Speech perception in the field

Dr. Christian DiCanio
cdicanio@buffalo.edu

University at Buffalo

5/29/18

Meta-outline for the lectures

1. The analysis of complex tonal systems: motivations, methods, and analysis
2. **Speech perception in the field**
3. Creating and working with endangered language corpora
4. Higher-level prosody and tone

Outline

1. Production and perception of tone/phonation
2. Methodological concerns
3. Categorical perception of tone (DiCanio, 2012b)
4. Perceptual cue-weighting of non-modal phonation (DiCanio, 2014)
5. Recap - future directions

Please feel free to ask questions.

How does our experience with language influence our ability to produce and perceive certain speech sounds?

'*fat*' [fæt] vs. '*vat*' [væt]

'*thin*' [θɪn] vs. '*fin*' [fɪn]

Not all minimal pairs are equally distinguishable. Why?

Some possibilities

1. Psychoacoustics. Some contrasts are just harder to distinguish for the human ear.

2. Language background. Exposure to certain phonological contrasts trains the listener/speaker to use the maximally useful phonetic cues for distinguishing them.

How might a listener's native language influence their sensitivity to certain phonetic cues?

Speech perception in a fieldwork context

- One may investigate cross-linguistic variation in perception with well-known languages, but these account for a small overall percentage of the diversity of human languages.
- Minority languages tend to have more "unusual" or complex properties that we do not fully understand (Whalen, 2004; Whalen and McDonough, 2015).
- How might being a native speaker of such a language influence the perception of well-known contrasts?
- How might the more unusual contrasts be perceived?

Challenges and rewards

Investigating perception in a fieldwork context poses challenges in terms of infrastructure, experimental design, and resources, but it is also rewarding.

- Gujarati listeners have lower JNDs for voice quality than listeners who do not speak a language with a phonation contrast (Kreiman et al., 2010).
- Sub-allophonic differences in sound inventories in Dravidian languages influence the perceptual space of nasal consonants (Harnsberger, 2001).
- Native listeners of Mixtec are sensitive to F_0 changes of just 1 Hz in their perception of glottalization (Gerfen and Baker, 2005).

Topics of investigation

Investigation of the perception of some of these "unusual" sound types found in Itunyoso Triqui, a native language of Mexico (DiCanio, 2008, 2010).

1. The influence of language background on tonal discrimination and sensitivity to tonal cues (DiCanio, 2012b).
2. Weighting of acoustic cues for the perception of glottal contrasts (DiCanio, 2014).

What does this work reveal that is relevant to general questions in speech perception?

Tones in Oto-Manguean: a typological abundance

- Of the world's tone languages, 60% (132/220) have only 1-2 lexical tone contrasts and 40% have three or more tonal contrasts (88/220).
- Among the tone languages with large inventories, languages with between 3-6 tonal contrasts are relatively common, e.g. Thai (5), Mandarin (4), Vietnamese (6), Cantonese (6), Yoruba (3).
- Languages with greater than 6 tones are rarer, but many are Oto-Manguean, e.g. Itunyoso Triqui (9) (DiCanio, 2008), Yoloxóchitl Mixtec (10) (DiCanio et al., 2012), Chatino (10) (Cruz and Woodbury, 2005), Tlacoatzintepec Chinantec (7) (Thalin, 1980), Chiquihuitlan Mazatec (17) (Jamieson, 1977).

Itunyoso Triqui

One such Oto-Manguean language with a complex tonal system is *Itunyoso Triqui*, spoken in Oaxaca, Mexico by about 2,500 speakers (DiCanio, 2008, 2010, 2012a,b,c, 2014).

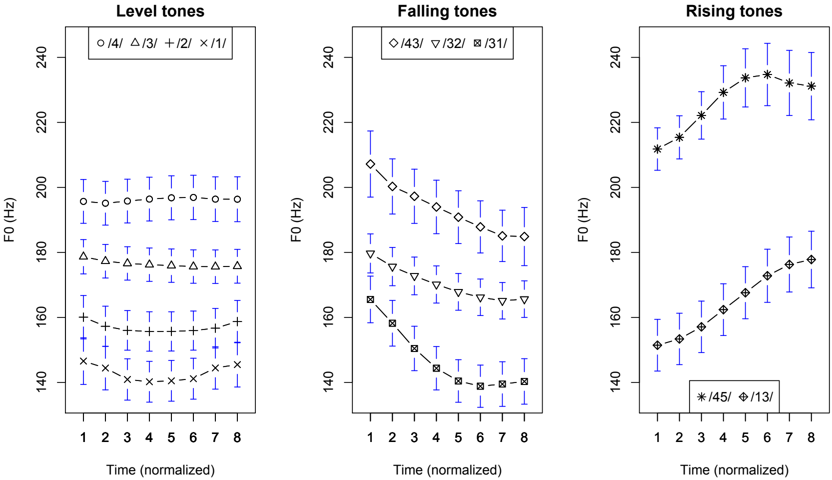
Triqui tone

There are **nine** tones in Triqui, including level, falling, and rising tones (DiCanio, 2008).

Tone	IPA	Gloss	Tone	IPA	Gloss
4	ββe^{4}	'*hair*'	43	li^{43}	'*small*'
3	nne^{3}	'*plough*'	32	nne^{32}	'*water*'
2	nne^{2}	'*to lie (tr.)*'	31	nne^{31}	'*meat*'
1	nne^{1}	'*naked*'	45	yoh^{45}	'*my forehead*'
			13	yo^{13}	'*light, quick*'

Numbers are used to mark tone, as if it were a musical scale, with "1" being lowest and "5" highest.

Triqui tones are distinguished by pitch level and slope.



