise, [ᵐbu] ‘bu ‘insect’ is predictably prenasalized because the root letter *b* is

preceded by the prefix *a.chung* (represented by an apostrophe in the Wylie

transliteration). The high tone of [lō] *glo* ‘cough’ corresponds to the fact that *l* here is

a subscript to the root letter *g*, rather than a root letter itself, as it is in [lō] *lo* ‘year’.

The high, checked falling tone of [tāk] *stag* ‘tiger’ corresponds to the combination of the *s* superscript and the *g* suffix.

These Tokpe Gola tone patterns are quite typical. All of the Innovative dialects display a distinction between high and low register tones⁴⁸; some distinguish a mid tone, too. Many exhibit high, checked falling tones and low rising tones under the same predictable circumstances, and some are reported to contrast level and moving contours. In some cases, authors offer different opinions regarding the significance of contour tones of a single dialect; this may be the result of working with different speakers or different sociolinguistic registers, or of simply coming to different conclusions about tonemic contrasts. Local variations on the pattern are usually systematic and can be correlated to Written Tibetan forms.

There is a significant body of literature reporting these general characteristics of tone on monosyllabic words in Lhasa Tibetan, including, among many others:

Chang and Shefts (1964), Dawson (1980), Denwood (1999), Goldstein and Nornang

⁴⁸ Because tone in Tokpe Gola – and other Innovative dialects – can be regarded as a bundle of features associated with the entire syllable, a contrast in tone is generally accompanied by a contrast in onset phonation, or vowel phonation, or both. Such feature-bundling in Tibetan is sometimes referred to as tone register – e.g., Hari (1980).

(1984), Haller (1999), Hari (1980), Hu (1982), Kjellin (1975, 1976), Mazaudon (1977), Meredith (1990), Sprigg (1955), and Yang (1974).

Bielmeier (1988a) discusses similar tone patterns in the Drokpa, Lhasa, Kyirong, Dzar, and Kagate dialects.

Huang (1995) comments on the tone patterns of Lhasa, Shigatse, Sherpa, Dege, Zhongdian, Muyu, Zhouqu, and Ruoergai. (Note that some of the Chinese dialect names she gives may have alternate Tibetan forms.)

Sun (2001) describes tone register and contour in the Lhasa, Zaiduo, Dege, Baima, Zhangla, Qiuji, Zhuoni, Rikeze [Shigatse], Zhibo, and Langkazi dialects of Tibetan. (Again, Chinese names may have alternate Tibetan forms.)

Other dialects in which tone register and tone contour are relevant include Kyirong (Bielmeier 1982; Huber 2005) and Shigatse (Haller 2000).

6.3.1.2 Disyllabic words

Disyllabic nouns and adjectives in Tokpe Gola also fall into two tone categories: low-high (LH) and high-high (HH), as evidenced by perceptual and acoustic data described in Caplow (2004). Minimal pairs contrasting only in terms of tone are exceedingly rare; in fact, I do not believe I have any examples at all in my data.

Table 6.6 below includes examples of monomorphemic, compound, and reduplicated forms. Crucially, while the first syllable may be either low or high, the second syllable is always high. It is exactly this prohibition against *LL and *HL tone

patterns that I account for, in section 6.4 below, in terms of the stress patterns and stress correlates reconstructed for Proto-Tibetan.

Example (a) in Table 6.6 shows that, when two disyllabic words are simplified and combined to form a disyllabic compound (a very common phenomenon in Tibetan), σ_2 will always have a high tone, regardless of the tone in the component parts. When two low-toned monosyllabic words are joined to form a compound, the second element will take on a high tone, as shown in example (b). Example (c) shows that the restriction to LH and HH tone patterns pertains to reduplicated forms as well, even when σ_1 has a low tone.

Table 6.6 Tokpe Gola Tibetan: Tone patterns in disyllabic non-verbs

	Tone	IPA	WT	Gloss
a.	HH	kāŋ.pā	<i>rkang.pa</i>	leg, foot
	LH	lak.pā	<i>lag.pa</i>	arm, hand
	HH	kāŋ.lāk	<i>rkang.lag</i>	limbs (arms & legs)
b.	L	mē	<i>me</i>	fire
	L	tā	<i>mda'</i>	arrow
	LH	mēn.dā	<i>me.mda'</i>	gun, rifle
c.	HH	ḥūk.ḥūk	<i>lhug.lhug</i>	loose, floppy
	LH	lēp.lēp	<i>leb.leb</i>	flat
	LH	ṣḥ̄:ṣḥ̄	<i>gzhon.gzhon</i>	young; youth

Another important feature of Tokpe Gola disyllabics is that tone in σ_1 is always level – if low, it is never rising; if high, it is never falling – even though rising and falling contours are seen in monosyllabic words. The H tone in σ_2 of disyllabics

may be level, or may have a falling tone in certain phonological environments – i.e., with certain codas.

The words in (b) above illustrate another phenomenon typical of the Innovative dialects: consonant clusters that are reduced in word-initial and word-final positions are preserved at the syllable juncture. That is, in Tokpe Gola [t̪a] *mda* ‘arrow’, the nasal prefix of the WT form is elided (though it is preserved in Balti and Rebkong Amdo, in both of which the word ‘arrow’ is pronounced [nda]). But in the Tokpe Gola compound [m̪ɛ̃n.d̪ā] *me.mda* ‘gun, rifle’, the nasal prefix is present, now re-syllabified as the coda of the preceding syllable. This retention of word-medial consonant clusters is relevant to consideration of the development of tone in Tibetan, to be discussed in sections 6.3.2 and 6.4.

Disyllabic verb complexes in Tokpe Gola may show different tone patterns than nouns and adjectives, and may even contrast in terms of tone, as illustrated in Table 6.7 below.

Table 6.7 Tokpe Gola Tibetan: Tone patterns in disyllabic verbs

Tone	IPA	WT	Composition	Gloss
LH	ny̥iː.l̥ɔq	<i>nyi.log</i>	<i>nyi.ma</i> LH ‘day’ + <i>log</i> L ‘to return’	to go and return in one day
HH	ny̥iː.l̥ɔq	<i>gnyid.log</i>	<i>gnyid</i> H ‘to sleep’ + <i>log</i> L ‘to return’	go to sleep!

In describing tone patterns in disyllabic words in other Innovative dialects, previous researchers have generally not attended to the potential for a distinction

between non-verbs and verbs. This can yield inaccurate or ambiguous descriptions and analyses: if sufficient information is not provided, a reader may not be able to determine, for instance, whether HL tone is actually observed in nouns – which would be noteworthy – or if it only occurs in verbs. In one or two publications, a reported phonemic tone contrast is illustrated by minimal pairs or triplets which compare a noun, a verb, and/or a noun stem + enclitic. I do not believe such examples constitute satisfactory evidence for Tibetan. For the most part, however, the descriptions offered for other Innovative dialects are explicitly based on non-verbs only, since these can be elicited very easily and very naturally. In these cases, the tone patterns reported are usually the same as those I observed in Tokpe Gola – i.e., disyllabic nouns and adjectives can only be HH or LH; the tone in σ_1 must be level, while in σ_2 it may be level or falling. Thus when I account for elements of Tokpe Gola tone in terms of historical stress patterns (in section 6.4), my conclusions are of broad relevance across the Tibetan dialects.

There is a significant body of literature reporting the same (or similar) tone restrictions in Lhasa Tibetan as in Tokpe Gola: in disyllabic words, tone is contrastive (L vs. H) only in σ_1 ; tone in σ_2 is always H, with some authors distinguishing between level and falling contours. Among the authors who describe such patterns for the Lhasa dialect are Chang and Shefts (1964, 1968: 3), Dawson (1980), Denwood (1999: 77), Edmondson et al (n.d.), Hu (1982), Mazaudon (1977: 82), Saxena (1991), Sprigg (1955), and Sun (1997). The pitch patterns described by Kjellin (1975, 1976)

are also consistent with this tone pattern.⁴⁹ Goldstein and Nornang's (1984: xv) description of tone is slightly different: They mark tone (either L or H) only on $\sigma 1$ of disyllabic words, and consider $\sigma 2$ to have a mid tone, which is higher than a $\sigma 1$ low tone, and lower than a $\sigma 1$ high tone.

In Tsang, as well, $\sigma 1$ may have a H or L tone, but $\sigma 2$ can only have a H tone (Ossorio, 1982). In Refugee Standard Tibetan, disyllabic nouns have only LH or HH tone patterns (Meredith 1990). The tone on $\sigma 1$ is always level, while the high tone on $\sigma 2$ may be either level or falling. Haller (1999, 2000) reports the same patterns for Shigatse, another Central dialect. In Kyirong (Huber 2005), $\sigma 1$ of disyllabic words may have low, mid, or high registers, but $\sigma 2$ is again always high. Dolpo (Watters 2002) also allows only a high tone in $\sigma 2$. In southern Mustang (Bielmeier 1988c), a phonemic H vs. L tone contrast only occurs in $\sigma 1$.

To the east, in Sichuan province, Lin (2002) reports that Thewo Tibetan also has contrastive H and L register tones in $\sigma 1$ of disyllabic nouns, while tone is always high in $\sigma 2$. For Bathang, a Khams dialect, Haller (1999) does not include an explicit phonological description, but the disyllabic examples provided all have only LH and HH tone patterns.

Thus for many of the dialects that have been described – and there are other cases that I have not mentioned here – the dominant pattern observed in disyllabic non-verbs is that tone is contrastive in $\sigma 1$ (low vs. high register), but is always high in

⁴⁹ Though in his view, Lhasa Tibetan is not a tone language; he considers the pitch patterns to be entirely derivable from segmental characteristics.

σ_2 . This dominant pattern is what I account for in terms of Proto-Tibetan stress in section 6.4.

There are some exceptions – and it is possible that more will be reported in future. Local varieties of Sherpa can have a low tone on σ_2 (Graves 2007; Kelly 2005; Sun 2003; Watters 2002). In fact, Watters (2002) notes that a high tone in σ_2 is actually rare in Sherpa. And in Bartee’s (2007) grammar of the Dongwang variety of Khams, the glossary includes convincing examples of nouns and adjectives with HL register tone patterns, along with forms that can be analyzed as LH and HH. Accounting for these rare patterns is beyond the scope of the present study.

6.3.2 Previous research on Tibetan tonogenesis

An adequate theory of Tibetan tonogenesis must account for the tone patterns described above. A number of hypotheses have been put forward by previous researchers, all focusing primarily on the relationship between consonants and tone. What is new in this dissertation is that I consider not only historical segmental features, but also historical suprasegmental features.

As noted previously, Proto-Tibetan (like the Archaic dialects and like Written Tibetan) has been reconstructed as rich in syllable-initial onset clusters and in syllable-final codas and coda clusters. Historically, these corresponded to prefixes, suffixes, and post-suffixes which may have had morphological significance as causative markers, tense/aspect markers, and classifiers, among other functions (e.g., Beyer 1992).

