Phonetic correlates of stress in Mapudungun

Michael O'Keefe
Rutgers University

1. Introduction

(1) The aim of this talk[1] is to accomplish the following:
   a. Provide a clear phonetic characterization of the stress pattern of Mapudungun
   b. Justify a phonological representation of Mapudungun stress
   c. Consider the implications of this representation

(2) Background on Mapudungun
   a. Language isolate (proposed: Penutian, Andean, Mayan, Arawakan)
   b. Central Chile, parts of Argentina
   c. ~440,000 speakers, but not highly promoted or regularly used
   d. Called "Araucanian" in the literature; this has negative connotations
      i. *La Araucana*: poem about the conquest of Chile (from the Mapuche)

(3) Why study Mapudungun?
   a. Presented as a canonical iambic stress language by Hayes (1995)
   b. There's a CD of recordings and lots of living speakers (not tremendously common for iambic languages)
   c. Widely varied portrayal in the literature

(4) The literature:

	Echeverria and Contreras (1965)	Smeets (1989)	Zúñiga (2006)
CVCV words	Free placement of stress	Lexical stress	Final stress
Secondary stress (default pattern)	Every second vowel	Every second vowel and every final vowel	Either the first or second syllable
Main stress	Second syllable	• Penultimate • Two main stresses in longer words (?)	Penultimate or ultimate
Open/closed	Final closed syllables attract secondary stress	Final closed or penultimate open	Closed syllables attract main stress
Lexical stress	In particles	In some suffixes and some CVCV words	No

(5) Conclusion: do some acoustic analysis to figure this out!

(6) Data source: Audio CD from Zúñiga (2006), *Mapudungun: El habla Mapuche* (has examples of phones, words, stories, and poems)[2]

[1] Many thanks to Crystal Akers, Will Bennett, Carlos Fasola, Paula Houghton, Paul de Lacy, Seunghun Lee, Alan Prince, and Fernando Zúñiga. Any remaining errors are my own.

<u>*General phonological overview*</u>

(7) Vowels

	front	central	back
high	i	ɨ	u
mid	e		o
low		a	

(8) [ɨ] is the "unmarked" vowel
  a. Shorter duration in sentence contexts than the others
  b. Omitted from some forms of orthography

(9) [ɨ] is not *just* a product of neutralization in non-stress positions
  a. [ŋɨ.ɹɨ] 'zorro' (fox)
  b. [ɨl] 'canto' (song)
  c. [pun] 'noche' (night)
  d. [wɨn] 'amanecer' (dawn)

(10) Consonants

	labial	(inter)dental	alveolar	palatal	retroflex	velar
Stop	p	t̪	t	t͡ʃ	t͡ɹ	k
Fricative	ɸ	ð̪	s			ɣ
Nasal	m	n̪	n	ɲ		ŋ
Liquid		l̪	l	ʎ	ɹ	ɰ
Glide	w			j		

Note: [ɰ] tends to pattern with glides.

(11) Syllable structure: (C)V(N/L/W)
  a. [pun] 'noche' (night)
  b. [ŋɨ.ɹɨ] 'zorro' (fox)
  c. [l̪aj] 'murió' (s/he died)

(12) Onsetless syllables word-initially:
  a. [ɨl] 'canto' (song)

[2] Please note: any English translations are my own and are generally for mnemonic assistance. As I am translating from Spanish to English, they should not be taken to accurately reflect the semantic subtleties of the original Mapudungun.

(13) No minimal pairs for vowel length.

(14) No minimal pairs for stress.

Simplifying assumptions (with respect to the analysis of stress)

(15) Observation: no minimal pairs for vowel length.
  a. Underlying forms for length don't matter.
  b. Any difference in duration is potentially relevant.

(16) Observation: no minimal pairs for stress.
  a. Outside of the fixed cases identified by Smeets (1989) stress must be surface-predictable.

(17) Assumption: vowel quality is unimportant with respect to stress, other than (potentially) the reduced vowel.
  a. Mapudungun isn't exactly diverse with respect to the vowel space.
  b. There are no immediately apparent restrictions on vowel placement.
  c. Nobody has previously observed anything of the sort.
  d. This gets complicated.