Tone perception

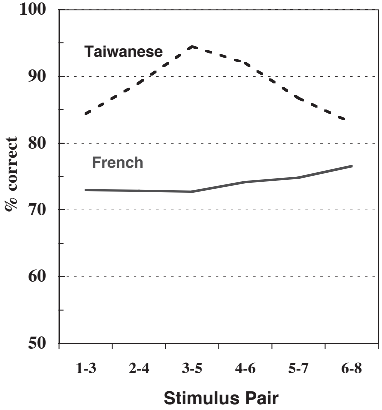
Most of what we know about tone perception is based on East and Southeast Asian languages (Chinese, Cantonese, Thai, Vietnamese).

In relation to tonal categorization, the literature shows mixed findings:

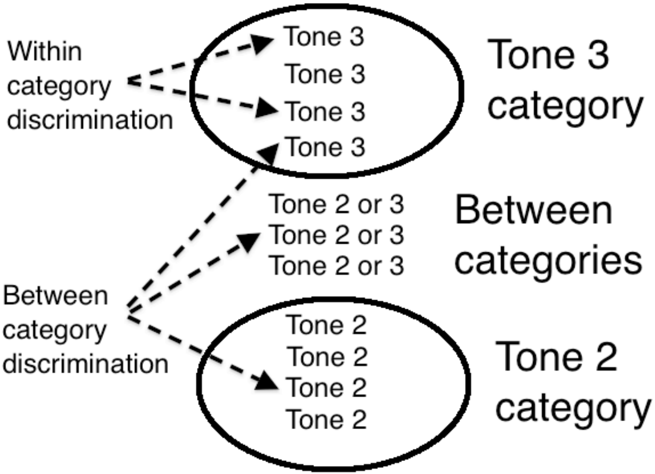
- Tone is perceived categorically (Chan et al., 1975; Stagray and Downs, 1993; Lee et al., 1996; Hallé et al., 2004; Xu et al., 2006; Peng et al., 2010).
- Tone is **not** perceived categorically (Abramson, 1979; Avelino, 2003; Francis et al., 2003).

Categorical effects in discrimination

Given a continuum from tone A to tone B (changing F_0), native listeners are better at discriminating *between* category stimuli than *within* category stimuli. Figure from Hallé et al. (2004).



Within and between category discrimination



What accounts for the mixed findings?

- Like vowels, tones have fuzzier boundaries, which means that their categorization functions will be less steep in identification tasks and discrimination better than predicted by identification (Hallé et al., 2004; Xu et al., 2006).

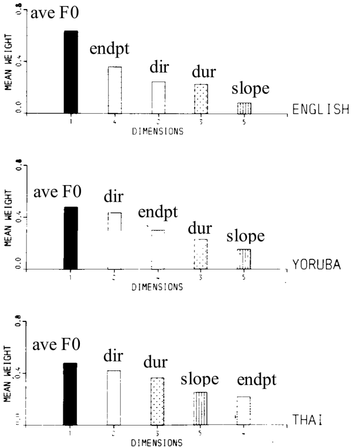
- Difficult to determine categorical effects in discrimination from a language-internal baseline (within-subjects design). Studies looking at cross-linguistic perception (between-subjects design) have found clear effects (Hallé et al., 2004; Peng et al., 2010).

- Different effects for level vs. contour tones due to normalization.

The effect of language experience

- Tone language listeners may be more sensitive to dynamic cues in tone perception than non-tone language listeners (Gandour and Harshman, 1978; Lin and Repp, 1989; Krishnan et al., 2005).
- Native tone language listeners (Mandarin Chinese, Cantonese, Thai) are better at discriminating pitch than listeners of non-tonal languages (English, German) (Burnham et al., 1996; Lee et al., 1996; Hallé et al., 2004; Peng et al., 2010; So, 2006).
- Mandarin Chinese listeners are *worse* at discriminating pitch than English listeners, but only *within* categories, not between them Stagray and Downs (1993).

Weight of different tonal cues by language



(Gandour and Harshman, 1978)

Questions/predictions 1

What do native tone speakers and listeners pay attention to that non-natives might not?

- If there are categorical effects in tonal discrimination, Triqui listeners will perceive between-category stimuli more accurately than within-category stimuli.
- Are native Triqui listeners better at pitch perception than speakers of a non-native tone language or just better at categorical boundaries?

Questions/predictions 2

What types of contrasts are harder to distinguish if you speak a tone language?

If speaking a tone language confers an advantage to using dynamic cues, like slope, then native listeners should better distinguish between contour tones than non-native listeners do.

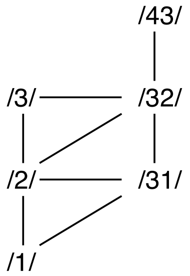
Discrimination experiment (DiCanio, 2012b)

Tested discrimination of Triqui tones by native listeners and French listeners following AXB discrimination task described in Hallé et al. (2004).

- 8 Blocks of 48 trials preceded by 1 practice block of 32 trials.
- All tonal stimuli appeared in carrier sentence:
  ka^3*tah*^{3} <target> , 'He says <*target*>.'
- Subjects: 18 native speakers of Itunyoso Triqui (all bilingual Triqui-Spanish), 20 native speakers of French.
- Location: Oaxaca, Mexico and Lyon, France. Experiment run in Spanish or French.

Comparisons

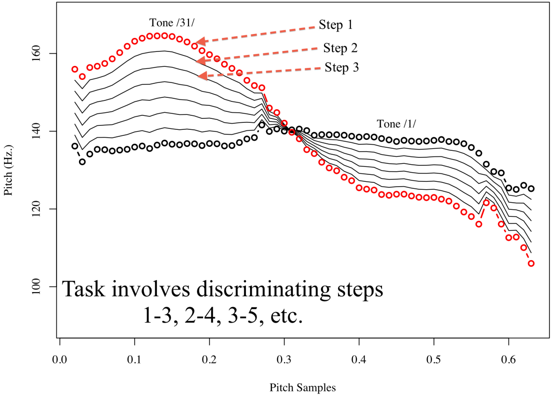
Resynthesized pitch on tokens using 8 steps for each of 8 tonal pairs.