Both Matisoff (1973) and Mazaudon (1977) refer to the effects which onset and coda consonants – such as Proto-Tibetan root letters and affixes – may have on nuclear vowels. That is, voiced onsets correspond to lower pitch register; voiceless onsets correspond to higher pitch register; and laryngeal final consonants correspond to a falling contour in the preceding vowel. These patterns are largely attributable to universal laryngeal and physiological factors, as discussed in Hombert, Ohala, and Ewan (1979), and so are wide-spread cross-linguistically, manifested as both contrastive and non-contrastive pitch differences. In Tibetan, over time, these non-contrastive tones took on a phonemic function, yielding a tonal split.

6.3.2.1 Matisoff (1973)

Matisoff (1973) suggests that languages which are basically monosyllabic are most prone to the development of tone. Their cohesive structure favors the constant influence of various parts of the syllable upon one another, and the shift of contrastive function from one feature to another (a phenomenon which Mazaudon 1977 refers to as “transphonologization”).

For the development of tone in Tibetan, Matisoff proposes a recurring cycle of word structure, tone birth, and tone decay. Looking in on the cycle at an arbitrary starting point, we observe monosyllabic words with prefixes and suffixes, in which consonants have an effect on pitch, but pitch patterns are redundant rather than contrastive.⁵⁰ In the next stage, initial and final consonants break down, affixes are dropped or absorbed into the root morpheme, and “homophony rears its ugly head”

⁵⁰ Some authors – including Huang (1995) – refer to this as “natural tone”.

(1973:83). At this point, tone becomes distinctive. Matisoff suggests that compounds are formed here to help disambiguate between potentially homophonous forms. Next, syllables in compounds are reduced: vowels lose their stress, reduce to schwa, lose their tone, and start to become absorbed into neighboring syllables. As syllables become fused, new consonant clusters develop, and the functional load of vowel quality differences wanes. Words tend once again towards monosyllabism, and new affixes arise, co-existing with some of the older ones. The cycle then repeats.

Other scholars generally agree with Matisoff's hypothesis. Among the many papers which contribute to the discussion of the development of tone in Tibetan are Sprigg (1972), Yang (1974), Kjellin (1975), Nishida (1975), Takata (1981), Mazaudon and Michailovsky (1988), DeLancey (1989), Saxena (1991), and Haller (1999).

Huang (1995) and Sun (2001) discuss the diversification of tone patterns in the Tibetan dialects, and the varying degrees to which tone has developed as a contrastive feature. Mazaudon (1977) elaborates on certain details and shows how similar patterns can be observed in other Tibeto-Burman languages.

6.3.2.2 Mazaudon (1977)

This latter paper includes a discussion of tone in Lhasa Tibetan. Mazaudon is in accord with previous researchers in her view that tone in monosyllabic words can be explained in terms of the influence of syllable onsets on tone height (or register) and the influence of syllable codas on tone contour.

However, she argues against Matisoff's view of the temporal relationship between the formation of compounds and the development of tone. As noted above, Matisoff suggests that the process of compound formation was initiated just as tone began to serve a contrastive function. Mazaudon, on the other hand, suggests that tone can arise on polysyllabic words as well as monosyllabic words, and that compounds must have existed before the phonemicization of tone. (Though Mazaudon does not point it out, the evidence in support of her view is abundant and clear: the many compound words which occur in the lexicons of both non-tonal Archaic dialects and tonal Innovative dialects offer proof that compound formation pre-dates tone in Tibetan.)

This leads Mazaudon to address the development of tone on polysyllabic words, a subject which is neglected by most other authors. What requires explanation, for disyllabic words, is why the first syllable can manifest only high level or low level tones, and the second syllable can manifest only high level or high falling tones. Her solution is to suggest that the limited possible tone patterns observed in disyllabics can be explained by the hypothesis that the domain of tone is the word, in the sense that "... the main tonal contrast is the high/low correlation which starts at the very beginning of the first syllable and then develops its melody up to the end of the word" (1977: 83). However, Mazaudon does not elaborate fully on what she means by tone "develop[ing] its melody up to the end of the word". My impression is that she means that σ_1 of a disyllabic word can only have a low or high register (the registers observed in monosyllabic words, corresponding diachronically with onset

consonants), and σ_2 can only have a level or falling contour (the contours observed in monosyllabic words, corresponding diachronically with coda consonants). It seems, though, that the high register observed in σ_2 is left unexplained.

Mazaudon does *not* seem to mean that all of the possible combinations of register and contour observed in monosyllabic words can be observed in disyllabic words – at least, not in the ways we might anticipate if we are familiar with word-tone and tone spreading typical of African languages. For instance, her schemas for Lhasa Tibetan (p. 82) include a monosyllabic word with a high falling tone. One might expect such a “melody” to be manifested on a disyllabic word as a $\sigma_1.\sigma_2$ / H.L tone pattern. But of course, such a pattern is generally not found in the Innovative dialects, as discussed in section 6.3.1.2. Instead, her corresponding disyllabic schema has a high level tone on σ_1 , and a high falling tone on σ_2 . The register of σ_1 and the contour of σ_2 are accounted for, but the high register of σ_2 is not addressed. Mazaudon herself seems resigned to this, observing that, compared to monosyllabic words, “the tonal system on polysyllabic items is more straightforward, and at the same time more difficult to account for diachronically” (1977: 81).

I agree with Mazaudon that “[i]t is impossible to explain the modern system if one supposes that tone developed on a still monosyllabic language, or on each syllable of a polysyllabic language. The idea that Tibetan used to have, and still underlyingly has, a tone per syllable, leads to embarrassing situations...” (ibid, p. 89). However, I do not agree that the facts can be best accounted for by a word-level tonal

melody that spreads from the beginning of a word to its end. Instead, as I go on to discuss in section 6.4, I suggest that tone is templatic.

6.3.2.3 Sun (1997, 2001)

Sun (1997) also rejects the notion of tone spreading as an account for the tone patterns observed in Tibetan disyllabics. He notes (on p. 489) that “[o]ne of the most important generalizations on Tibetan tone... is that the primary register contrast is realized only on the initial syllable of the phonological word; all other syllables are predictably high-registered.” He concludes (p. 513) that “[t]his type of tone system is so restricted that tone in polysyllabic Tibetan words may be viewed as adhering by and large to a simple tone template.” Secondary contour distinctions (such as those described in the sections above) are then superimposed on this template (1997: 515).

According to Sun, the tone contrasts and patterns observed in a dialect like Lhasa Tibetan evolved from an original state in which “all syllables were normally produced in the high register” (1997: 508). Later, he claims that “Old Tibetan did not seem to have tone, even at the phonetic level ... In this assumed purely atonal stage ... syllables that are not inflectional suffixes are generally high-registered” (Sun 2001:2).

From here, Sun suggests that “[t]he first significant change altering this incipient state was the emergence of the **non-contrastive** low register in Amdo

Tibetan conditioned either by voiced initials or by syllable quantity, depending on the dialect” (1997: 508; bold in the original) ⁵¹.

Sun invokes autosegmental phonology to account for the fact that tone is always H on σ_2 of a disyllabic non-verb.⁵² That is: “The high tone in Tibetan non-initial syllables, in other words, results ... from phonological neutralization reducing the original tonal contrast to a non-distinctive high register. In our analysis, this generalization is conveyed by a tone-deletion rule which cancels (neutralizes) the original lexical tones on non-initial tone-bearing syllables, and a **default-tone rule** which fills the empty tone slots with the default value H” (Sun 1997: 503; bold in the original).

So, illustrating with examples from Lhasa Tibetan, Sun says that, in σ_1 , H and L tones reflect a fundamental register distinction. He attributes the predictable H tone on σ_2 to tone neutralization under an autosegmental analysis, and suggests that high tone is the unmarked default tone in Tibetan. An apt example from Tokpe Gola Tibetan was included in Table 6.6 above: the compound noun [mɛ̃n.dā̃] *me.mda* ‘gun, rifle’ is composed of the monosyllabic words [mɛ̃] *me* ‘fire’ and [tā̃] *mda* ‘arrow’, both of which have a low tone. Sun would suggest here that when the two

⁵¹ Sun (2001: 3-9) suggests that L register in σ_1 can be fostered in various dialects not only by onset voicing, but also by rhyme length and onset aspiration.

⁵² Sun does not refer specifically to disyllabic non-verbs, and his analysis is intended to apply to words of up to three syllables. But the examples he presents are all non-verbs, and there is no difficulty in simplifying his ideas to talk about disyllabic words here.

elements are compounded, the tonal association of σ_2 is de-linked and defaults to a high tone, yielding a LH pattern.⁵³

Sun offers a diachronic explanation for his proposed default high tone (1997: 508): “Comparative evidence presented in §1.1 suggests that Tibetan originally must have been in a state where, *the effects of stress and intonation aside*, all syllables were normally produced in the high register”. The emphatic italics here are mine: I do not believe that the effects of stress (and intonation) can be put aside. In section 6.2, I reconstructed the stress patterns and the acoustic correlates of stress for Proto-Tibetan, based on my analysis of stress in Balti and Rebkong Amdo. This reconstruction shows unequivocally that pitch was *not* equally high in both syllables. Rather, for Proto-Tibetan non-verbs, pitch in σ_1 was low compared to pitch in σ_2 (and the opposite is the case for verbs). An account of Tibetan tonogenesis must begin from that state.

Furthermore, it is not clear to me that the data which Sun refers to in his §1.1 actually suggests an “original state” in which all syllables were produced with a high tone. What he presents in this section is a tonal continuum for Tibetan, ranging from dialects which are completely atonal to dialects in which the tone system is highly evolved, with phonemically contrastive contours as well as registers. For the atonal dialects – which include Ndzorge Amdo (mdzod.dge; Sun 1986) – he says “all syllable types [he refers to monosyllabic words here - NJC] carry a high (falling) tone

⁵³ Yip (1993) and Meredith (1990) also offer autosegmental accounts for the H tone which is always found in σ_2 , and Meredith similarly accounts for the absence of contour tone on σ_1 .

when uttered in isolation, whereas the initial syllable of polysyllabic words are predictably low-pitched.” (This is just what I have documented for Rebkong Amdo.) Since these are reported to be atonal dialects, and they have a low pitch on σ_1 , I am not able to see how this demonstrates that the original, default tone is high for all syllables.

