



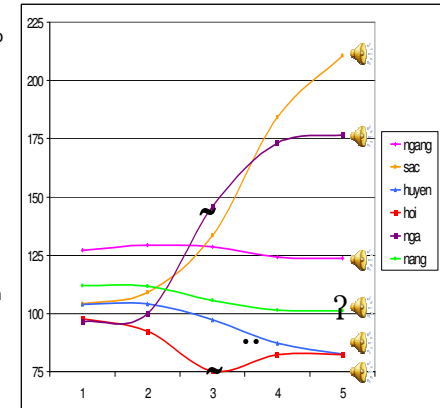
Tone perception in Vietnamese dialects

Marc Brunelle
University of Ottawa

TIE2, ZAS, Berlin
September 7th, 2006

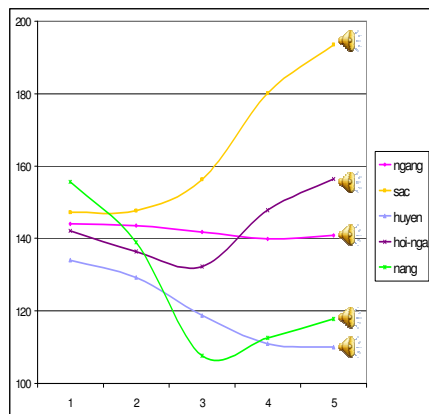
Acoustics of Northern Vietnamese tones

- Standard dialect
 - Everybody is exposed to it through TV and radio
 - Northern settlers in all regions of the country
- Well-described instrumentally
 - Han 1969
 - Han and Kim 1974
 - Earle 1975
 - Vũ 1982
 - Nguyễn and Edmondson 1998
 - Brunelle 2003
 - Michaud 2004
 - Vũ, Alessandro and Michaud 2005



Acoustics of Southern Vietnamese tones

- Regional standard
 - Can be used in speeches and in the media
 - Most speakers adopt some northern features in formal contexts
- Few instrumental descriptions (Trần 1967, Gsell 1980, Vũ 1982)
 - No quantitative data on voice quality



Models of tone features in VN

- Proposed tone features
 - **Voice quality:** Laryngeal Effect (Earle 1975), Phonation (Hoàng 1986), Register (Phạm 2001)
 - **Contour:** Rising (Earle 1975), Modulé (Gsell 1980), Concave, Contour (Ngô 1984), Level, Simple or Complex contour (Phạm 2001)
 - **Pitch height:** High (Earle 1975, Gsell 1980, Burton 1992, Alves 1997), Low, All in one reg. (Gsell, 1980, Hoàng 1986), Register (Burton (1992)
 - **Duration:** Short (Alves 1997)
- Based on acoustic evidence and controversial phonological processes
- Most models assume that all VN dialects have the same tonal features and representation as NVN
 - Except Earle (1975) and Gsell (1980)
- **We need more perceptual evidence**
 - Bits and pieces in Gsell (1980), Vũ (1981), Nguyễn and Ingram (2005)

Perception, features and dialectal diversity

- How much intelligibility across dialects in the absence of semantic context?
 - Experiment 1
- What are the perceptual cues that are used for identification in each of the two major dialects?
 - Experiment 2
- Can perceptual cues be equated with features?

Experiment 1

Identification of tones across dialects

Perception of natural tones

- 36 natural stimuli
 - 6 “real” tones uttered on syllable /ma/ by a male native speaker of Hà Nội Vietnamese
 - 6 repetitions each
- Real stimuli mixed with resynthesized stimuli and presented to subjects during a more comprehensive experiment
 - Experiment 2

Subjects

- 40 subjects between 20-35
 - 20 Northerners
 - Both parents from the North
 - Living in Hà Nội for at least 5 years
 - 10 “Pure” Southerners
 - Both parents from the South
 - Living in Hồ Chí Minh City for at least 5 years
 - 10 “Mixed” Southerners
 - At least one parent from Northern or Central Vietnam
 - Born and raised in the South
 - Southern accent (at least when talking to me)
 - Living in Hồ Chí Minh City for at least 5 years