Not all tones were analyzed; those that were were all produced on open syllables.

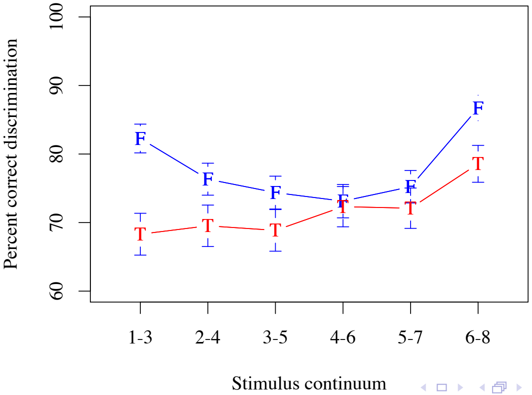
Resynthesis

8-step linear interpolation of pitch and intensity of original tones using Praat (Boersma and Weenink, 2016) and Matlab. Tokens were matched for duration; e.g. /nne^{31}/ ‘meat’ vs. /nne^{1}/ ‘naked’. Resynthesized speech



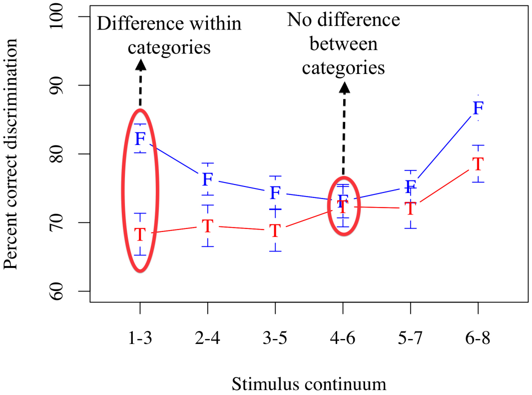
Results I

French speakers performed better (78.0%) overall at tonal discrimination than Triqui speakers (71.6%), but Triqui listeners showed better between-category discrimination than within-category discrimination.

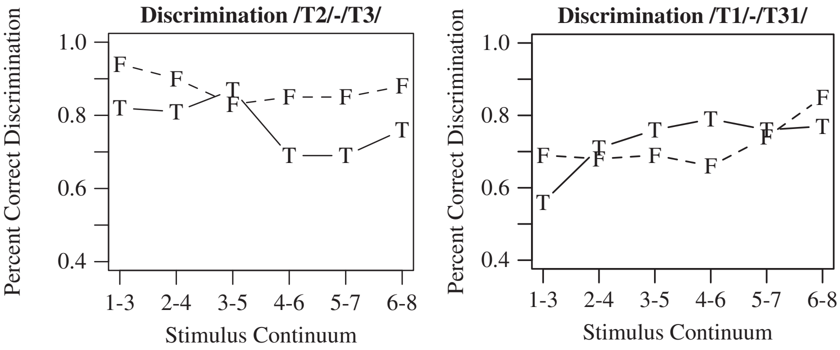


Results I: Categorization effects

French speakers performed better (78.0%) overall at tonal discrimination than Triqui speakers (71.6%), but Triqui listeners showed better between-category discrimination than within-category discrimination.



Specific tonal comparisons



At the continuum endpoints, Triqui listeners are poorer at discriminating stimuli than French listeners. At the midpoints, they improve.

Results 2: Tone type

Though Triqui listeners were poorer at discriminating Triqui tones than French listeners, were there certain tonal pairs they were better at?

Tone type comparison	Triqui accuracy	French accuracy
Level x Level	63.8%	77.6%
Level x Contour	68.7%	75.9%
(Difference)	-5.1%	2.7%

Triqui listeners were significantly better at discriminating level vs. contour tones than level vs. level tones. French listeners showed a slight preference in the opposite direction.

Results 3: Psychoacoustic Effects

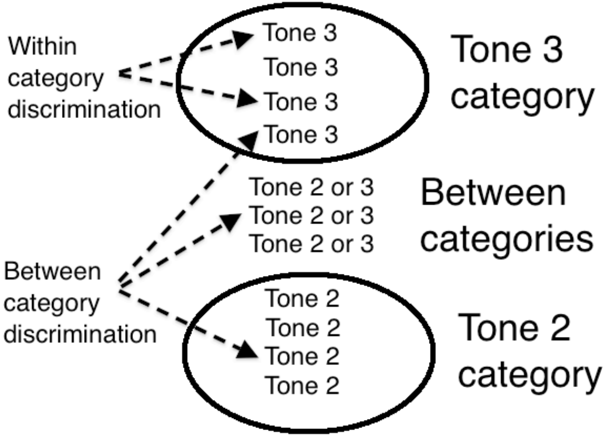
- Correspondence between the raw psychoacoustic distance between each stimulus pair and the degree of discriminability by listeners.
- Psychoacoustic distance between stimuli = average difference in semitones between each stimulus pair.

Tonal Comparison	Psychoacoustic Distance (semitones)	Discrimination Accuracy Triqui	French
/32/ - /31/	1.02	71.6%	87.2%
/2/ - /3/	0.81	72.3%	87.5%
/43/ - /32/	0.75	58.5%	71.1%
/2/ - /32/	0.64	69.1%	81.1%
/2/ - /31/	0.63	77.4%	83.4%
/1/ - /31/	0.58	67.1%	72.0%
/2/ - /1/	0.46	60.6%	74.2%
/3/ - /32/	0.25	61.0%	67.2%

Interim discussion

- Overall, French listeners are better at discriminating pitch, but they show no sensitivity to categorical boundaries between stimuli.
- Triqui listeners are better at distinguishing pitch when doing so distinguishes words in the language; i.e. at categorical boundaries.
- Triqui listeners seem to ignore within-category pitch differences. This is in line with Stagray and Downs (1993), but contra some recent work on tone perception Burnham et al. (1996); Hallé et al. (2004); Peng et al. (2010).