(18) Assumption: *statistically* significant patterns are *perceptually* significant patterns.
  a. This is almost certainly false in general.
  b. Determining whether these patterns are perceived requires a totally different sort of analysis.
  c. We can probably make reasonable guesses.
    i. "Position A tends to be ~0.2 dB louder than position B" = unlikely to be phonological.
    ii. "Position A tends to be ~20 Hz greater than position B" = plausibly phonological.

(19) Assumption: there is no interference by consonants other than the closed/open distinction.
  a. There is no good reason to assume this besides the fact that no pattern is particularly apparent.
  b. If the analysis were to be inconclusive, this would be good to consider.

Stress hypotheses to consider

Where *P* is some phonetic property:

(20) Null: all syllables have the same range of P

(21) Positional: compare all initial and final syllables in OO and KK words, etc.

(22) Grammar-based: there will be a significant difference for P between syllables predicted to be stressed in Zúñiga's account (for instance) and those predicted to be unstressed.

I will consider (20) and (21) for now and apply (22) to my own work in the future.

What is stress?

(23) I want to refrain from answering this fully for now.

(24) Standard view (e.g. Hayes 1995): stress corresponds principally to a particular phonetic property (duration, pitch, etc.), possibly with associated effects.

(25) Generally accepted diagnostics (Gussenhoven 2004):
  a. Spectral tilt: unstressed vowels have lower intensities for higher frequencies
  b. Tendency towards central unrounded in unstressed vowels
  c. Duration
  d. *Potentially* pitch, but this interacts with sentence intonation

(26) Why not to answer this directly for now:

Suppose we have a pattern characterized by "stress the second open syllable in any string of open syllables": CV.**CV**.CV.**CV**

How do we know if a closed syllable interferes with this pattern?

a. CV.**CV**.CVC.**CV**.CV or b. CV.**CV**.CVC.CV.**CV**

The phonetic nature of the CVC syllable doesn't actually matter in answering this question.

(27) ∴ While we would hope that CVC syllables behave like CV syllables, we don't actually have to look at them to make generalizations.
  a. If (purely hypothetically) CVC syllables marked prominence with stress but CV syllables marked it with duration, we wouldn't be able to compare them at all.
  b. If (more realistically) duration is the only significant property, we still have no way of comparing CVC and CV syllables to each other.

2. Method

Source materials

(28) Recording of *Feychi ngürü afngünengelu* from Zúñiga 2006
  a. Speaker information:

<u>*Analytical procedure*</u>

(29) Syllable nuclei were analyzed in Praat (Boersma and Weenink, 2007) for:
  a. Duration (s)
  b. Mean pitch (Hz)
  c. Min pitch (Hz)
  d. Max pitch (Hz)
  e. Mean intensity (dB)

(30) Praat settings
  a. Spectrograms
    i. Range: 0-5000 Hz
    ii. Window length: 0.005 seconds
    iii. Dynamic range: **35 dB**
    iv. Analysis method: Fourier, Gaussian window
  b. Pitch
    i. Range: 75-500 Hz
    ii. Optimized for intonation (AC method)
    iii. Silence threshold: 0.03
    iv. Voicing threshold: 0.45
    v. Octave cost: 0.01
    vi. Octave-jump cost: 0.35
    vii. Voiced/unvoiced cost: 0.14
  c. Intensity
    i. Range: 50-100 dB
    ii. Averaging: Mean energy - mean pressure

(31) Syllabic nuclei were identified visually as the range on the spectrogram across which F1, F2, and F3 were visible with the above settings.
  a. This is an source of potential (human) error.
    i. Algorithms for this are probably less reliable than I am.
  b. The *dynamic range* setting for spectrograms is the most crucial one here, as it determines where F1, F2, F3 are actually visible.

<u>*Phonological targets*</u>

(32) Assumption: in general, the (continuous) phonetic signal will be targeting some (discrete) representation
  a. e.g. "Stressed syllable = 400 Hz"
  b. If a local optimum is observable, it will coincide with the target
  c. Targets occur at a roughly consistent position in the syllable
    i. e.g. "Peak 50ms into the nucleus"
  d. If a local optimum can't be found, check the "target" position.