Procedure

- Forced-choice identification task in Praat

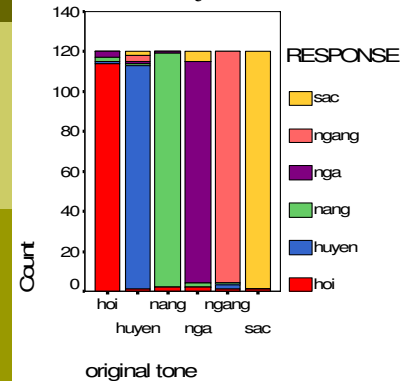
- Tone (real words)
- Goodness rating (1 to 5)

- One subject at a time on a laptop computer with Sennheiser headphones
 - Practice run with 10 tokens



Results - Northerners

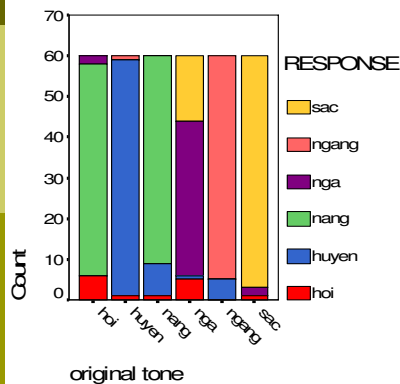
20 Subjects



- Few errors: More than 95% of tones identified correctly in every tone category

Results – Southerners with southern parents

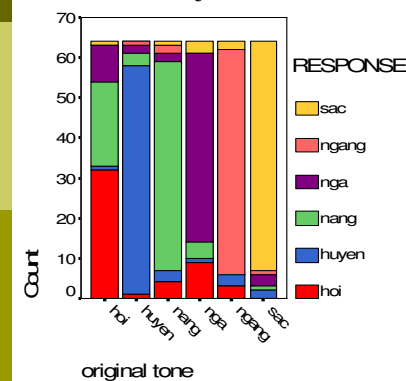
10 Subjects



- NVN **hoi** is low, falling-rising and glott: SVN **nặng**, ignoring glott.
- NVN **nặng** is low and glott: SVN **huyền**, ignoring glott.
- NVN **ngã** is rising and glott: SVN **sắc**, ignoring glott.
- SVN **hoi-ngã** merged

Results – Southerners with at least one non-southern parent

10 Subjects



- Fewer misidentifications
 - Same errors as other Southerners, but lower prop.
- More idiosyncratic patterns of misidentification




Cross-dialectal perception - summary

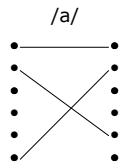
- "Pure" Southerners often identify NVN glottalized tones as their non-glottalized equivalents
 - But they are not totally glottalization-deaf (media?)
- "Mixed" Southerners identify NVN tones more accurately (but more idiosyncratic mistakes)
 - Exposure to more dialectal diversity
- Tones in isolation are often confused by listeners of other dialects. However, in real life:
 - Subjects can rely on non-tonal cues to determine a speaker's dialect and adjust their perception if they are familiar with it.
 - Rarely any possible confusion if word classes and semantic context are taken into account.

Experiment 2

Perceptual cues

Resynthesized stimuli

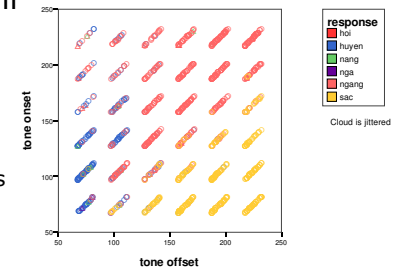
- Stimuli resynthesized from 3 instances of NVN syllable /ma/
 - Modal voice (from ngang) 
 - Glottalization in middle part of the tone (from hỏi) 
 - Glottalization at the end of the tone (from nặng) 
 - Durational differences neutralized
- Pitch resynthesis
 - Simple contours (36)
 - Two targets: onset and offset
 - 6 pitch heights: 75, 105, 135, 165, 195, 225
 - Complex contours: all falling-rising (5)
 - Three targets: onset, midpoint, offset
 - 3 pitch heights: 75, 165, 225
- 123 stimuli X 6 repetitions = 738 tokens (± 55 min.)