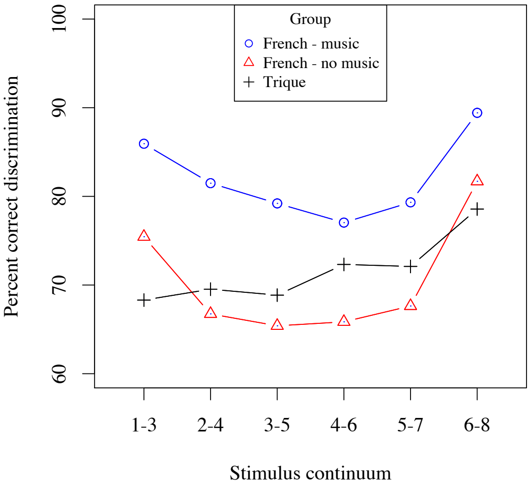
Triqui listeners are better at between category discrimination because the *linguistic boundaries* are here.



But why aren't they better at pitch discrimination?

- Unlike Mandarin, Thai, and Cantonese speakers, there is no literacy in Triqui.
- Moreover, Triqui subjects are not familiar with experimental tasks. Unfamiliarity with experimental procedures may explain some of the language differences in discrimination accuracy.
- Most studies do not control for music experience, which influences one's ability to perceive tone (Deutsch et al., 2009).
- While many of the French subjects (13/20) had some music training, such training is rare for Triqui listeners.

French listeners with no music training (70.5%) discriminated tone similarly to Triqui listeners (71.6%), but improved if they had training (82.1%)



The studies which found better pitch perception for tone language listeners did not specifically examine whether native tone language listeners had a music background.

Yet, the exception, Stagray and Downs (1993), involved native Mandarin listeners who specifically *had no musical experience*. Subsequent work found this too (So and Best, 2010).

Suggests that musical experience trains listeners to attune to non-contrastive phonetic detail.

What influences one's ability to perceive tone in Triqui?

1. Language background
   - Native listeners better discriminate between-category stimuli than within-category stimuli.
   - Native listeners pay more attention to the change in F_0 over time (slope) than absolute F_0 level (c.f. Kewley-Port et al. (1983); Mirman et al. (2004)).
2. Music background (c.f. Kühnis et al. (2013), Bidelman et al. (2014)).
3. Psychoacoustic distance between stimuli (c.f. Francis et al. (2003)).

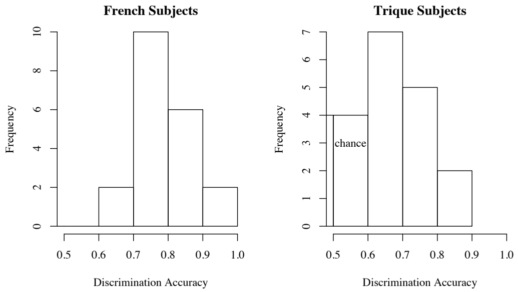
General findings

Language experience does not enhance the ability to discriminate phonetic details, but rather assists in the formation of clear perceptual boundaries between meaningful contrasts.

The goal of speech perception is categorization (Best et al., 2003; Xu et al., 2006).

But... language differences in discrimination

- Strong subject effect in discrimination for all tonal comparisons, mean $G^2(36) = 171.1$, $p < .001$ ***.
- Language effects on discrimination partly explainable by differences among individual listeners.



On perceptual cues

- Phonological contrasts are multidimensional and involve many distinct cues which vary in their perceptual importance (Keating, 1984; Kingston and Diehl, 1995).
- Tone is multidimensional, but we have not specifically tested listener sensitivity to specific cues.
- We can tease apart which cues are more important for listeners by investigating their relative weight in a labelling task.
- Rather than look at tone for this, we will focus instead on cue ranking among another phonological contrast in Triqui.

Glottal contrast

- In addition to tone, Itunyoso Trique has a 3-way rime-type contrast: /Vː/, /Vh/, /Vʔ/.
- Coda /h/ is usually realized as vocalic breathiness spread across the latter half of the rime. Coda /ʔ/ is realized with an abrupt glottal closure (DiCanio, 2012a).
- Duration of modal vowel before laryngeal coda is shorter than duration of vowel without coda.

nne^{3} ‘*plough*’ nneʔ^{3} ‘*fiber rope*’ nneh^{3} ‘*toothless*’

Glottal Consonants & Non-Modal Phonation

- Substantial research on the production mechanisms and the acoustics of non-modal phonation (Kirk et al., 1984; Ladefoged et al., 1988; Gordon and Ladefoged, 2001; Blankenship, 2002; Wayland and Jongman, 2003; Pennington, 2005; Esposito, 2006; Keating and Esposito, 2006; Kreiman et al., 2007; DiCanio, 2009, 2008:among others).
- Until recently, little work on the perceptual status of different acoustic properties.

Perceptual Cues for Phonation Type

- **Spectral Slope**
  - H1-H2 (OQ measure) (Esposito, 2010; Kreiman et al., 2010)
  - Global spectral shape (H1-A3, A1-A3, etc.) (Esposito, 2006; Kreiman et al., 2010)
- **Noise-related**
  - Cepstral Peak Prominence (Kreiman et al., 2010)
- **Prosodic / Suprasegmental**
  - Intensity Contour (Hillenbrand and Houde, 1996; Gerfen and Baker, 2005)
  - Pitch (Hillenbrand and Houde, 1996; Gerfen and Baker, 2005)
  - Duration (Lyon, 2008)

Perceptual findings with respect to voice quality come from:

- Linear discriminant analyses on the discrimination performance of listeners evaluating natural stimuli (Esposito, 2010).
- Discrimination tasks where prosodic cues were manipulated using resynthesized natural speech (Hillenbrand and Houde, 1996; Gerfen and Baker, 2005).
- Discrimination tasks where spectral cues were manipulated using synthetic speech (Kreiman et al., 2010).