Thus while I agree with Sun’s ideas about templatic tone with superimposed consonantal effects at the boundaries, my views differ from his on several points: (a) I do not agree with his assumption that pitch was originally high on all syllables of disyllabic words. On the contrary, my reconstruction of Proto-Tibetan shows that pitch was only high on σ_2 , functioning as a correlate of stress. (b) I find his argument about the origin of a “default” high tone to be unconvincing. (c) Sun (1997, 2001) suggests that the first change to this system was the development of low register, conditioned by onset voicing, onset aspiration, and rhyme length. What I have shown for Balti and Rebkong Amdo is that the first syllable of disyllabic non-verbs in Amdo almost *always* has a lower pitch (except for a minority of the AR_05 isolation forms, in which pitch slope was the more important acoustic correlate of stress), and there is no evidence that this was conditioned by anything. It occurs as a contrast to σ_2 stress.

6.3.2.4 Bielmeier (1988a)

Bielmeier (1988a) makes an original and important contribution to the development of ideas regarding Tibetan tonogenesis. It seems there was no precedent to his suggestion that stress may play a role in the process – nor has there been any subsequent follow-up, as far as I know, until the present study.

Bielmeier proposes that stress originated in Tibetan on σ_2 . Noting an occasional σ_1 stress in some of the Innovative dialects provoked the idea that a genuine shift in stress may have taken place, and that this may correspond somehow to observed tone patterns.

In lieu of the impossible task of tracing such potential changes through time, Bielmeier traces them through space, by describing perceived patterns of stress and tone in disyllabic non-verbs along a linguistic transect from Balti, in the west, to Kyirong and Dzar, near the border between central Nepal and Tibet.

A historical pattern of second-syllable stress in Tibetan is demonstrated by the pattern which occurs in Balti, classified by Bielmeier as a phonologically conservative Western Archaic dialect. The next stage is illustrated by the apparently less conservative Western Archaic dialects of Ladakh. Here, stress is perceived sometimes on the first syllable, sometimes on the second, and is sometimes even across syllables.

Continuing eastward, Bielmeier finds that stress is still traceable in the southwestern and central tonal dialects, in which it is "... weakened and tends to shift from the second to the first syllable" (p. 49). Finally, in Innovative dialects such as Kyirong (in Tibet, north of central Nepal) and Dzar (Jharkot, in southern Mustang, Nepal), he observes that, among "disyllabic words where a morpheme boundary is not evident or only historical" (p. 49), those with a voiceless unaspirated onset (like [p]) and H tone on the first syllable are also stressed on σ_1 . In these more "evolved"

dialects, then, he suggests there may be a fairly regular correlation between onset type, stress, and tone.

In words with a voiceless aspirated onset (like [p^h]) and H tone on the first syllable, the pattern is less consistent, and stress may vary across dialects: sometimes stress is even across syllables, and sometimes there is a slight preference for $\sigma 2$ stress. Words with L tone on the first syllable have $\sigma 2$ stress. Stress also tends to remain on $\sigma 2$ in compound words. Thus it is only in some dialects – and with specific $\sigma 1$ onsets – that the stress shift appears to be most consistent.

Bielmeier does not make a claim that this stress shift is responsible for the high tones observed in $\sigma 1$ of so many disyllabic non-verbs. The laryngeal and physiological correlations between voiceless onsets and higher pitch – and between voiced onsets and lower pitch – are recognized as playing a fundamental role, and this is, in fact, the focus of much of the paper. But Bielmeier observes that the stress pattern – though sometimes weak and not entirely consistent – often matches the tone pattern, and wonders if there is a causal relationship.⁵⁴

This notion of a causal relationship provokes a number of questions. Is a shift in stress from $\sigma 2$ to $\sigma 1$ triggered by particular $\sigma 1$ onset types? If such a shift occurred, did it somehow induce a H tone on $\sigma 1$? This seems unlikely, as there are plenty of words with a H tone on $\sigma 1$ in which stress has *not* shifted – words with a voiceless

⁵⁴ These statements reflect the research questions which Bielmeier was considering at the time. He has long since rejected the notion that there might be a correlation between stress and register tone. (Bielmeier p.c., 2009)

aspirated onset (like [p^h]). Here, since stress has not shifted, it cannot be credited with the H tone observed, which must instead be attributed to the characteristics of the onset.

Conversely, where stress *has* shifted to $\sigma 1$ (with an onset like [p]), any effect of stress on pitch could be considered redundant. The correlation between $\sigma 1$ onset type and tone has already been established, and there is no need to invoke stress to account for it.

A third possibility is that stress began to shift to $\sigma 1$ *because* it has H tone. If this were the case, then it would be impossible to claim stress has played a role in motivating the genesis of tone after all.⁵⁵

These questions cannot be approached without a careful acoustic analysis of the dialects considered. Such an analysis would also be helpful in determining how consistent and how robust the patterns described by Bielmeier are. As I discussed in detail in section 3.1.2, contrasts in vowel height across syllables can significantly impact our perception of relative stress. In that case, Sprigg (1966, 2002) perceived stress on the first syllable of certain disyllabic words in Balti, and on that basis identified particular suffixes as non-stress-bearing. What I demonstrated, instead, is

⁵⁵ In fact, I *do* think this is the case. I think that the erratic $\sigma 1$ stress Bielmeier perceives in some of the Innovative dialects reflects not a shift in stress, but a new, emergent stress. In my view, the function of this incipient prosodic element is to highlight $\sigma 1$ for the benefit of the listener; this is the only syllable on which tone is contrastive – it can be either H or L only here. Since it is important to convey this information, the speaker may lend prominence to this syllable. I suspect such prominence would be conveyed by vowel duration, vowel quality, or intensity – but not by fundamental frequency, since this acoustic resource is already operating at its functional capacity, in conveying tone. But this is a research project for another day.

that in nearly all of the cases Sprigg considered, there was a [low.high] vowel contrast across syllables. His perception of σ_1 stress may thus have been attributable to the fact that lower vowels have an intrinsically higher intensity. This prominence should then be considered of phonetic – but not of phonological – interest. It is possible that such incidental factors as a contrast in vowel height across syllables, or a contrast in syllable template, may play a role in Bielmeier’s perceived stress patterns in Kyirong and Dzar as well. Thus a closer examination of these and similar dialects must control for such incidental factors.

While I think a shift in stress from σ_2 to σ_1 is unlikely, I am in complete agreement with Bielmeier’s core concepts: that stress has played an important role in Tibetan tonogenesis; that “...in internal proto-reconstruction of Tibetan we should start with stress” (1988a: 53); and that what we have in Tibetan is “...an ‘underlying accentual representation turning into a surface tonal representation’ (Goldsmith 1980:415)” (1988a: 52). Indeed, I agree that the modern spoken Innovative dialects preserve a trace – a continuity – of a historical σ_2 stress pattern (which Bielmeier infers from his observations of conservative Western Archaic dialects and which I have reconstructed here for Proto-Tibetan). However, as I show in section 6.4 below, this trace is actually manifested as σ_2 tone, not as shifting stress.

My approach differs from the concept that Bielmeier put forth in two important respects: First, while Bielmeier was pondering a potential shift in stress from σ_2 to σ_1 , I focus instead on the preservation of stress features on σ_2 . Second, as

described in section 6.3.1.2, disyllabic non-verbs in the Innovative dialects generally exhibit only LH and HH tone patterns. Bielmeier's idea was focused on the origin of the H tone on $\sigma 1$ of HH words. My work addresses the H tone which occurs on $\sigma 2$ of both LH and HH patterns.

Finally, aside from the originality of taking stress into consideration, Bielmeier's approach differs from the studies described above in another important element: implicit in his analysis is the notion that pitch was historically low on $\sigma 1$. This is what my reconstruction of stress correlates in Proto-Tibetan confirms, but it is exactly the opposite of the common assumption that pitch was historically high on all syllables.

6.4 A new view of the role of stress in Tibetan tonogenesis

In my view, the tone patterns observed on disyllabic words in many dialects of Tibetan can be explained diachronically as a combination of two processes.

In part, the possible tone patterns are constrained by an acoustic template inherited from the stress patterns of Proto-Tibetan. Specifically, for disyllabic non-verbs, a historical $*\sigma\acute{\sigma}$ pattern has yielded only LH and HH tone patterns; $\sigma 2$ can never have a L tone.⁵⁶ This restriction on $\sigma 2$ tone is a direct reflection of the acoustic history of disyllabic non-verbs. As reconstructed in section 6.2, non-verbs in Proto-

⁵⁶ For disyllabic verbs, a historical $*\acute{\sigma}\sigma$ stress pattern allows HL tone patterns. (LH tone patterns also occur on verbs.) Since my data for verbs in Balti and Rebkong Amdo Tibetan was much more limited, my claims are correspondingly more limited and more tentative. Thus the analysis of Tibetan tonogenesis here focuses on nouns.

Tibetan were stressed on σ_2 , and this stress was conveyed primarily by fundamental frequency. This prominence of F0 on σ_2 has evidently been quite robust and has persisted over time, even as its function changed, and regardless of variations in pitch (or tone) in σ_1 . This σ_2 F0 prominence is now manifested in the Innovative dialects as a σ_2 H tone. In this way, Proto-Tibetan stress patterns are reflected in – and even constrain – the tone patterns observed in the modern spoken Innovative dialects.

Superimposed on this templatic tone are the consonantal “edge” effects discussed by Matisoff (1973), Mazaudon (1977), and others, which I suggest operate from both ends of a word towards its middle (just as they do on monosyllabic words). That is, the erosion of word-initial consonants induces a H/L register contrast on the initial syllable, the erosion of word-final consonants induces a level/falling contour contrast on the final syllable, and consonants and consonant clusters remain intact at medial syllable junctures. My hypothesis does suggest that the domain of tone is the word, but is much different than Mazaudon’s (1977) view that a tonal melody spreads from the beginning of the word to its end, and does not call on tone sandhi as a factor.

All of the features commonly observed in the Innovative dialects can thus be explained by superimposing consonantal edge effects on a historical acoustic template.

My analysis is quite similar to that of Sun (1997, 2001), who also regards tone in the Innovative dialects as templatic, with superimposed edge effects. As discussed in section 6.3.2.3 above, Sun suggests that [for disyllabic non-verbs] the original, lexical tone of σ_2 is deleted, and the empty tone slot then defaults to a H tone.

As discussed, I do not find Sun's diachronic explanation for the default H tone to be very convincing. Instead, I am able to offer acoustic evidence of an alternative explanation – the H tone observed on σ_2 is a relict of Proto-Tibetan, preserved as part of an acoustic template.

6.5 Implications regarding the direction of tone split in Tibetan

Except for Bielmeier (1988a), previous analyses of Tibetan tonogenesis are based on an assumption that a high tone was the original and default tone on all syllables. Sun (1997, 2003) in particular is quite succinct and explicit in this assertion. As I quoted previously: “Tibetan originally must have been in a state where... all syllables were normally produced in the high register” (1997: 508), with low tone emerging later. And again (2003: 37), he suggests that the history of tone in Tibetan languages “can be characterized by the genesis of the low register, which has steadily invaded the former territory of the high register”.