Presentation of results

- For each group of subjects, we must look at 6 types of stimuli

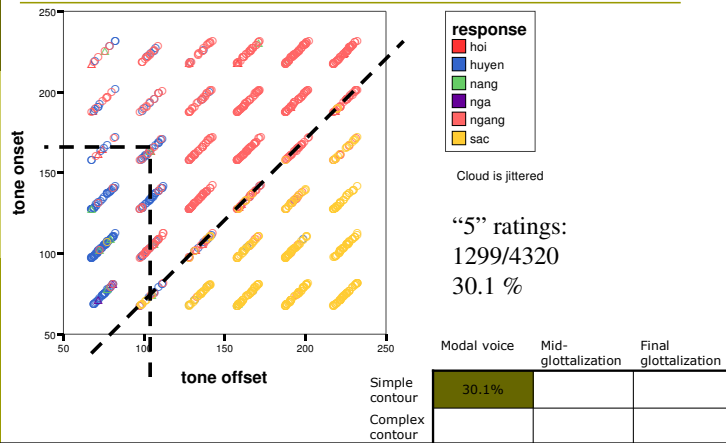
- Simple contours
 - Modal
 - Mid-glottalization
 - Final glottalization
- Complex contours
 - Three voice qualities in single chart



- For the sake of simplicity, I will only report on tokens rated "5"