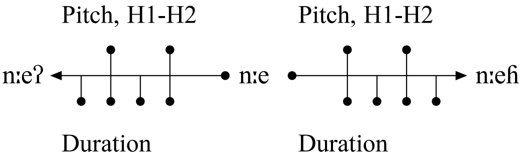
Perceptual Cue Weight

- So far, little work on the relative weight of distinct acoustic cues for glottal contrasts.
- Listeners pay attention to certain cues more than others in speech perception (Broersma, 2005; Cho and McQueen, 2006; Escudero, 2005; Gottfried and Beddor, 1998; Harnsberger, 2001; McGuire, 2007).
- Production data suggests many possible cues, such as spectral tilt, pitch, and duration, may be used (DiCanio, 2008, 2012a).
- What cues are more important for listeners in their perception of glottal consonants?

Method - Multidimensional resynthesis (DiCanio, 2014)

- AX labelling task with 14 native speakers of Itunyoso Trique.
- Manipulated duration, H1-H2, and pitch on modal vowel word (nne^{3} '*plough*') to match values on laryngealized rimes (nneh^{3} '*toothless*' and nneʔ^{3} '*fiber rope*') with help of Praat and Matlab scripts.
- 3 sets of two dimensions: duration X pitch, duration X H1-H2, H1-H2 X Pitch.
- 2 laryngeal conditions were tested: /V:/ vs. /Vh/, /V:/ vs. /Vʔ/.
- Total of 6 blocks (3 x 2).

- Manipulated duration using 6 steps, other cues using 4 steps. For the H1-H2 X Pitch condition, manipulated H1-H2 using 6 steps. Total of 24 stimuli, each repeated twice (48 trials).



- Targets spliced into original carrier sentence:
  <target> ka^{4}tah^{4} ri^{3}ə̃^{32}ɾeʔ^{1}‘, **<target>** I told you!’.

- Listeners pressed right or left button on keyboard corresponding to visual target on screen.

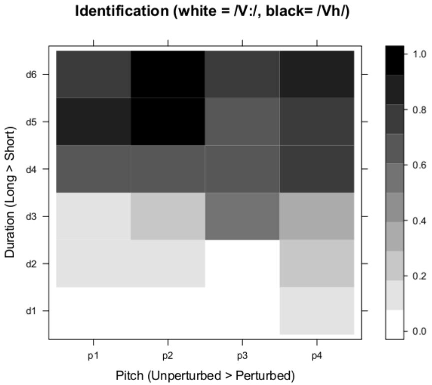

© La Ferme des Acacias - Gabriel R. 2004 http://www.lesacacias.net

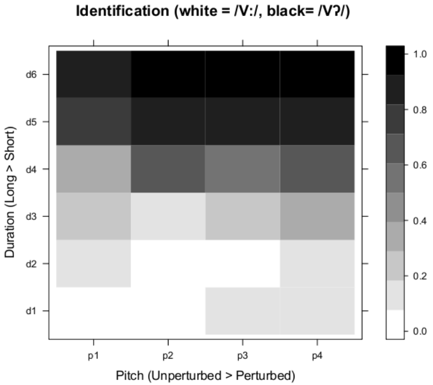
Results

- Identification results analyzed with a linear mixed effects model with manipulated dimensions as fixed effects and subject as a random effect.
- Response time results analyzed in a two-way repeated measures ANOVA with subject as an error term.

Results I: Duration and Pitch Manipulation

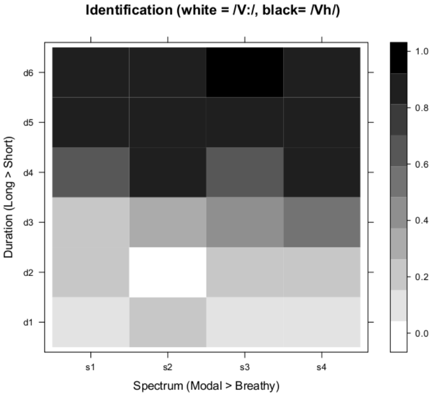
For both comparisons, duration, but not pitch, was a significant cue.

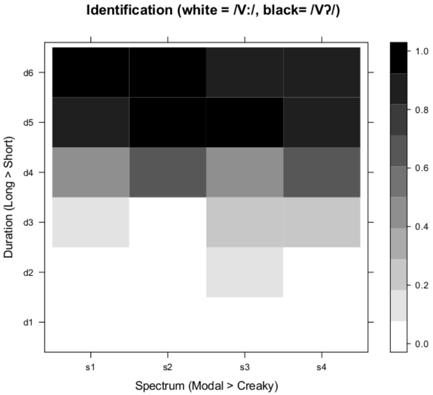




Results 2: Duration and H1-H2 Manipulation

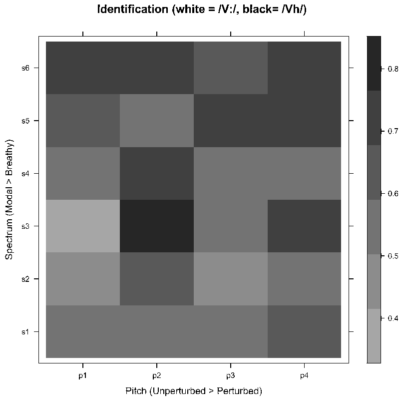
For both comparisons, duration, but not H1-H2, is a significant cue.
For the /V:/ - /Vh/ comparison, a significant interaction between duration and H1-H2 was found.