Thus the general consensus has been that a historical *HH tone pattern split to yield both LH and HH tone patterns in the modern Innovative dialects.⁵⁷

But this assumption of an original, default high tone on all syllables is not consistent with what I have documented and reconstructed here. For both Balti speakers and for Rebkong Amdo speaker AR_04, pitch is consistently and significantly lower in σ_1 than in σ_2 , in both isolation and frame forms. For Rebkong Amdo speaker AR_05, this is the case for the frame forms; in the isolation forms, the

⁵⁷ Though it is hardly ever stated explicitly, these views are based on consideration of disyllabic non-verbs.

pitch difference is not so dramatic because pitch slope is the primary stress correlate. Logically, then, it is impossible to reconstruct a *HH tone pattern for Proto-Tibetan. A *LH tone pattern is far more plausible.

Thus I suggest here that hypotheses regarding tonogenesis in Tibetan must begin with a *LH tone pattern on disyllabic non-verbs. This *LH tone pattern must then have split to yield the LH and HH patterns observed today in the modern Innovative dialects.

Both starting points yield the same result: only σ_1 has the potential to manifest a tone contrast – L or H – while in σ_2 tone remains H. In previous analyses, the σ_1 L tone is derived; in this new analysis, the σ_1 H tone is derived. The ramifications of this view will require further consideration.

6.6 For future research...

This dissertation suggests a number of topics which merit further investigation:

- Pitch accent: Will consideration of words longer than two syllables provide evidence that Tibetan can productively be analyzed as some type of pitch accent language?
- Tone as syntagmatic (related to the issue of pitch accent): Is it always appropriate to regard tone as a paradigmatic prosodic feature? In this dissertation, I reconstruct syntagmatic stress patterns for disyllabic words in Proto-Tibetan. I then demonstrate how the tone patterns of the modern spoken

Innovative dialects have evolved from this Proto-Tibetan stress. Because fundamental frequency in these dialects now functions primarily to convey tone, rather than stress, does this mean we instantly need to describe prosodic contrasts as paradigmatic, rather than syntagmatic? Does this question have broader cross-linguistic relevance?

- Stress “shift” in Tibetan: Bielmeier (1988a) reports σ_1 stress on disyllabic words in *tonal* dialects of Tibetan. This has been reported by other scholars, as well. I think this may represent not a shift in stress from its historical σ_2 locus, but a new, emergent prosodic feature, whose function is to highlight σ_1 since it is the only syllable on which tone is distinctive.
- Historical comparative acoustics: What other types of acoustic reconstruction have been described in the phonetics literature? In what other ways can this methodology be employed in future?
- The role of stress in tonogenesis: Are there other language groups in the world in which acoustic traces of a proto stress pattern can be found in modern tone languages? Are there language groups in the world in which acoustic traces of a proto tone pattern can be found in modern stress languages? And, conversely, can proto acoustic patterns be reconstructed for other language groups by comparing reflexes of stress and tone in various daughter languages?

I look forward to pursuing some of these questions in my future research, and hope they are intriguing to others as well.

REFERENCES

- Bartee, Ellen. 2007. *A grammar of Dongwang Tibetan*. PhD Dissertation, University of California Santa Barbara.
- Beyer, Stephan V. 1992. *The Classical Tibetan language*. Albany: State University of New York Press.
- Bielmeier, Roland. 1982. Problems of Tibetan dialectology and language history with special reference to the sKyid-grong dialect. *Zentralasiatische Studien* 16: 405-425.
- Bielmeier, Roland. 1985a. *Das Märchen vom Prinzen Čobzan. Eine tibetische Erzählung aus Baltistan. Text, Übersetzung, Grammatik und westtibetisch vergleichendes Glossar*. Sankt Augustin: VGH Wissenschaftsverlag.
- Bielmeier, Roland. 1985b. A survey of the development of western and southwestern Tibetan dialects. In B.N. Aziz and M. Kapstein (eds.), *Soundings in Tibetan Civilization*. New Delhi: Manohar.
- Bielmeier, Roland. 1988a. On tone in Tibetan. In Helga Ueback and Jampa L. Panglung (eds.), *Tibetan Studies: Proceedings of the 4th Seminar of the International Association for Tibetan Studies*. Munich: Kommission für Zentralasiatische Studien, Bayerische Akademie der Wissenschaften.
- Bielmeier, Roland. 1988b. The reconstruction of the stop series and the verbal system in Tibetan. In *Languages and History in East Asia. Festschrift for Tatsuo Nishida on the occasion of his 60th birthday*, pp. 15-27. Kyoto / Shokado.
- Bielmeier, Roland. 1988c. A preliminary survey of the dialect of Mustang. In A. Wezler et al. (eds.), *Journal of the Nepal Research Centre VIII*: 31-37. Wiesbaden: Franz Steiner Verlag GmbH.
- Bielmeier, Roland, et. al. 2002. *Comparative Dictionary of Tibetan Dialects (CDTD)*. Fourth preprint, University of Berne: Tibetan Dialects Project.
- Bielmeier, Roland. n.d. *Tibetan dialects: Overview of the main dialect groups in their geographical distribution as used for the Comparative Dictionary of Tibetan Dialects (CDTD)*. Unpublished manuscript, University of Berne: Tibetan Dialects Project.
- Boersma and Weenink 2007. *Praat: doing phonetics by computer*. Institute of Phonetic Sciences, University of Amsterdam. <http://www.praat.org>.

- Bradley, David. 1997. Tibeto-Burman languages and classification. Papers in Southeast Asian linguistics No. 14: Tibeto-Burman languages of the Himalayas. pp. 1-72. Pacific Linguistics, A-86.
- Caplow, Nancy. 2004. Fundamental frequency and intensity in Tokpe Gola, a tonal dialect of Tibetan. Unpublished manuscript, University of California Santa Barbara.
- Chang, Kun and Betty Shefts (with the help of Nawang Nornang and Lhadon Karsip). 1964. *A manual of spoken Tibetan (Lhasa dialect)*. Seattle: University of Washington Press.
- Chang, Kun and Betty Shefts. 1968. Vowel harmony in spoken Lhasa Tibetan. University of California, Berkeley: Project on Linguistic Analysis, Phonology Laboratory, second series 7.
- Dawson, Willa. 1980. *Tibetan phonology*. PhD dissertation, University of Washington.
- DeLancey, Scott. 1989. Contour tones from lost syllables in Tibetan. *Linguistics of the Tibeto-Burman area* 12(2): 33-34.
- Denwood, Philip. 1999. *Tibetan*. Amsterdam, Philadelphia: John Benjamins.
- van Driem, George. 1993. *A grammar of Dumi*. Berlin, New York: Mouton de Gruyter.
- van Driem, George. 2001. Languages of the Himalayas: An ethnolinguistic handbook of the Greater Himalayan region. Volumes I and II. Leiden / Boston / Koln: Brill.
- Edmondson, Jerold A., Deji-Sezhen Geziben, and Michael Phillipini. n.d. A cross-lectal study of Tibetan tones: analysis and representation. Unpublished manuscript, 20 pages.
- Goldsmith, John. 1980. Review of tone: a linguistic survey. *Language* 56(2): 413-418.
- Goldstein, Melvyn C. and Nawang Nornang. 1984 (3rd edition). *Modern spoken Tibetan: Lhasa dialect (with Written Tibetan appendix)*. Kathmandu: Ratna Pustak Bhandar. (1st ed. 1970).
- Goldstein, Melvyn C. with Ngawangthodup Narkyid. 1984. *English-Tibetan dictionary of modern Tibetan*. Dharamsala: Library of Tibetan Works and Archives.

- Goldstein, Melvyn C. 2001. *The new Tibetan-English dictionary of modern Tibetan*. University of California Press.
- Gordon, Matthew and Peter Ladefoged. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29: 383-406.
- Grierson, George A. 1909. *Linguistic Survey of India, Vol. III, Part I. Tibeto-Burman family: Tibetan dialects, the Himalyan dialects, & the North Assam group*. Delhi: Motilal Banarstdas.
- Haller, Felix. 1999. A brief comparison of register tone in Central Tibetan and Kham Tibetan. *Linguistics of the Tibeto-Burman Area* 22 (2): 77-98.
- Haller, Felix. 2000. *Dialekt und Erzählungen von Shigatse*. [Beiträge zur tibetischen Erzählforschung, 13.] Bonn: VGH Wissenschaftsverlag GmbH.
- Haller, Felix. 2004. *Dialekt und Erzählungen von Themchen: Sprachwissenschaftliche Beschreibung eines Nomadendialektes aus Nord-Amdo*. [Beiträge zur tibetischen Erzählforschung, 14.] Bonn: VGH Wissenschaftsverlag GmbH.
- Hari, Anna Maria. 1980. *An investigation of the tones of Lhasa Tibetan*. [Language data, Asian Pacific Series, No. 13.] Huntington Beach, CA: Summer Institute of Linguistics.
- Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55(1): 37-58.
- Hu Tan. 1982. Recherches sur les tons du Tibetain (dialecte de Lhasa). *Cahiers de Linguistique Orientale* 11(1): 11-46.
- Huang Bufan. 1995. Conditions for tonogenesis and tone split in Tibetan dialects. *Linguistics of the Tibeto-Burman area* 18(1): 43-62.
- Huber, Brigitte. 2005. *The Tibetan dialect of Lende (Kyirong): a grammatical description with historical annotations*. Bonn: VGH Wissenschaftsverlag GmbH.
- Jaeschke, Hermann A. 1871. *Handwörterbuch der tibetischen Sprache*. Gnadau: Universitätsbuchhandlung.
- Jaeschke, Hermann A. [1881] 1958. A Tibetan-English dictionary with special reference to the prevailing dialects; to which is added a Tibetan-English vocabulary. London: Routledge and Kegan Paul.

- Kjellin, Olle. 1975. How to explain the “tones” in Tibetan. *Annual Bulletin* 10: 127-143. Tokyo: Research Institute of Logopedics and Phoniatics.
- Kjellin, Olle. 1976. A phonetic description of Tibetan, with a review of the literature. *Annual Bulletin* 9: 151-166. Tokyo: Research Institute of Logopedics and Phoniatics.
- Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1970. *Suprasegmentals*. MIT Press.
- Lin You-Jing. 2002. Phonological profile of Thewo Tibetan. Paper presented at the 8th Himalayan Languages Symposium. Berne, September 2001.
- Makley, Charlene, Keith Dede, Hua Kan, and Wang Qingshan. 1999. The Amdo dialect of Labrang. *Linguistics of the Tibeto-Burman Area* 22(1): 97-127.
- Matisoff, James A. 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study. *Journal of the American Oriental Society* 90(1): 13-44.
- Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In L. Hyman (ed.), *Consonant types and tone*, pp. 71-96. Southern California Occasional Papers in Linguistics. Los Angeles: University of Southern California.
- Mazaudon, Martine. 1977. Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman Area*, 3(2): 1-123.
- Mazaudon, Martine and Boyd Michailovsky. 1988. Lost syllables and tone contour in Dzongkha (Bhutan). In David Bradley, Eugenie J.A. Henderson, and Martine Mazaudon (eds.), *Prosodic analysis and Asian linguistics: to honour R.K. Sprigg*, pp. 115-136. Pacific Linguistics C-104.
- Meredith, Scott. 1990. *Issues in the phonology of prominence*. PhD dissertation, MIT.
- Nishida, Tatsuo. 1975. On the development of tones in Tibetan. *Acta Asiatica / Bulletin of the Institute of Eastern Culture* [Special issue: Tibetan studies in Japan] 29: 43-55.
- Norberg-Hodge, Helena and Gyelong Thupstan Paldan. 1991. *Ladakhi-English English-Ladakhi Dictionary*. Leh, Ladakh: The Ladakh Ecological Development Group.
- Ossorio, Janet R. W. 1982. *Tsang Tibetan Phonology*. PhD dissertation, University of Colorado.