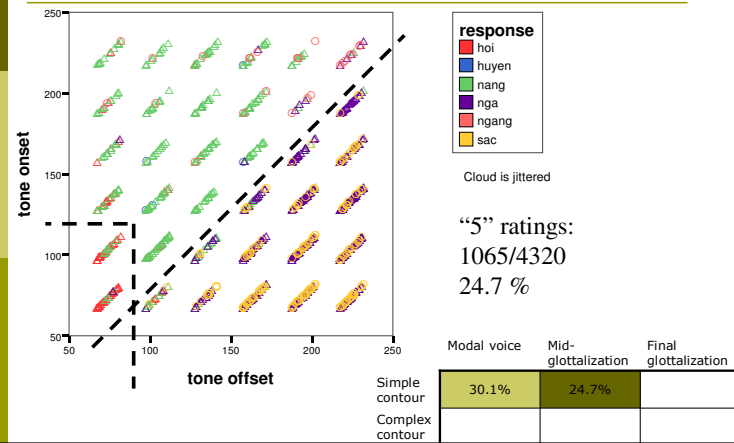
Perception – Northerners

Modal voice – simple contour



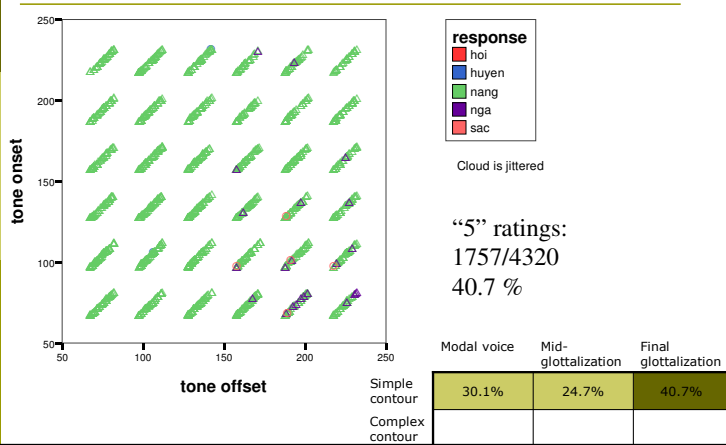
Perception – Northerners

Mid-glottalization – simple contour



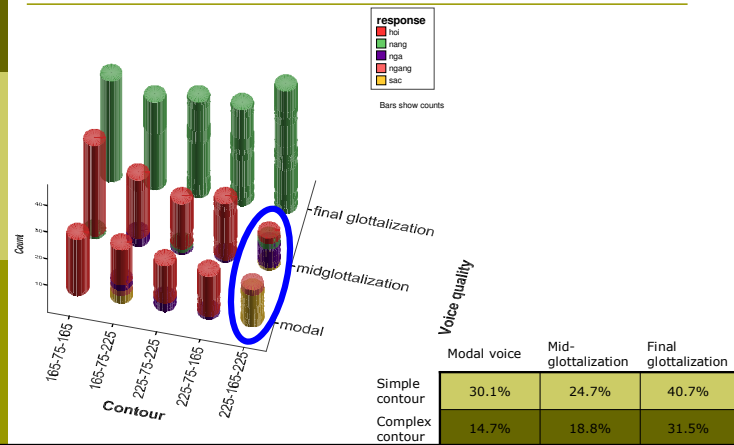
Perception – Northerners

Final glottalization – simple contour



Perception – Northerners

Complex contours – three voice qualities



Northern Vietnamese perceptual cues

	Modal	Mid-glottalized	Final Glottalization
Falling-Rising	Hỏi		Nặng
Low	Huyền	Hỏi	
Rising	Sắc	Ngã	
Else	Ngang		