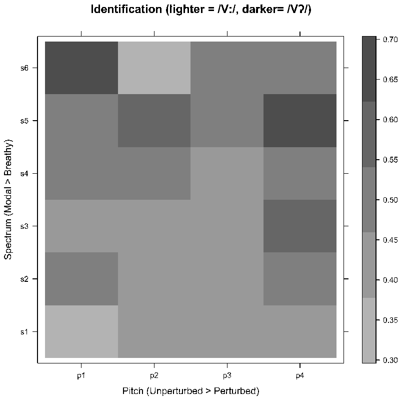




Results 3: Pitch and H1-H2 Manipulation

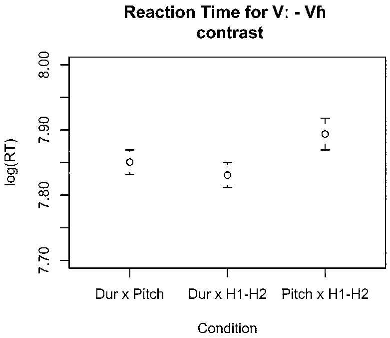
For the /V:/ - /Vh/ comparison, pitch, but not H1-H2, is a significant cue. For the /V:/ - /Vʔ/ comparison, only large H1-H2 differences were significant.

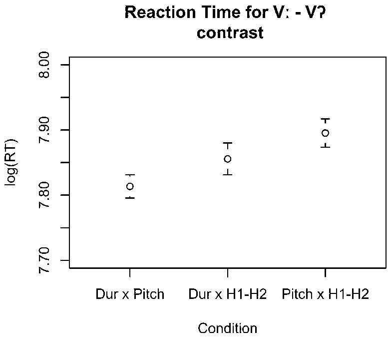




Results: Response Time

Significant effect of available cues on log(RT). Longer RT occurred in conditions where duration was ambiguous.





Summary

- In all cases where duration is not ambiguous, it is used as a strong cue for /Vh/ and /Vʔ/ rimes. Shorter duration stimuli are identified faster than longer duration stimuli.
- When H1-H2 is ambiguous, laryngeally-induced pitch raising caused a small shift in identification of /V:/ > /Vh/.
- When pitch is ambiguous and duration is ambiguous, H1-H2 acts as a cue to distinguish /V:/ and /Vh/ rimes.
- When duration is ambiguous, only H1-H2 acts as a cue to distinguishing /V:/ and /Vʔ/ rimes.

Discussion

What are the cues used to distinguish the glottal consonant contrast in Itunyoso Trique?

Contrast	Ranking
/Vh/ vs. /V:/	Duration > H1-H2 > Pitch
/Vʔ/ vs. /V:/	Duration > H1-H2

The Perception of Coarticulation

- Final syllables in Trique are bimoraic, consisting of either a long vowel /Vː/ or a vowel followed by a glottal coda (/h/, /ʔ/).
- While duration is prosodic cue related to the presence or absence of a coda consonant, changes in voice quality and accompanying pitch perturbations are unambiguously the result of vowel-glottal coarticulation.
- Supports the hypothesis that coarticulatory cues in speech production are useful for the perceptual identification of phonological contrasts (Beddor and Krakow, 1999; Holt and Lotto, 2006; Nowak, 2006; Beddor, 2009).

Laryngeal cues and cue-weight

- As per DiCanio (2012a), greater coarticulatory overlap in /Vh/ rime causes more salient changes in pitch and spectral tilt on the vowel. Such effects are not compensated for, but are directly related to the perception of laryngeal contrasts. Listeners use coarticulatory effects in perceptual identification.
- When a dominant cue is available, subordinate cues are not used by listeners. When a dominant cue is not available, listeners attend to subordinate cues. This is an example of *adaptive plasticity* in speech production (Holt and Lotto, 2006).

Conclusions: Triqui tone perception

- Language experience facilitates perceptual sensitivity to tonal boundaries, but not greater sensitivity to pitch.
- Native Triqui listeners largely ignore within-category phonetic differences; they're not meaningful.
- Native tone language listeners pay relatively more attention to dynamic tonal cues than static ones when compared with non-native listeners.
- Sensitivity to pitch mediated by musical training and natural psychoacoustic distance among stimuli.

Conclusions: Triqui glottal perception

- Glottal contrasts cued mainly by duration, but spectral tilt and pitch are weaker secondary cues.
- Listeners actively adapt to use the cues available in context, but such adaptations are sensitive to overall acoustic salience.
- The strength of the coarticulatory cues used by listeners mirror closely the strength of those cues in production.

Field speech perception

- Research in the field is challenging but can engage with existing debates within the speech perception literature.
- Exploratory research on cue weighting reveals adaptation to subordinate cues *only* when dominant cues are unavailable. Speech perception is an active, adaptive process (Heald and Nusbaum, 2014).
- Linguistic patterns in endangered and minority languages are a fertile ground for research on speech perception.
- Theories of speech perception are only generalizable to all languages if they are examined in relation to a diversity of languages.

Methodological constraints

Non-linguistic tasks (non-word-level perception) may be difficult for listeners.

Tasks involving literacy may be difficult for listeners.

Complex discrimination tasks like 2I2AFC or 4IAX, etc. are usually not possible.

But labelling and identification tasks are possible.

Future directions

1. Investigate the influence of contextual effects on attention to tonal cues.

2. Investigate the interaction of tone and prosody in Mixtec and Triqui corpus data using forced alignment DiCanio et al. (2013) with NSF grant #1360670.