- Peterson, G. and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32: 693-703.
- Pike, Kenneth L. 1948. *Tone languages*. Ann Arbor: University of Michigan Press.
- R Development Core Team (2007). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Read, A.F.C. 1934. *Balti grammar*. London: The Royal Asiatic Society.
- Roerich, Georges de. 1958. *Le parler de l'Amdo: étude d'un dialecte archaïque du Tibet*. [Serie Orientale Roma XVIII.] Rome: Instituto Italiano per il medio ed estremo oriente.
- Róna-Tas, Andras. 1966. *Tibeto-Mongolica: The Tibetan loanwords of Monguor and the development of the Archaic Tibetan Dialects*. [Indo-Iranian Monographs VII.] The Hague: Mouton.
- Saxena, Anju. 1991. Tone in PaTani and Central Tibetan: parallel developments? *Linguistics of the Tibeto-Burman Area* 14(1): 129-135.
- Shafer, Robert. 1974. *An introduction to Sino-Tibetan*. Wiesbaden: Harrassowitz.
- Sprigg, R.K. 1955. The tonal system of Tibetan (Lhasa dialect) and the nominal phrase. *Bulletin of the School of Oriental and African Studies* 17(1): 133-153. Reprinted in Palmer (ed.), *Prosodic Analysis*, pp. 112-132. London: Oxford University Press.
- Sprigg, R.K. 1966. Lepcha and Balti Tibetan: tonal or non-tonal languages? *Asia Major* 12: 185-201.
- Sprigg, R.K. 1972. A polysystemic approach, in Proto-Tibetan reconstruction, to tone and syllable-initial consonant clusters. *BSOAS* 35: 546-587.
- Sprigg, R.K. 2002. *Balti-English English-Balti Dictionary*. New York: RoutledgeCurzon.
- Sun, Jackson T.-S. 1986. *Aspects of the phonology of Amdo Tibetan: Ndzorge Śæme Xõra dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Sun, Jackson T.-S. 1997. The typology of tone in Tibetan. In *Chinese Languages and Linguistics IV: Typological studies of languages in China*. Symposium Series of the Institute of History and Philology, Academia Sinica, Number 2. Taipei: Academia Sinica.

- Sun, Jackson T.-S. 2001. Variegated tonal developments in Tibetan. Paper presented at the 34th International Conference on Sino-Tibetan languages and linguistics. Kunming, October 2001.
- Sun, Jackson T.-S. 2003. Phonological profile of Zhongu: a new Tibetan dialect of northern Sichuan. *Language and linguistics* 4(4): 769-836. Taipei: Academia Sinica.
- Takata, Tokio. 1981. Sur la naissance des tons du Tibétain. *Journal Asiatique* 269(1,2): 277-285.
- Thup bstan tbang po, with Bsod stobs and Lu'o ping hphen. 2002. *Lha sa'i kha skad sbyong deb*. [Spoken Lhasa Tibetan textbook]. Lhasa: bod ljongs mi dmangs dpe skrun khang.
- Tournadre, Nicolas. 2005. L'aire linguistique tibétaine et ses divers dialectes. *Lalies* n°25. Presse de l'école normale supérieure.
- Tournadre, Nicolas. 2008. Arguments against the concept 'Conjunct' / 'Disjunct' in Tibetan. In Brigitte Huber, Marianne Volkart and Paul Widmer (eds.), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag. Band I: Chomolangma*, pp. 281-308. Halle (Saale): International Institute for Tibetan and Buddhist Studies (IITBS) GmbH.
- Volkart, Marianne. 2003. Types of compounds in Written Tibetan. In Tej Ratna Kansakar and Mark Turin (eds.), *Themes in Himalayan languages and linguistics*, pp. 233-248. Heidelberg: South Asia Institute, and Kathmandu: Tribhuvan University.
- Watters, Stephen A. 2002. The sounds and tones of five Tibetan languages of the Himalayan region. *Languages of the Tibeto-Burman Area* 25(1): 1-65.
- Wylie, Turrell V. 1957. A standard system of Tibetan transcription. *Harvard Journal of Asiatic studies* 22: 261-267.
- Yang Ho-Chin. 1974. The Tibetan alphabet and its tonal system in the Lhasa dialect. *Bulletin of the Institute of China Border Area Studies* 5: 19-27.

APPENDIX – WORD LIST

The list on the following pages provides the Written Tibetan and phonetic forms of words recorded from the four speakers considered in this study: two speakers of Balti Tibetan (BSh_03 and BM_01), and two speakers of Rebkong Amdo Tibetan (AR_04 and AR_05). Only those words and tokens which could be analyzed acoustically are included. There were hundreds of other forms recorded that could not be analyzed because they could not be reliably segmented (as I discuss in Chapter 2). These words are not on this list.

The Written Tibetan (WT) forms shown in this table have been drawn from a number of sources, which are not always in complete accord. In some cases, there was a very obvious standard form. In other cases I chose a spelling alternant which in my view most closely matches the phonetic form. More knowledgeable scholars may not always agree with my choices. I provide different spellings for the Balti and Rebkong Amdo spoken forms – separated by a semi-colon “;” – when they evidently correspond to different historical roots. A slash “/” indicates that there are two equally plausible WT forms.

The sources consulted include:

Goldstein, 1984
Goldstein, 2001
Jaeshke, [1881] 1958
Norberg-Hodge and Gyelong Thupstan Paldan, 1991
Sprigg, 2002
Roerich, 1958

Some spellings were also kindly provided to me by Roland Bielmeier, by Nicolas Tournadre (often referring to the *tsig dzo chen mo*, a Tibetan-Chinese source which I am not able to read for myself), and by my research assistant Sangye Gyatso. In some cases I was not able to confirm their suggestions in the sources available to me; I have marked these forms as follows:

- * Reconstructed Proto-Tibetan form provided by Roland Bielmeier (p.c. 2008).
- † Written Tibetan form or etymology suggested by Roland Bielmeier (p.c. 2008).
- †† Written Tibetan form suggested by Sangye Gyatso (p.c. 2004).
- # Written Tibetan form provided by Nicolas Tournadre (p.c. 2008).

There were a few instances in which – despite these thoughtful suggestions – I chose to use a WT form from a different source. For instance, some Balti words seemed to me to be best represented by the forms offered in Norberg-Hodge and Paldan’s dictionary of Ladakhi.

The phonetic transcriptions shown in this word list are based on elicitation and repetition in the field, double-checking of the recordings, and close examination of waveforms and spectrograms. Working with only two speakers of each dialect, and with a limited vocabulary, it is not possible to draw conclusions about underlying phonological forms, or even to know what a “standard” pronunciation of a word might be in a particular dialect. Each speaker has their own idiosyncrasies, and these are captured in the transcription. For instance, speaker BM_01 produced the form for the verb ‘to close a door’ (# 585) as [ʼzgo.tɕuk^h], with final aspiration. Final aspiration does not occur phonologically in Balti, but it clearly occurred phonetically in this instance. Because the acoustic measurements in this study were based on these particular tokens, such phonetic variations are included in the transcription.

Question marks are used in various ways to indicate places where I am not confident that I have identified a WT form which corresponds to the phonetic form(s):

- # 267 ‘frog’ ??
Two question marks (“??”) indicates that I was not able to identify a possible WT correlate for either syllable of the word, and that the etymology is unknown.
- # 18 ‘harvest’ *thog.btsas* ?
A question mark followed by a space after a full WT form indicates that this seemed like the most logical Written Tibetan correlate, but I am not completely confident of this.
- # 76 ‘lid, cover’ *kha.thum?* (for Rebkong Amdo [k^ha.'tu])
A question mark following either syllable – without a space – indicates that I am confident of the WT form for one syllable, but not the other.
- # 205 ‘sweat, perspiration’ *rngul.?* (for Rebkong Amdo [ʼŋuv.'tsʌ])
A question mark in place of either σ_1 or σ_2 (before or after the “.” at the syllable boundary) indicates that I was not able to identify a WT correlate for that syllable.

WORD LIST – CAPLOW 2009

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
8	N	pearl	mu.tig (<Sanskrit †)	mu.'tik	mu.'tik					
10	N	school	'school'; slob.grwa	su.'kul		su.'kul	ʃop.'ʃa	ʃop.'ʃa	ʃov.'ta	ʃop.'ta
14	N	horse saddle	rta.sga	ʃtaʌz.'ga	ʃtaʌz.'ga		xtar.'ka	xtar.'ka	xtar.'ga	tar.'ga
15	N	horse race	rta.rgyugs				ʃtar.'tɕix	ʃtar.'tɕi		
16	N	stable	rta.khang; rta.rwa	ʃta.'kʰaŋ	ʃta.'kʰã		xta.'ra	xta.'ra	xta.'ra	ta.'ra
17	N	horse year	rta.lo	ʃta.'lo	ʃta.'lo		xta.'lo	xta.'lo	xta.'lo	ta.'lo
18	N	harvest	thog.btsas ?				to.'tsiɯ			
19	N	astrological year	lo.ril				lo.'ri	lo.'ri	lo.'ri	lo.'ri
21	N	new year	lo.gsar	lo.'sar	lo.'sar				ʒo.'siŋ	
22	N	horse whip	rta.mthur; rta.lcag	ʃta.'tʰur			xtax.'tɕɤɣ			
25	N	blacksmith	mgar.ba	gar.'ba	gar.'ba	ŋgar.'va	ŋka.'ra	ŋga.'ra	ŋga.'ra	ŋga.'ra
27	N	arrow and bow	mda'.gzhu	ndaɤ.'zu			nta.'zɤ	nta.'zɤ	nda.'zɤ	nda.'zɤ
29	N	star	skar.ma	skar.'ma	skar.'ma		ɤxkar.'ma	ɤxkar.'ma	kar.'ma	kar.'ma
30	N	shooting star, meteor	skar.zla	skar.'da	skar.'da		xkan.'ta	xkan.'ta	xkarn.'da	kan.'da
33	N	gun, rifle	me.mda'	mɛ.'ða	mɛ.'ða					
35	N	flame	me.rlabs	mɛ.'lɰap	mɛ.'lɰav					