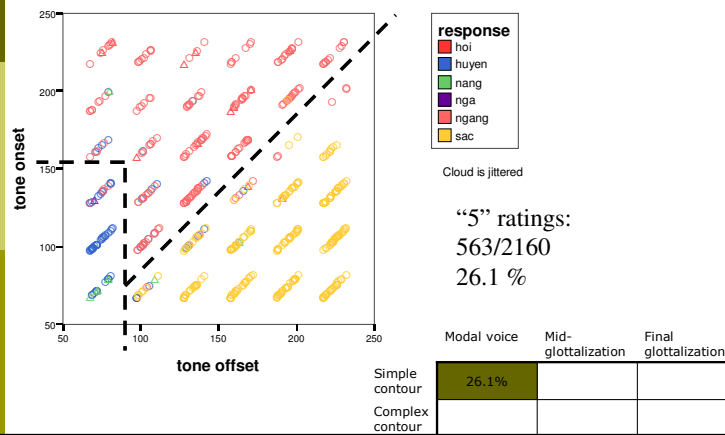
NVN perceptual cues

Falling-rising: Formal ~ Elicitation Hỏi

	Modal	Glott.
Low	Huyền	Hỏi
Rising	Sắc	Ngã
Else	Ngang	Nặng

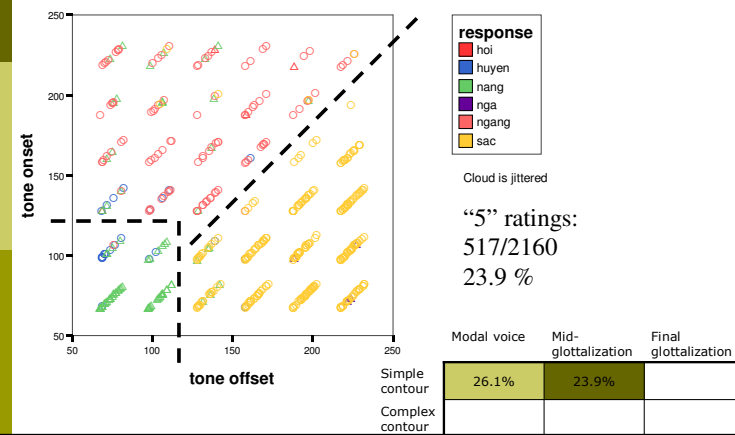
Perception – “Pure” Southerners

Modal voice – simple contour



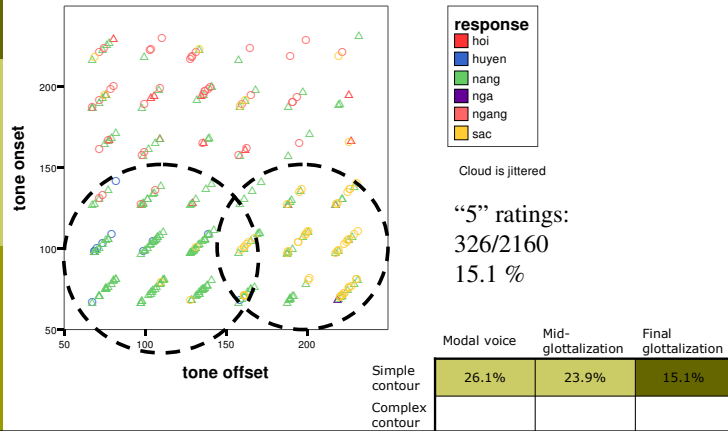
Perception – “Pure” Southerners

Mid-glottalization – simple contour



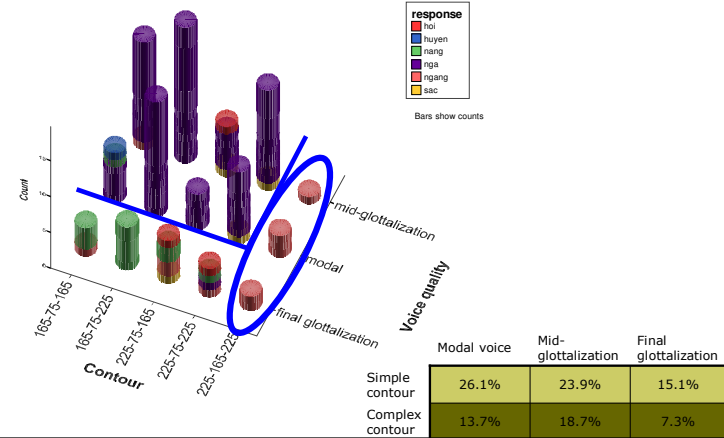
Perception – “Pure” Southerners

Final glottalization – simple contour



Perception – “Pure” Southerners

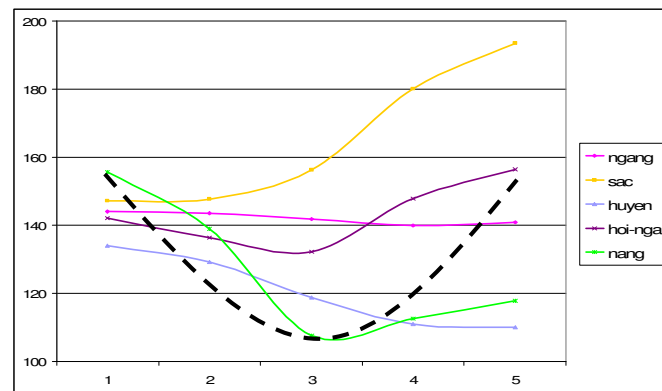
Complex contours – Three voice qualities



Southern Vietnamese Perceptual Cues

	Modal	Mid-glottalized	Final glottalization
Falling-Rising	Hỏi-Ngã		
Low	Huyền	Nặng	
Rising	Sắc		
Else	Ngang		

Falling-Rising Tones in SVN



SVN perceptual cues summarized

	Simple Contour	Falling-Rising
Low	Huyền	Nặng
Rising	Sắc	Hỏi-Ngã
Else	Ngang	

What about “mixed” Southerners?

- Overall “Mixed” Southerners use the same perceptual cues as other Southerners
 - Slightly higher sensitivity to glottalization
- More idiosyncratic variation in responses
 - Origin of parents was not systematically controlled for

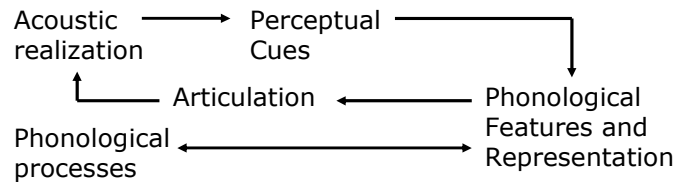
Perceptual cues ~ Tone features ??

Tone	NVN	SVN
Ngang	[]	[]
Huyền	[low]	[low]
Sắc	[rising]	[rising]
Nặng	[glott.]	[low, curve]
Hỏi	[low, glott.] <i>[curve] in form. speech</i>	[curve]
Ngã	[rising, glott.]	

Things to keep in mind when proposing a model of tone features...

- Perception tells us which elements of the acoustic signal are important, but perceptual cues and features are not trivially identical.
 - Perception translates fine acoustic detail into coarse phonological categories
 - Ex. 1: A pitch rise could be analyzed as [H]
 - Ex. 2: A low pitch at rime offset could be interpreted as [L], but not a low pitch at rime onset
- Phonological evidence is crucial (Phạm 2001)
 - Spreading, sandhi > Evidence about natural classes
 - However, Vietnamese mostly has “dead” tonal phonology
 - Reduplication: Productive or not?
 - Poetry: Codification or phonology?
 - Treatment of borrowings: Dependant on a theory of markedness
- Still unclear if tone features are articulatorily (Bao 1999, Duanmu 2000) or acoustically invariant (tacit assumption)
 - Features could be relatively abstract