3. Results

(33) Sample text (track 41 from Zúñiga 2006)

Bold text: CVCV words for comparison.
Italic text: CVCV sequences for comparison.
Underlined: CVC syllables

IPA: **kiɲe** **paŋi** ka **kiɲe** **ŋɨɹɨ** mɨlej **kiɲe** **ɹuka** mew
Orthography: Kiñe pangi ka kiñe ngürü müley kiñe ruka mew.
Glosses:
Una leona y una zorra vivían en la misma casa.
A lioness and a vixen lived in the same house.

IPA: niej **kɨla** pɨɲeɲ fejtʃi **paŋi**
Orthography: Niey küla püñeñ feychi pangi.
Glosses:
La leona tenía tres hijos.
The lioness had three children.

IPA: fejmew **kiɲi** antɨ fejpi **paŋi**
Orthography: Feymew kiñe antü feypi pangi:
Glosses:
Cierto día dijo el puma a la zorra:
One day the puma said to the vixen:

IPA: iɲtʃe kin*tume*an iloal ejmi *mɨle*ajmi *pitʃile*wej ilo dewmajaymi iɲtʃe ipatuan pi **paŋi**
Orthography: "Iñche kintumean iloal", pingey nguru, "eymi müleaymi; pichilewey ilo, dewmayaymi; iñche ipatuan", pi pangi.
Glosses:
"Voy a ir en busca de carne; quédate aquí; hay todavía un resto de carne; la prepararás. A mi vuelta voy a servírmela."
"I am going to go in search of meat; stay here; there is still some meat left; you will prepare it. When I return I will serve myself."

CV.CV words

(34) Simplest hypothesis: there will be some statistically significant difference for some property between the first and second vowel.

Duration of CVCV words in non-sentence-final positions

(35) Data from audio tracks 41 and 42:

	V1 duration (s)	**V2 duration (s)**
kiɲe	0.087548	0.10214
paŋi	0.115689	0.138618
kiɲe	0.074617	0.106805
ŋɨɹɨ	0.066265	0.065866
kiɲe	0.056102	0.089075
ɹuka	0.06088	0.077011
kɨla	0.033313	0.094508
kiɲi	0.058831	0.102517
paŋi	0.11014	0.123848
fitʃi	0.049297	0.053622
paŋi	0.145667	0.1368
kɨla	0.038515	0.085105
tʃaʎa	0.087442	0.084003

(36) t-Test result

t-Test: Paired Two Sample for Means		
	V1 duration	**V2 duration**
Mean	0.075715846	0.096916769
Variance	0.001067819	0.000651649
Observations	13	13
Pearson Correlation	0.760910974	
Hypothesized Mean Difference	0	
df	12	
t Stat	-3.603437608	
P(T<=t) one-tail	**0.001811491**	
t Critical one-tail	1.782287548	
P(T<=t) two-tail	**0.003622982**	
t Critical two-tail	2.178812827	

Summary: this tests the hypothesis that the mean of V2 is *either higher or lower* than the mean of V1. Low values for P are good; as this is less than .05; it is significant.

More specific:

The two-tail P value answers the following:
What is the probability that randomly selected samples will have means as far apart as we observe?

The one-tail P value answers the following:
What is the probability that, in a randomly selected sample, the mean V1 duration will be lower than the mean V2 duration?

Both of these are quite low. The average difference between the samples is about 20 milliseconds; this is really small but the correlation is very strong.

(37) Sentence-final vowels often exhibit finality effects; they appear to behave the same here (at most, they skew things further in the direction we'd expect)