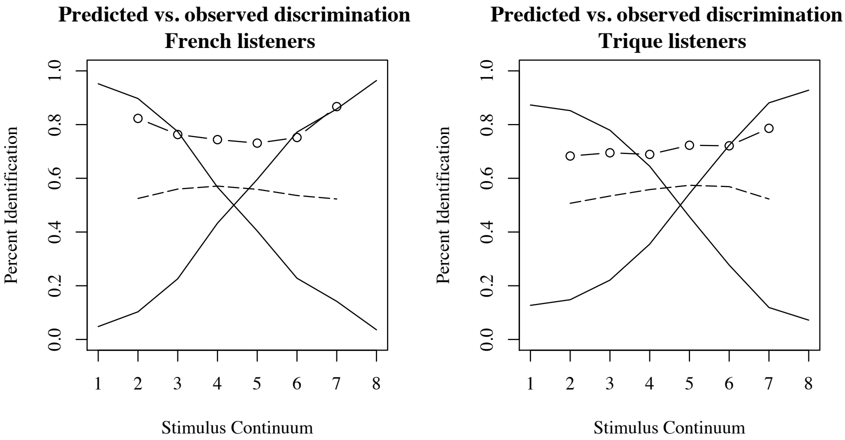
Acknowledgements

National Science Foundation
WHERE DISCOVERIES BEGIN

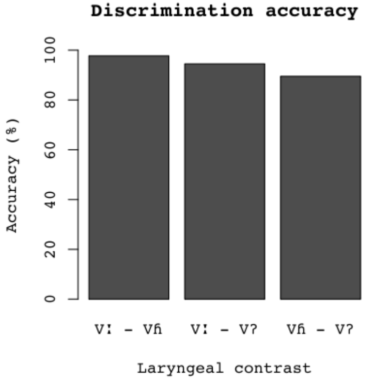
- The work was partly funded through a grant to Haskins Laboratories (Douglas Whalen, PI) on phonetic documentation in endangered languages.
- Laboratoire Dynamique du Langage / CNRS, François Pellegrino
- The Itunyoso Triqui community, kùruaa nihírèh! ([ku^{2}ɾu^{4}a^{43} ni^{3}ʔi^{4}ɾeʔ^{1}])!

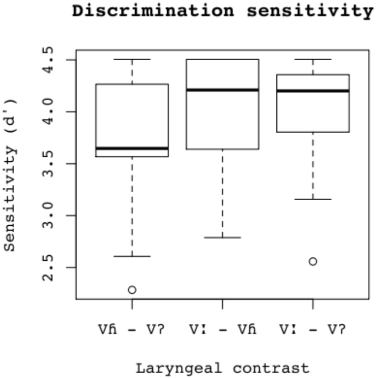
Thank you!

Predicted vs. observed discrimination accuracy



Good overall baseline identification of contrast





Abramson, A. S. (1979). The noncategorical perception of tone categories in Thai. In Lindblom, B. and Ohman, S., editors, *Frontiers of speech communication*, pages 127–134. London: Academic Press.

Avelino, H. (2003). Categorical Perception of Phonemic Tone in Yalálag Zapotec. In *Proceedings of the 15th International Congress of the Phonetic Sciences*.

Beddor, P. S. (2009). A Coarticulatory Path to Sound Change. *Language*, 85(4):785–821.

Beddor, P. S. and Krakow, R. A. (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *Journal of the Acoustical Society of America*, 106(5):2868–2887.

Best, C. T., Traill, A., Carter, A., Harrison, K. D., and Faber, A. (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. In Solé, M.-J., Recasens, D., and Romero, J., editors, *Proceedings of the 15th International Congress of the Phonetic Sciences*, pages 853–856. Barcelona: Causal Productions.

Bidelman, G. M., Weiss, M. W., Moreno, S., and Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience*, 40:2662–2673.

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, 30:163–191.

Boersma, P. and Weenink, D. (2016). Praat: doing phonetics by computer [computer program]. www.praat.org.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117(6):3890–3901.

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F., and Keller, P. (1996). Perception of lexical tone across languages: evidence for a linguistic mode of processing. In *Proceedings of the 4^{th} International Conference on Spoken Language Processing*, volume 4, pages 2514–2517.

Chan, S. W., Chuang, C.-K., and Wang, W. S.-Y. (1975). Cross-linguistic study of categorical perception for lexical tone. *Journal of the Acoustical Society of America*, 58:S119.

Cho, T. and McQueen, J. M. (2006). Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English Consonants. *Journal of the Acoustical Society of America*, 119(5):3085–3096.

Cruz, E. and Woodbury, A. (2005). El sandhi de los tonos en el Chatino de Quiahije. In *Memorias del Congreso de Idiomas Indígenas de Latinoamérica-II*. University of Texas, Austin.

Deutsch, D., Dooley, K., Henthorn, T., and Head, B. (2009). Absolute pitch among students in an American music conservatory: Association with tone language fluency. *Journal of the Acoustical Society of America*, 125(4):2398–2403.

DiCanio, C., Amith, J., and Castillo García, R. (2012). Phonetic alignment in Yoloxóchitl Mixtec tone. Talk Presented at the Society for the Study of the Indigenous Languages of the Americas Annual Meeting.

DiCanio, C., Nam, H., Whalen, D. H., Bunnell, H. T., Amith, J. D., and Castillo García, R. (2013). Using automatic alignment to analyze endangered language data: Testing the viability of untrained alignment. *Journal of the Acoustical Society of America*, 134(3):2235–2246.

DiCanio, C. T. (2008). *The Phonetics and Phonology of San Martín Itunyoso Trique*. PhD thesis, University of California, Berkeley.

DiCanio, C. T. (2009). The Phonetics of Register in Takhian Thong Chong. *Journal of the International Phonetic Association*, 39(2):162–188.

DiCanio, C. T. (2010). Illustrations of the IPA: San Martín Itunyoso Trique. *Journal of the International Phonetic Association*, 40(2):227–238.