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
36	N	disease, illness	na.tsha						na.'tsa	na.'tsa
39	N	sick person, patient	nad.pa	nʌt.'pa	nʌt.'pa		nai.'pa	na.'pa	na.'pa	na.'pa
42	N	pig	phag.pa						hʌʔ.'ʒɣ	
43	N	pork	phag.sha		p ^h aq.'ʃa		həʒ.'xa	həʒ.'xa	hʌʒ.'xa	
44	N	house	khang.pa						k ^x o.'ŋa	k ^x o.'ŋa
45	N	restaurant	za.khang	za.'k ^h aŋ	za.'k ^h ã:					
47	N	land, place, property	sa.cha				s ^h a.'tɕa	s ^h a.'tɕa	s ^h a.'tɕa	s ^h a.'tɕa
49	N	chicken, hen	mo.byā				mou.'ʃa	mou.'ʃa	mou.'ɕa	mou.'ɕa
50	N	rooster, cock	pho.byā				p ^h ou.'ɕa	p ^h ou.'ɕa	hou.'ɕa	hou.'ɕa
51	N	eagle	bya.glag ? †			ʔm.'ðak ^x				
54	N	crow, raven	pho.rog				ho.'roʔ	ho.'roʔ		
56	N	snuff	sna.tho? sna.rdul? †† sna.tha?				na.'to	na.'to	ʔna.'do	na.'do
58	N	hand, arm	lag.pa	laq.'pa	laq.'pa	laʒ.'pa				
64	N	foot, leg	rkang.ma	kaŋ.'ma	kã.'ma					
67	N	on foot	rkang.thang						xkʌŋ.'t ^h ʌŋ	kʌŋ.'t ^h ʌŋ
72	N	boss, head man	bdag.po	ndaʒ.'po	daʒ.'po		ʔtʌq.'qu	tʌq.'qo	ʔdo.'ɕo	do.'ɕo
76	N	lid, cover	kha.thum?				k ^h a.'tu	k ^x a.'tu	k ^h a.'to	k ^h a.'to

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
77	N	hair on head	mgo.ral			ŋgo.'rɛl				
80	N	obstacle	bar.chad	bar.'tʃad	bar.'tɕad				var.'tɕʰɪl	var.'tɕʰɪl
81	N	Buddha	sangs.rgyas						sʰaŋ.'dʒi	
82	N	axe	sta.re; sta.ri	sta.'re	sta.'re		xta.'ri	xta.'ri	xta.'ri	ta.'ri
86	N	prayer, blessing	smon.lam				mɕ.'ʒi	mɕ.'ʒim		
89	N	snow-covered peak	gangs.ri	ŋgã.'ri	gã.'ri					
90	N	hibivore, ibex	ri.dwags	ri.'daɣ	ri.'dax					
95	N	work	las.ka				lex.'ka	le.'ka	lex.'ka	le.'ka
96	N	knife sheath	gri.? gri.shubs	dʒi.'ʃuk			tɕɤ.'xʰiv		tɕi.'ɕiv	tɕi.'ɕiv
98	N	field	sa.khyad	sa.'kʰet	sa.'kʰɛd		ɕʰɬŋ.'kʰa	ɕɬŋ.'kʰa		
99	N	field work, agriculture	zhing.las				ɕaŋ.'li	ɕaŋ.'li		
100	N	forest, jungle	'jungle'			dʒɬŋ.'gɐl				
101	N	tree	sdong.bu						ʔdõ.'ŋu	dõ.'ŋu
102	N	wooden yoke	ske.shing						ce.'xuŋ	ce.'xuŋ
104	N	willow tree	lcang.ma			ʔtsay.'ma	xtɕãi.'ma	ɕtɕãi.'ma	xtɕɬŋ.'ma	xtɕɬŋ.'ma
105	N	alpine willow tree	glang.ma	ɣlãŋ.'ma	xlãi.'ma		ʰɔʔɕõi.'ma	ʒõi.'ma	ʰɔʔɬŋ.'ma	lõi.'ma
106	N	birch tree	stag.pa	staq.'pa	staq.'pa					
107	N	walnut	star.kha?, star.ka?				xtar.'ka	xtar.'ka	xtär.'ga	tar.'ga

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
108	N	fruit	shing.tog				s ^h i.'toɤ			
109	N	flower	me.tog		min.dɔ:		me.'toɤ			
112	N	salt	tshwa.khu				ts ^h a.'kɣ ts ^h a.'kɣ	ts ^h a.'kɣɣ	ts ^h a.'k ^h ɣ	
115	N	thermos bottle	'thermos'			t ^ə ar.'muɕ	ts ^h a.'tim ts ^h a.tĩ	ts ^h a.'dim	ts ^h a.'dim	
117	N	tea churn	ja.mdong	tɕa.'doŋ		tɕ ^h a.'ðõ:				
118	N	tea w/ tsampa	dul.ma ? ††				ɕtuu.'ma xtuu.'ma	ɕtu.'ma	t ^h u.'ma	
121	N	big river	rgya.mtsho			ʔrgyam.'so				
129	N	chest (body part)	brang.kha				tɕŋ.'k ^x a tɕŋ.'k ^x a	tɕɕŋ.'ka		
130	N	dusty wind	rlung.tshub / rlung.'tshub				ʔlõ:.'ts ^h ɣu lõ:.'ts ^h ɣu	ʔlõ:.'ts ^h ɣu	lɔŋ.'ts ^h ɣv	
133	N	cave, rock shelter	brag.khog?, brag.khung		bla.'q ^h oq				tɕɕ.'kuŋ tɕɕ.'kuĩ	
134	N	cow (female); cattle	mo.bang, ba.lang; ba.mo	mo.'baŋ mo.'βã:	mba.'laŋ		ua.'mo	va.'mo	va.'mo	
135	N	cow (male), bull, ox	pho.bang	p ^h o.'βaŋ					ʔla.'ŋɣ la.'ŋɣ	
136	N	elephant	glang.chen	ɬlan.'tɕ ^h ɛn						
138	N	family members	nang.mi			naŋ.'mi				
139	N	household, family	khyim. tshang				tɕ ^h ɣn. tɕ ^h ɣn. 'ts ^h ɣuɰ 'ts ^h ɔ	tɕ ^h ɛm. tɕ ^h ɛn. 'ts ^h ɕŋ 'ts ^h ɕŋ		

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
145	N	nomad, pastoralist	'brog.pa	mbloq.'pa	bloq.'pa	broq.'pa				
146	N	nomad area	'brog.sa				ɖok.'s ^h a		ndoɣ.'s ^h a	ndoɣ.'s ^h a
148	N	rope	thag.pa	t ^h aq.'pa						
149	N	chain	lcags.thag						xtɕɤɤ.'t ^h ɤɤ	tɕɤɤ.'t ^h ɤɤ
153	N	pasture, grazing area	rtswa.thang				xtsa.'t ^h ʊŋ	ɕtsa.'t ^h ʊ	xtsa.'t ^h ʌŋ	
155	N	medicinal plant	rtswa.sman						xtsa.'hmin	xtsar.'mĩ:
156	N	hospital	sman.khang	sman.'k ^h aŋ	sman.'k ^h ã:		hmãĩ.'kxɤŋ	mãĩ.'kxɤ		
159	N	birthplace	skye.sa				ɕtɕe.'sa	ɕtɕe.'sa	xeye.'sa	tɕe.'sa
161	N	dialect	skad.?				xka.'tuɤ	xka.'tɤx		
166	N	mandala	dkyil.'khor						xtɕiŋ.'k ^x ur	tɕiŋ.'k ^x ur
168	N	goddess; Tibetan opera	lha.mo			ɬla.'ŋo	ɬaa.'mu	ɬaa.'mu	ɬa.'mo	ɬa.'mo
169	N	temple	lha.khang	ɕla.'k ^h aŋ	ɕa.'kã:	ɬla.'k ^x aŋ				
173	N	war	dmag.'khrug				^{hə} m ^ə ɤ.'t ^h ix			
176	N	footpath, one- person trail	rkang.lam	kã:.'lam	kã:.'lam				xkʌŋ.'lam	xkʌŋ.'lam
177	N	highway	rgya.lam				^{hə} dʒya.'lɤm			
181	N	cat	byi.la (<Middle Indo- Aryan?†); lu.li		bi.'la	bi.'la			lu.'li	lu.'li

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
183	N	heavy thing (?)	ljid.mo				ʒi.'mo	ʒi.'mo		
187	N	merit, luck	bsod.nams				so.'nim		s ^h ʕ.'nim	s ^h o.'nim
188	N	prison	btson.khang						tsɔiŋ.'k ^x ʌŋ	
189	N	prisoner	btson.ma				tsɔi.'ma	tsɔi.'ma	tsɔiŋ.'ma	tsɔi.'ma
190	N	chorten	mchod.rten				tɕ ^h ou.'tẽ	tɕ ^h ou.'tẽ	tɕ ^h or.'tɪn	tɕ ^h or.'tɪn
200	N	sacred pile of 100,000 stones	rdo.'bum				h ^a tim.'bɣm	ədim.'bɣ	h ^a dim.'bɣm	dim.'bɣm
205	N	sweat, perspiration	rngul.chu; rngul.?		ɣmul.'tɕ ^h u		ʔŋuʊ.'tsɣ	ŋu.'tsɣ		
209	N	local language, father tongue	pha.skad						p ^h ar.'kiɭ	
210	N	bone	rus.pa	ʔrus.'pa	ʔrus.'pa	ʔrus.'pa	ʔrɣ.'pa	rɣ.'pa	zɹ.'pa	
211	N	bald head; animal w/o horns	mgo.rdo				ŋgor.'tu	ŋgor.'tu	ŋgor.'do	ŋgor.'do
213	N	lamp	*dkar.lto				kar.'tu	kar.'tu		
221	N	water source, spring	chu.mig				tɕ ^h ʌŋ.'ku	tɕ ^h ʌŋ.'ku	tɕ ^h ɣŋ.'go	tɕ ^h ɣŋ.'go
226	N	eyeglasses	mig.srab			miɣ.'sɽap				
227	N	key	lde.mig	le.'mik	le.'mig					
230	N	goats and sheep	ra.lug						h ^a ra.luɣ	ra.'luɣ