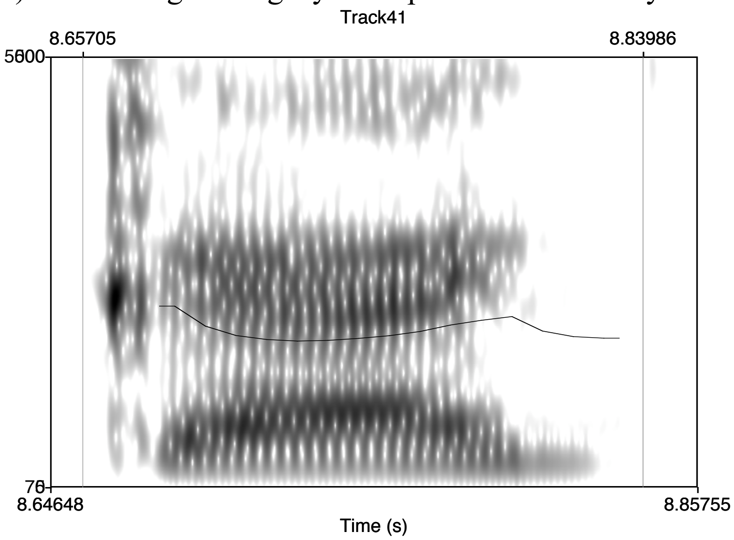
paŋi	0.104554	0.157116
paŋi	0.114373	0.115552
paŋi	0.092415	0.132669
ŋɨɹɨ	0.071668	0.139355

(38)

t-Test: Paired Two Sample for Means		
	V1 duration	**V2 duration**
Mean	0.080430353	0.106153529
Variance	0.000941078	0.000838739
Observations	17	17
Pearson Correlation	0.686469006	
Hypothesized Mean Difference	0	
df	16	
t Stat	-4.481630076	
P(T<=t) one-tail	0.00018881	
t Critical one-tail	1.745883669	
P(T<=t) two-tail	0.000377621	
t Critical two-tail	2.119905285	

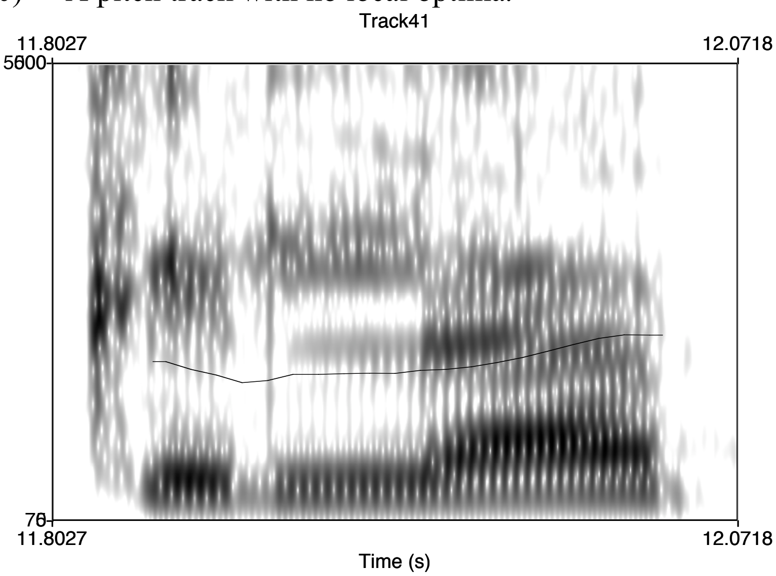
<u>Pitch of CVCV words in non-sentence-final position</u>

(39) Irritating finding: syllable pitches are virtually all level.



[ni]

(40) A pitch track with no local optima.



[**kɨla**]

(41) What does this tell us?
   a. It *probably* means that there isn't any phonetic target here with respect to pitch.
   b. Meaning: the phonology doesn't care where the pitch track is, and any variations are due to consonantal pressures.
      i. To be precise: pitch has nothing to do with stress.

(42) In practice, this means that we're stuck with mean pitch.
   a. Not a fantastic measure.

(43) Data from audio tracks 41 and 42

	V1 pitch (Hz)	V2 pitch (Hz)
kiɲe	206.0681535	251.5575276
paŋi	269.0901117	255.9521862
kiɲe	245.0976014	239.7995819
ŋɨɹɨ	215.0163328	220.7105671
kiɲe	195.0933412	206.0276298
ɹuka	201.5165844	173.3021566
kɨla	218.1325192	231.2783313
kiɲi	216.4654301	227.6614599
paŋi	206.7779406	228.232738
fitʃi	177.4596465	285.1738701
paŋi	204.3577497	255.2165338
kɨla	183.5591527	203.0889474
tʃaʎa	229.7645171	240.1235318

(44)

t-Test: Paired Two Sample for Means		
	V1 pitch	**V2 pitch**
Mean	212.9537755	232.1634663
Variance	603.6371623	797.1655813
Observations	13	13
Pearson Correlation	0.177352006	
Hypothesized Mean Difference	0	
df	12	
t Stat	-2.038206023	
P(T<=t) one-tail	0.032096766	
t Critical one-tail	1.782287548	
P(T<=t) two-tail	**0.064193532**	
t Critical two-tail	2.178812827	

(45) The distance in means between these two samples is statistically significant (.032).