DiCanio, C. T. (2012a). Coarticulation between Tone and Glottal Consonants in Itunyoso Trique. *Journal of Phonetics*, 40:162–176.

DiCanio, C. T. (2012b). Cross-linguistic perception of Itunyoso Trique tone. *Journal of Phonetics*, 40:672–688.

DiCanio, C. T. (2012c). The Phonetics of Fortis and Lenis Consonants in Itunyoso Trique. *International Journal of American Linguistics*, 78(2):239–272.

DiCanio, C. T. (2014). Cue weight in the perception of Trique glottal consonants. *Journal of the Acoustical Society of America*, 135(2):884–895.

Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the attainment of optimal phonological categorization*. PhD thesis, Utrecht University.

Esposito, C. (2006). *The Effects of Linguistic Experience on the Perception of Phonation*. PhD thesis, UCLA.

Esposito, C. (2010). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, 38:306–316.

Francis, A. L., Ciocca, V. C., and Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception and Psychophysics*, 65:1029–1044.

Gandour, J. and Harshman, R. (1978). Cross-language differences in tone perception. *Language and Speech*, 21(1):1–33.

Gerfen, C. and Baker, K. (2005). The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, 33:311–334.

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

Gottfried, T. L. and Beddor, P. S. (1998). Perception of temporal and spectral information in French vowels. *Language and Speech*, 31:57–75.

Hallé, P. A., Chang, Y. C., and Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3):395–421.

Harnsberger, J. D. (2001). The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics*, 29:303–327.

Heald, S. L. M. and Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience*, 8(35):1–15.

Hillenbrand, J. M. and Houde, R. A. (1996). Role of F_0 and Amplitude in the Perception of Intervocalic Glottal Stops. *Journal of Speech and Hearing Research*, 39:1182–1190.

Holt, L. L. and Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119(5):3059–3071.

Jamieson, A. R. (1977). Chiquihuitlan Mazatec Tone. In Merrifield, W. R., editor, *Studies in Otomanguean Phonology*, pages 107–136. Summer Institute of Linguistics, University of Texas at Arlington.

Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 60:286–319.

Keating, P. A. and Esposito, C. (2006). Linguistic voice quality. *UCLA Working Papers in Phonetics*, 105:85–91.

Kewley-Port, D., Pisoni, D. B., and Studdert-Kennedy, M. (1983). Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America*, 73(5):1779–1783.

Kingston, J. and Diehl, R. L. (1995). Intermediate properties in the perception of distinctive feature values. In Connell, B. and Arvaniti, A., editors, *Phonology and Phonetic Evidence: Papers in laboratory phonology 4*. Cambridge University Press.

Kirk, P. L., Ladefoged, P., and Ladefoged, J. (1984). Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics*, 59:102–113.

Kreiman, J., Gerratt, B., and Antoñanzas Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research*, 50:595–610.

Kreiman, J., Gerratt, B. R., and ud Dowla Khan, S. (2010). Effects of native language on perception of voice quality. *Journal of Phonetics*, 38(4):588–593.

Krishnan, A., Xu, Y., Gandour, J. T., and Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25:161–168.

Kühnis, J., Elmer, S., Meyer, M., and Jäncke, L. (2013). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: An EEG study. *Neuropsychologia*, 51:1608–1618.

Ladefoged, P., Maddieson, I., and Jackson, M. (1988). Investigating phonation types in different languages. In Fujimura, O., editor, *Vocal Physiology: voice production, mechanisms and functions*, pages 297–317. Raven Press, Ltd., New York.

Lee, Y.-S., Vakoch, D. A., and Wurm, L. H. (1996). Tone Perception in Cantonese and Mandarin: A Cross-Linguistic Comparison. *Journal of Psycholinguistic Research*, 25(5):527–542.

Lin, H.-B. and Repp, B. H. (1989). Cues to the perception of Taiwanese tones. *Language and Speech*, 32(1):25–44.

Lyon, J. (2008). A Perception Study of Glottalization in Gitksan Resonants. *manuscript*, University of British Columbia.

McGuire, G. L. (2007). *Phonetic Category Learning*. PhD thesis, The Ohio State University.

Mirman, D., Holt, L. L., and McClelland, J. L. (2004). Categorization and discrimination of nonspeech sounds: Differences between steady-state and rapidly-changing acoustic cues. *Journal of the Acoustical Society of America*, 116:1198–1207.

Nowak, P. M. (2006). The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics*, 34:139–152.

Peng, G., Zheng, H.-Y., Gong, T., Yang, R.-X., Kong, J.-P., and Wang, W. S.-Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38:616–624.

Pennington, M. (2005). *The Phonetics and Phonology of Glottal Manner Features*. PhD thesis, Indiana University.

So, C. K. (2006). Perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. In Warren, P. and Watson, C. I., editors, *Proceedings of the 11th Australian International Conference on Speech Science & Technology*. University of Auckland, New Zealand.

So, C. K. and Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech*, 53(2):273–293.

Stagray, J. and Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and nontone language. *Journal of Chinese Linguistics*, 21:143–163.

Thalin, A. (1980). Tlacoatzintepec Chinantec Syllable Structure. *SIL-Mexico Workpapers*, 4:1–8.

Wayland, R. and Jongman, A. (2003). Acoustic correlates of breathy and clear vowels: the case of Khmer. *Journal of Phonetics*, 31:181–201.

Whalen, D. H. (2004). How the study of endangered languages will revolutionize linguistics. In *Linguistics today: Facing a greater challenge*, pages 321–342. Amsterdam: John Benjamins.

Whalen, D. H. and McDonough, J. (2015). Taking the laboratory into the field. *Annual Review of Linguistics*, 1:395–415.

Xu, Y., Gandour, J. T., and Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, 120(2):1063–1074.