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
231	N	baby sheep	lug.phrug						ʕʕ.'ʕʕ	lʕ.'ʕʕ
235	N	body.HON	sku.lus				kʕ.'ʕi	kʕ.'ʕi	xkʕ.'li	xkʕ.'li
237	N	incarnation	sprul.sku				ʕtʕʕ.'kʕ		xtiʕ.'kʕ	xtiʕ.'kʕ
241	N	nose blood, bloody nose	sna.khrag				hna.'tʕʕʕ			
242	N	border	mu.rdo; sa.mtshams			smor.'ðo	s ^h an.'tsim		s ^h an.'tsim	s ^h an.'ts ^h im
243	N	shopping, things to be bought	nyo.cha, *nyo.bca'				nyou.'tʕ ^h a	nyou.'tʕ ^h a		
249	N	grape	rgun.'brum				ʔgʕn.'dʕʕm	gʕn.'dʕʕm	^h ʔgʕn.'dʕʕm	gʕn.'dʕʕ
253	N	khata, offering scarf	kha.btags						k ^h a.'tʕʕʕ	k ^h a.'tʕʕʕ
256	N	nit, lice egg	sro.ma						ʕo.'maŋ	ʕo.'maŋ
257	N	wool felt	phying.pa			p ^h uŋ.'ma				
259	N	rain	char.pa; gnam.chu			tʕ ^h ar.'pa	nam.'tʕ ^h i		hnam.'tʕhu	nam.'tʕ ^h u
261	N	dust	sa.rdul				s ^h ar.'tuv		s ^h ar.'du	s ^h ar.'du
262	N	bride	bag.mo, mna'.ma	ʕaʕ.'mo	baʕ.'mo	na.'ma	na.'ma	na.'ma	nã.'ma	na.'ma
263	N	bridegroom, son-in-law	bag.po, mag.pa	ʕaʕ.'p ^h o	baʕ.'po	mak.'pa				
267	N	frog	??			ʕa.'tek				
268	N	lama	bla.ma	la.'ma	la.'ma	la.'ma	ʔla.'ma	la.'ma	ʔla.'ma	la.'ma

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
273	N	bridge	zam.ba	sam.'ba		nzam.'ba	ɕsam.'ba	ɕsam.'ba	xsam.'ba	sam.'ba
274	N	wolf	spyang.ku	spyaŋ.'ku		ɕaŋ.'ku	ɕtɕaŋ.'kxɣ	ɕtɕaŋ.'kʰɣ		
278	N	leaf	lo.ma				lo.'ma	lo.'ma	lo.'ma	lo.'ma
279	N	brain	klad.pa / glad.pa	xlat.'pa		ɣlat.'pa	ʰlai.'pa	ɣa.'pa	ʰlai.'pa	lai.'pa
282	N	finger nail, toenail	ser.mo, sen.mo			ʰzer.'moŋ			sʰe.'mo	sʰe.'mo
283	N	show, spectacle	ltad.mo				ɕtai.'mo	ctai.'mo	xtai.'mo	xtai.'mo
285	N	cream	ʰo.spri			ʰo.'spis				
286	N	smell, scent, odor	dri.ma				tʃo.'ma	tʃo.'ma	tʃi.'ma	tʃi.'ma
287	N	liver	mchin.pa			tɕʰin.'ma			tɕʰin.'ba	tɕʰin.'ba
291	N	animal droppings (pellets)	ril.ma						ʰzi.'ma	ri.'ma
292	N	drawing, picture	ri.mo				ʰri.'mō	ri.'mo		
293	N	round pill	ril.bu				ʰrx.'lx	rx.'lx		
295	N	knee (kneecap ?)	pus.mgo				ʰvuŋ.'go			
298	N	milk	ʰo.ma						o.'ma	ʰo.'ma
299	N	forehead, skull	dpral.ba; thod.pa	spal.'ba	spal.'ba	ʰspal.'ba	tʰo.'pa	tʰo.'pa	tʰo.'pa	tʰo.'pa
302	N	sickle	zor.ba	ɤzor.'ba	ɤzor.'ba		xso.'ra	so.'ra	xso.'ra	xso.'ra

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
303	N	official, chief	dpon.po						χɔim.'bo	χɔim.'bo
304	N	charcoal, coal	sol.ba; *sol.gas	sol.'ba	χsol.'ba		s ^h oŋ.'kxe	s ^h oŋ.'kxe		
305	N	evening	dgong.dag dgong.phyi; dgong.mo	gøŋ.'t ^h aχs		ŋgɔ̃:.'ɕim	^h a'kɔ̃:.'mu	kɔ̃:.'mu		
306	N	female dzo	mdzo.mo				dzo.'mu	dzo.'mu	ndzo.'mu	ndzo.'mu
307	N	guest	mgron.po	ɖrøn.'pa		ŋgrøn.'pa				
309	N	plan	char.gzhi						tɕ ^h ar.'zɣ	tɕ ^h ar.'zɣ
310	N	queen	rgyal.mo				ntɕai.'mo	tɕai.'mo		
315	N	smoke	dud.ba	tut.'pa	tut.'pa					
319	N	thread	skud.pa				xkɣ.'pa	xkɣ.'pa	xkɣ.'pa	kɣ.'pa
321	N	thief	rkun.ma			ʂkøn.'ma				
324	N	girl, daughter	bu.mo; sras.mo			ɸo.'ŋo	ʂi.'mo		ʂɣ.'mo	ʂɣ.'mo
327	N	donkey	bong.bu	mbɔ̃:.'bu	mbɔ̃:.'bu	ɸɔ̃ŋ.'bu				
328	N	digit (finger, toe)	sen.mo; mdzug.gu	sen.'mo	sen.'mo	sen.'ma	ntsɣ.'ɣɣ	ndzɣ.'ɰɣ	ndzɣ.'ɣɣ	ndzɣ.'ɣɣ
329	N	paper, book	shog.shog, shog.bu	ɕoq.'ɕoq	ɕoɕ.'ɕoɕ	ɕoɣ.'ɕu				
331	N	bellows	sbud.pa			zbuk.'pa				
334	N	tsampa	ja.phye, rtsam.pa	tɕa.'p ^h e	tɕa.'p ^h e		xtsam.'ba	xtsam.'pa	xtsam.'ba	xtsam.'ba
337	N	faith, practice	dad.pa						tai.'pa	tai.'pa

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
342	N	wolf (jackal?)	'phar.ba				p ^h a.'ra	p ^h a.'ra	p ^h a.'ra	p ^h a.'ra
343	N	matted hair, dreadlocks	ral.?				ra.'lɣ	ra.'lɣ	ʔra.'lɣ	ra.'lɣ
344	N	pea, bean, legume	sran.ma	stɾen.'ma	stɾen.'ma		ʂai.'ma	ʂai.'ma	ʂ ^h ai.'ma	ʂ ^h ai.'ma
365	N	length	ring.thung						h ^ə raŋ.'t ^h uŋ	raŋ.'t ^h uŋ
392	N	beggar, mendicant, poor person	ltsangs. mkhan ?, bslangs. mkhan †			ʈtsan.'k ^h an				
400	N	expert	shes.mkhan			ɕes.'k ^x an				
499	N	clothing	gon.chas	ŋgøn.'tɕes	ŋgøn.'tɕes					
519	N	white prayer flag	dar.lcog				taɤ.'tɕɣɤ			
520	N	water drop, leak	thig.chu	t ^h ik.'tɕ ^h u	t ^h ik.'tɕ ^h u		t ^h ɣɣ.'tɕ ^h ɣ			
522	N	medicinal plant (local term)	sman.rtswa				hmãi.'tsa	hmãi.'tsa	hmain.'tsa	main.'tsa
526	N	personal deity	gzhis.lha						xtɕ ^h i.'ʂa	xtɕ ^h i.'ʂa
543	N	bell (?) 'small bells worn around ankle of dancers' †	?.dril, dril.?			tʃiv.'ɕil				
544	N	stomach	grod.pa			krot.'pa				

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
555	N	bottle	'bottle'			mbʌ.'tʃɪl				
561	N	breast	<Burushaski †			tɕu.'tɕu				
576	N	carpet	cha.ra #			tɕ ^h a.'ra				
586	N	cloud	mun.ma			mun.'ma				
588	N	cold, flu	chams.pa			tɕ ^h am.'ba				
615	N	diamond	pha.lam			p ^h a.'lam				
616	N	dice	cho.lo			tɕ ^h o.'lo				
623	N	doctor	sman.pa	smʌn.'pa		sman.'pa				
625	N	wild donkey	thang.bong			t ^h aũŋ.'boŋ				
641	N	edge, limit	mtha'.ma			t ^h a.'ma				
642	N	egg	bya.bzhon †			byaβ.'zøn				
645	N	eweskin coat	thu.lu †			t ^h u.'lu				
655	N	film	'film'			pɸi.'lim				
663	N	foreigner	??			mɛr.'man				
667	N	friend	dga'.mkhan			ʔrɣa.'χan				
669	N	juniper	shug.pa			ɕuk.'pa				
686	N	hail	??			k ^h ar.'dzɪm				
690	N	hat	??			na.'tiŋ				
717	N	last year	na.ning			na.'niŋ				

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
764	N	morning	'gyogs.pa ?, mgyogs.pa ?†			gyoχ.'pa				
766	N	mouth	kha.'khor			k ^h a.'k ^h oɾ				
767	N	mule	<Urdu, Hindi †			k ^h a.'çãɾ				
876	N	teacher	'teacher'	ti.'tçəɾ	ti.'tçə					
922	N	woman, wife	bu.sring	bu.'striŋ	mbu.'strĩ:					
927	N	wound; scab	rma.ra ? sma.kha ?			sma.'ra				
963	N	bread	'k ^h ur.ba	k ^h ur.'βa						
968	N	eyebrows	smin.ma	smɪn.'ma						
971	N	archery festival	mda'.'phang / mda'.'phangs	ða.'p ^h aŋ	da.'p ^h aŋ					
977	N	goat (alt. form)	ra.ma	ra.'ma	ra.'ma					
978	N	conversation, talk	dpe.sgra	χpe.'ra	xpe.'ra					
984	N	forearm	phrag.pa	pɾak.'pa	p ^h ɾak.'pa					
986	N	standing tree	sta.?, stag.?	stax.'zi	stax.'zi					
990	N	wheat flour	bag.phye	βax.'p ^h e	mbax.'p ^h e					
991	N	black pea	nag.sran	naq.'stren	naq.'stren					
1002	N	twelve year cycle	lo.skor	lo.'skoɾ	lo.'skor					
1003	N	first year	lo.'go	lom.'go	lom.'go					