(46) The hypothesis that V1 has a lower pitch is *not* (.064).

(47) Either this is bad data (unlikely) or pitch isn't predictable in this manner.

(48) Note: the difference between 210 Hz and 230 Hz is easily perceptible.
  a. Generate a Praat sound: sin(2*pi*210*x)
  b. This is certainly a plausible distinction.

(49) Conclusion: pitch isn't relevant.
  a. It's impossible to find distinct targets (local optima).
  b. The means clearly differ, but it's unclear how.
  c. Duration looks promising, and it's generally held that duration and pitch are significant in different languages (e.g. Hayes 1995).

<u>Intensity of CVCV words in non-sentence-final position</u>

(50) Preface: nobody really thinks duration matters.
  a. Nobody really has a good universal definition of stress either, though.

(51) Conclusion: they're probably right not to care.

	V1 intensity (dB)	**V2 intensity (dB)**
kiɲe	81.22384222	81.69402536
paŋi	80.69679467	81.32448449
kiɲe	80.23464844	81.3991576
ŋɨɹɨ	80.11664759	80.16663494
kiɲe	78.75476963	79.51670707
ɹuka	82.01478748	78.11457089
kɨla	80.94328104	80.36425797
kiɲi	80.98508128	82.07989759
paŋi	81.12733016	82.28687049
fitʃi	77.43643597	84.84769271
paŋi	82.33104644	84.46624117
kɨla	80.54036571	81.44406417
tʃaʎa	81.78373943	82.37840306

(52) Conclusion: they're probably right not to care.

t-Test: Paired Two Sample for Means		
	V1 intensity	**V2 intensity**
Mean	80.62990539	81.54484673
Variance	1.759516084	3.354219041
Observations	13	13
Pearson Correlation	-0.155271134	
Hypothesized Mean Difference	0	
df	12	
t Stat	-1.361802551	
P(T<=t) one-tail	**0.099139407**	
t Critical one-tail	1.782287548	
P(T<=t) two-tail	**0.198278813**	
t Critical two-tail	2.178812827	

(53) This is well below significance.

(54) Even given that, the difference in means is under 1dB.
   a. Definition of a decibel: 1dB = minimum perceptible intensity difference.

<u>Interim summary</u>

For CVCV words:
(55) There is a strong correlation between position and duration.
(56) Second syllables have greater duration.
(57) Pitches tend to vary, but unpredictably.
(58) Intensity tells us nothing.

(59) If the duration story can't be extended, check on pitch and intensity.
   a. If it can, stick to duration, since it's the safest to measure.

<u>Adding closed syllables</u>

(60) An issue: what do we compare CVC (and VC) syllables to?

(61) Solution: make a hypothesis about stress

a. Hypothesis: closed syllables always attract stress and subsequent sequences of open syllables will show an canonical iambic pattern

i. [pa.tá] ii. [pák.ta] iii. [pa.ták.ta.pá]

b. Suppose we see a form [pa.tak.ta.pa.ka]. What vowel do we care about?

i. [pa.ták.ta.**pa**.ka]

c. If the closed syllable is messing with the stress of the word, then we can see it on a subsequent open syllable.

i. Advantage: we just compare open syllables to open syllables this way.