Phonological processes, perceptual invariance and variation in representation



Learners acquire features that:

- Can be correctly inferred from perceptual cues
- Generate proper acoustic cues via articulation
- Account for phonological processes

As long as these conditions are met, features and representations can vary across speakers

Previous models of VN tone features in the light of perceptual cues

- All models except Earle (1975) and Gsell (1980) assume identical tone features for all dialects
- Their features refer directly to acoustic properties, but make incorrect predictions about:
 - Voice quality
 - Contours
 - Pitch height
- We must propose new models for Vietnamese tone features
 - More perceptual work
 - Duration, breathiness, more complex contours...
 - Test the productivity of phonological processes experimentally
 - Articulatory evidence is needed to determine if we are dealing with acoustic or articulatory invariance

Conclusions

- Perception across dialects
 - Without context, there is a significant amount of confusion
 - Knowledge of the standard dialect affects perception of speakers of non-standard varieties
 - Exposure to diverse dialects leads to more "tolerance" to variation
- Perception and features
 - NVN and SVN do not use the same perceptual cues
 - Glottalization in NVN, complex contours in SVN
 - Perceptual cues impose restrictions on tone representation
 - Unlikely that tone features are identical in NVN and SVN

References

- Alves, M. (1995). "Tonal Features and the Development of Vietnamese Tones." *Hawai'i Working Papers in Linguistics* 27: 1-13.
- Bao, Zhiming (1999). *The Structure of Tone*. Oxford, Oxford University Press.
- Brunelle, M. (2003). Coarticulation Effects in Northern Vietnamese. *Proceedings of the 15th International of Phonetic Sciences Barcelona*: 2673-2676.
- Burton, Strang (1992). Reduplication and the representation of Vietnamese tone, ms., Brandeis.
- Duanmu, San (2000). *The Phonology of Standard Chinese*. New York, Oxford University Press.
- Earle, M. A. (1975). *An Acoustic Phonetic Study of Northern Vietnamese Tones*. Santa Barbara, Speech Communications Research Laboratory Inc.
- Gsell, René (1980). Remarques sur la structure de l'espace tonal en Vietnamien du Sud (parler de Saigon). *Cahiers d'Études Vietnamiennes*, 4, 1-26.
- Han, M. (1969). *Studies in the Phonology of Asian Languages VIII: Vietnamese Tones*. Los Angeles, University of Southern California.
- Han, M. and K.-O. Kim (1974). "Phonetic variation of Vietnamese tones in disyllabic utterances." *Journal of Phonetics* 2: 223-232.

References - 2

- Hoàng, C. C. (1986). "Suy Nghi Thêm về Thanh Điều Tiếng Việt." Ngôn Ngữ **3**: 19-38.
- Michaud, A. (2004). "Final Consonants and Glottalization: New Perspectives from Hanoi Vietnamese." Phonetica **61**: 119-146.
- Ngô, T. N. (1984). The Syllabeme and Patterns of Word Formation in Vietnamese. Linguistics. New York, New York University: 465.
- Nguyễn, V. L. and J. Edmondson (1997). "Tones and voice quality in modern northern Vietnamese: Instrumental case studies." Mon-Khmer Studies **28**: 1-18.
- Phạm, A. H. (2001). Vietnamese Tone: Tone is not pitch. Linguistics. Toronto, University of Toronto.
- Thompson, L. (1965). Vietnamese Reference Grammar. Seattle, University of Washington Press.
- Trần, H. M. (1967). Tones and Intonation in South Vietnamese. Series A - Occasional Papers #9, Papers in Southeast Asian Linguistics No.1. Đ. L. Nguyễn, H. M. Trần and D. Dellinger. Canberra, Linguistics Circle of Canberra.
- Vũ, N. T., C. d'Alessandro, et al. (2005). Using open quotient for the characterization of Vietnamese glottalised tones. Interspeech, Lisbon.
- Vũ, T. P. (1982). Phonetic Properties of Vietnamese Tones across dialects. Papers in Southeast Asian Linguistics. D. Bradley. Sydney, Australian National University. **8 - Tonation**: 55-75.