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
1007	N	restaurant, hotel	'hotel'	ho.'tʃil						
1010	N	socks	rkang.rtse?	kãi.'tse						
1016	N	cairn, pile of stones	rdo.byur †	ʔrdo.'pyur	ʔrdo.'pyur					
1022	N	polo accessory for horse	??	χmit.'paχ						
1027	N	mare, female horse	rgod.ma	ʔrgud.'mo	rgud.'mo					
1028	N	older person, mature	tshar.ma †	ts ^h ar.'ma	ts ^h ar.'ma					
1031	N	shame	khrel.ba	t ^h r̥ɛl.'ba	t ^h r̥ɛl.'ba					
1042	N	woman's hair	skra.lo			ʃka.'lo				
1043	N	bride (alt. form)	bag.mo			bar.'mo				
1044	N	brain, mind	mgo.glad			ŋgo.'rlat				
1045	N	crow, rave (alt. form)	khwa.ta				kxa.'ta	kxa.'ta		
1046	N	waist	sked.pa			sket.'pa				
1047	N	barley flour, tsampa	nas.phye			nas.'pʰe				
296	A	thick (like rope)	sbom.po	bom.'bo	bom.'bo	bom.'bo				
348	A	black	nag.po	naq.'po	naq.'po				no.'χo	no.'χo
349	A	white	dkar.po	kar.'p ^h o	kar.'p ^h o	kar.'po	ka.'ru	ka.'ru	ka.'ro	ka.'ro

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
350	A	yellow	ser.po			s ^h er.'po	s ^h e.'ru	s ^h e.'ru		
351	A	red	dmar.po	mar.'p ^h o						
352	A	blue color	sngo.rang (rang <Urdu <Persian †)			sŋo.'raŋ				
363	A	long	ring.po				ʔra.'ŋo			
371	A	short, few	chad.chad			tɕ ^h at.'tɕ ^h at				
372	A	flat	leb.leb						ʒev.'lev	lev.'lev
382	A	important	tsha.tsha ?			s ^h a.'s ^h a				
384	A	difficult	mkhregs.pa †		t ^h ɾis.'pa					
385	A	easy	bde.mo			ʔvdo.'ŋo				
387	A	clean	gtsang.ma			ʈtsaŋ.'ma			xtsaŋ.'ma	xtsaŋ.'ma
388	A	new	gsar.pa; so.ma	sar.'p ^h a	sar.'p ^h a	sar.'pa	s ^h o.'ma	s ^h o.'ma	s ^h o.'ma	s ^h o.'ma
393	A	steep	gzar.mo				ʔzar.'mō	zar.'mo	^h zar.'mō	zar.'mō
395	A	dry, dried	skam.po			skam.'bo	xkam.'bo	xkam.'bo		
396	A	hot	tsha.'di				ts ^h an.'di	ts ^h an.'di	ts ^h an.'di	ts ^h an.'di
403	A	alone, single, solo	kher.po †				k ^h ə.'ro		k ^x ɣr.'mō	k ^h ɣr.'mō
404	A	same, alike	tsogs.tsogs			tsors.'tsor				
408	A	high	mthon.mo	t ^h on.'mo		t ^h ɣn.'mo				
409	A	rich	phyug.po				ɕ ^h ɣ.'ko	ɕ ^h ɣ.'ko	ɕ ^h ɿ.'ko	ɕ ^h ɿ.'ko

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
410	A	thick	mthug.mthug			ʈuk.ʈuk				
411	A	rare	dkon.po ?	ʃkøn.'mo	ʃkøn.'mo		χkɔim.'bu	xkɔim.'bu	xkɔim.'bo	kɔim.'bo
412	A	big	chen.po				tɕ ^h e.po	tɕ ^h e.po	tɕ ^h e.po	tɕ ^h e.po
413	A	small_little	chung.tshad?; chung.chung			tɕ ^h un.'tse			tɕ ^h ɔŋ.'tɕ ^h ɔŋ	tɕ ^h ɔŋ.'tɕ ^h ɔŋ
415	A	fat	?? ; tshon.po			ʈ ^h äk.ʈ ^h äk	ts ^h ɔim.'bu	ts ^h ɔim.'bu		
529	A	all	gang.ma			ŋgaŋ.'ma				
581	A	cheap, inexpensive	<Urdu			sas.'ta				
584	A	clear	gsal.po			xsal.'po				
587	A	cold	grang.mo			dʒaR.'mo				
619	A	dirty	<Burushaski †			tɕ ^h a.'ru				
651	A	fast, quick	mgyogs.mo ?, *shogs.mo			ɕOR.'mo				
677	A	good	legs.mo			lear.'mo				
678	A	tasty, delicious	zhim.bo			ʔzum.'bo				
685	A	blue, green	sngon.po			sŋɔn.'po				
740	A	low	*babs.mo			ʃah.'mo				
747	A	many	mang.mo	mã:.'mo	mã:.'mo	mãũ.'mo				
422	Num	3rd	gsum.pa						ksum.'ba	
426	Num	7th	bdun.pa						ʔdʏn.'ba	

No.	Cat	Gloss	WT	BSh 03		BM 01	AR 04		AR 05	
				isol	frame	isol	isol	frame	isol	frame
432	Num	12	bcu.gnyis	tɕo.'ɲas						
433	Num	13	bcu.gsum	tɕuɣ.'sum						
435	Num	14	bcu.bzhi	tɕuβ.'zi						
436	Num	15	bco.lnga						tɕo.'ɲa	
437	Num	16	bcu.drug	tɕu.'ruk					tɕɤ.'ɲix	
438	Num	17	bcu.bdun	tɕub.'dun					tɕɤb.'dɤn	
439	Num	18	bco.brgyad	tɕop.'gyet						
440	Num	19	bcu.dgu	tɕur.'gu					tɕɤr.'gɤ	
443	Num	30	gsum.bcu	xsum.'tɕu					s ^h ʌm.'tɕɤ	
447	Num	40	bzhi.bcu						ʒɤb.'tɕɤ	
449	Num	50	lnga.bcu	ʔkaf.'tɕu					ɲap.'tɕɤ	
453	Num	70	bdun.cu						h ^o dɤn.'tɕɤ	
457	Num	90	dgu.bcu						kɤb.'tɕɤ	
28	V	to shoot an arrow	mda'.'phen				'nta.hen		'nda.hen	'nda.hɛi
32	V	to make a fire	spar.ba			'spar.va				
34	V	to shoot a gun	wu.brgyab †, wos.brgyab < Chinese? #						'vu.dzəp	'vu.dzɪp
88	V	to snow	gangs.babs ?, kha.ba.babs ? †						'k ^x ʌɲ.bʌp	'k ^x ʌɲ.bɤu
260	V	to rain	gnam.babs				h ^o 'nʌm.bʌu	'nʌm.biʊ	h ^o 'nʌm.bʌp	'nʌm.bau

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
333	V	to answer	lan.brgyab				'ʒẽ.ɕɣʋ	'ʒẽ.dʒɣʋ	'len.tɣv	'len.tʰɣv
532	V	to arrive	'thon.ma	'tʰɣn.ma	'tʰɣn.ma					
570	V	to call, to shout	skad.zer			'skʰat.zerɕ				
585	V	to close a door	sgo.bcug †			'zgo.tɕukʰ				
597	V	to cook food	zan.bcos †			'zan.ɸtɕos				
640	V	to eat food	zan.za	'za:n.za	'za:n.za	'zan.za				
654	V	to be afraid, to fear	'jigs.pa	'ʒik.pʰa	'ʒik.pʰa	'dʒik.pʰa				
660	V	to fly	'phur.ba			'pʰur.ba				
671	V	to give (by hand)	sbyin.ma ? †	'min.ma	'min.ma	'min.ma				
711	V	to keep, to protect	srung.ma			'stun.ma				
715	V	to know	shes.pa			'ɕes.pa				
718	V	to be late	'gor.ba			'ŋgor.va				
733	V	to listen	rna.byā			'sna.byā				
746	V	to spread manure	lud.btang			'lut.taŋ				
755	V	to meet	thug.pa			'tʰuk.pə				
814	V	to ride	zhon.ma ? bzhon.ma ?	'ʒun.ma	'dʒun.ma					
816	V	to rise (sun)	shar.ba	'ʃar.ba	'ʃar.va					

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
856	V	to get up, to rise	lang.ma	'laŋ.ma	'laĩ.ma					
870	V	to carry	'khur.ba	'k ^h ur.ba	'k ^h ur.βa					
957	V	to stretch	brkyang.ma †, rkyangs.ma ?	'škyaŋ.ma	'škyaŋ.ma					
958	V	to release, to set X on to	??	'škat.pa						
960	V	to fill	bkang.ma ?	'skaŋ.ma	'skāĩ.ma					
961	V	to clothe	skon.ma	'skøn.ma	'skøn.ma					
964	V	to be tired	*glad.pa	'χlat.pa	'xlat.pa					
967	V	to grind	'thag.pa	't ^h aq.pa	't ^h aq.pa					
970	V	to light, to ignite	spar.ba	'spar.ba	'spar.ba					
974	V	to walk	'grul.ba	'ndrul.ba	'drœl.ba					
983	V	to harm	gnod.pa	'χnot.pa	'χnot.pa					
985	V	to be clean, pure	dag.pa	'daq.pa	'dak.pa					
987	V	to assign a name	btag.pa	'taq.pa	'taq.pa					
992	V	to straighten	srang.ma	'strãŋ.ma	'strãĩ.ma					
993	V	to destroy, to raze	??	'çte.va	'xte.βa					
996	V	to ride (var)	??	'dzɹon.ma	'dzɹon.ma					
997	V	to laugh, to get wild	rgod.pa	ʳrgot.pa	'rgot.pa					

No.	Cat	Gloss	WT	BSh_03		BM_01	AR_04		AR_05	
				isol	frame	isol	isol	frame	isol	frame
998	V	to survive, to be nourished	gson.ma	'χsøn.ma	'xsøn.ma					
1000	V	to be ripe	smin.ma	'smin.ma	'smin.ma					
1013	V	to drip	'dzar.ba † / gzar.ba † / bzar.ba †	ʷɛzar.pa	'ɛzar.ba					
1014	V	to deprive	sgag.pa / bsgags.pa	ʷzgak.pa	ʷzgak.pa					
1018	V	to leak, to drip	thig.pa	'tʰik.pa						
1021	V	to devour, to swallow	mid.pa	'χmit.pa	'χmit.pa					
1023	V	to entangle	'khris.pa	'tʰɾis.pa	'tʰɾis.pa					
1024	V	to wear (clothes)	gon.ma	'gun.ma	'gun.ma					
1030	V	to smell (trans)	dri.byā	'tri.byā	'tri.byā					
1032	V	to shave	'breg.pa	'blaq.pa						
1034	V	to dry, to evaporate	rad.pa ? (Ladakhi: ras.pa) †	ʷɾas.pa	'ɾas.pa					
1039	V	to fight, to wrestle	'khril.ba	'tʰɾil.ba	'tʰɾil.ba					
	N			77	63	72	93	74	96	87
	A			8	7	26	12	10	12	12
	Num			9	0	0	0	0	11	0
	V			36	32	14	3	2	5	5
	Total			130	102	112	108	86	124	104