(62) Representative data

	V1 duration	**V2 duration**	**V3 duration**	**V4 duration**	**V5 duration**
kintuntɨkuɲe	0.040459	**0.091447**	**0.112508**		
kintutɨkumelu	**0.101907**	0.57474	0.58531	**0.110716**	**0.090261**
wajontɨkufi	0.02806	0.06441	**0.161982**		
wadkɨletʃi	0.043235	**0.068982**	0.058781		
ramtupatuj	0.045476	**0.08152**	CVC		
eluŋej	0.03197	**0.061219**	VC		
mɨtrɨmel	0.042	**0.049**	CVC		
metuen	**0.074878**	0.03663	CVC		
wɨnokintuj	0.048628	**0.08586**	CVC	CVC	
rejɨkɨnufi	0.063196	**0.076423**	0.06	0.06466	**0.166073**
pitʃilewetʃi	0.030535	0.040477	**0.079533**	0.067461	**0.111541**
pitʃike	0.062885	**0.068714**	0.06392		
ilokorɨ	**0.113246**	0.094372	**0.10562**	0.076566	

(63) Final syllables *tend* to be longer than others; second syllables *tend* to be longer; neither tendency looks particularly reliable.

<u>*Duration in CVC.CV.CV sequences*</u>

(64) Hypothesis: the second CV syllable after a closed syllable will have greater duration than the first.

t-Test: Paired Two Sample for Means		
	Duration 1	***Duration 2***
Mean	0.047586714	0.141616857
Variance	0.000613546	0.03666712
Observations	10	10
Pearson Correlation	0.970128299	
Hypothesized Mean Difference	0	
df	9	
t Stat	-1.484684768	
P(T<=t) one-tail	**0.094082115**	
t Critical one-tail	1.943180274	
P(T<=t) two-tail	**0.188164231**	
t Critical two-tail	2.446911846	

Result: this appears to be wildly inaccurate.

(65) Conclusion:
a. Mapudungun stress is not produced by quantity-sensitive left-aligning iambs. (CVC)(CV.CV)(CV.CV)

(66) Prediction: open syllables in even-numbered positions will have a correlation in duration, and open syllables in odd-numbered positions will have a correlation in duration.

t-Test: Paired Two Sample for Means		
	Odd positions	***Even positions***
Mean	0.086281964	0.099617214
Variance	0.010423598	0.009397479
Observations	28	28
Pearson Correlation	0.132706883	
Hypothesized Mean Difference	0	
df	27	
t Stat	-0.538132286	
P(T<=t) one-tail	**0.297447231**	
t Critical one-tail	1.703288423	
P(T<=t) two-tail	**0.594894462**	
t Critical two-tail	2.051830493	

Result: this is wildly inaccurate as well.

(67) Conclusion:
  a. Mapudungun stress is not produced by quantity-insensitive left-aligning iambs.

  (CVC.CV)(CV.CV)

(68) Thus, we have disproved the accounts of Echeverria and Contreras (1965) and Smeets (1989).

Discussion

What's going on? It's not entirely clear at present, but some patterns have been identified:

(69) It was shown that a phonologically heterogenous set of words similar only in syllable structure (CVCV) exhibit a correlation between duration and position.

(70) No other properties show such a correlation.
  a. Thus duration is the manifestation of stress.

(71) It was shown that there is no correlation between position relative to closed syllables and the duration of open syllables.
  a. This is unlike the case reported by Derbyshire (1979, 1985).

(72) It was shown that there is no strict correlation between even/odd syllable position and open syllable duration.
  a. Mapudungun is not a quantity-insensitive left-aligning iambic system, as has been reported.

References

Boersma, Paul & Weenink, David (2007). Praat: doing phonetics by computer (Version 4.6.36) [Computer program]. Retrieved November 2, 2007, from http://www.praat.org/

Derbyshire, Desmond C. 1979. *Hixkaryana*. Lingua Descriptive Studies 1. Amsterdam: North-Holland.

Derbyshire, Desmond C. 1985. *Hixkaryana and linguistic typology*. SIL Publications in Linguistics 76. Summer Institute of Linguistics.

Echeverría, Max S., and Heles Contreras. 1965. Araucanian phonemics. *IJAL* 31(2): 132-135.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

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

Smeets, Ineke. 1989. *A Mapuche Grammar*. Leiden, Holland.

Zúñiga, Fernando. 2006. *Mapudungun: El habla Mapuche*. Santiago, Chile: Centro de Estudios Públicos.

email: mcokeefe@eden.rutgers.edu