Tone slips in Cantonese: Evidence for early phonological encoding

John Alderete, Queenie Chan, Henny Yeung
Simon Fraser University

Thanks!
Data collectors: Gloria Fan, Tsz Ying Ng, Macarius Chan
Comments/assistance: Paul Tupper, Stephen Matthews

Slides available at: anderei.net

Tone languages

Contrastive tone - talkers must choose correct tone.

Cantonese for ‘two’

yi [22]

Cantonese tones

22	Low Level
33	Mid Level
55	High Level
23	Low Rising
25	High Rising
21	Low Falling

Tone languages

Contrastive tone - talkers must choose correct tone.

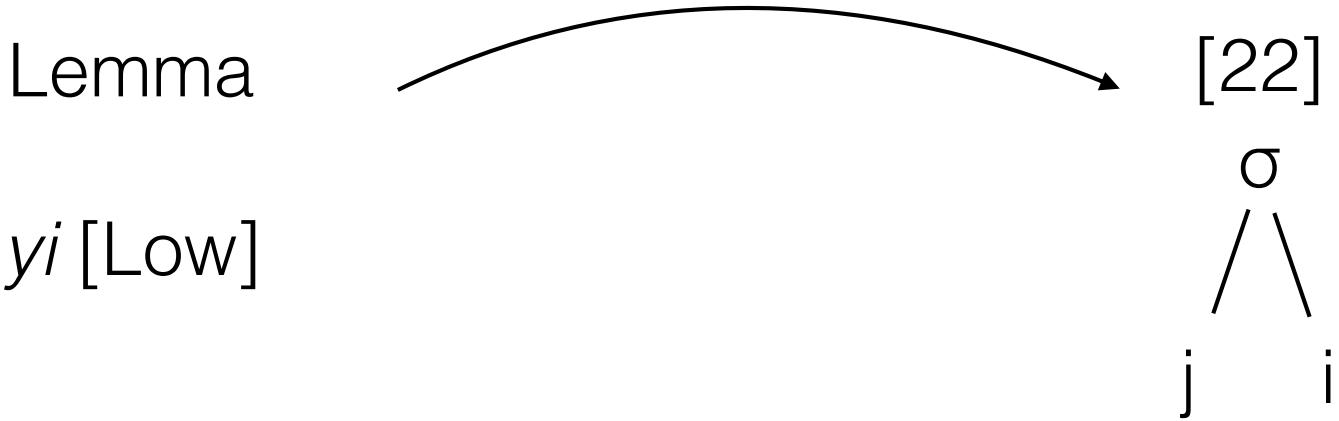
Cantonese for ‘two’

yi [22]

Cantonese tones

22	Low Level
33	Mid Level
55	High Level
23	Low Rising
25	High Rising
21	Low Falling

Phonological encoding



Focus

Key question: is tone actively selected in phonological encoding, or inherited from lemma level?

Two speech error studies of Mandarin

Yes
Wan & Jaeger (1998)

yi [Low]

22
33
55
23
25
21

j i [22]

j i [low level pitch]

No
Chen (1999)

yi [Low] **Lemma**

j i [Low] **Phonological Encoding**

j i [low level pitch] **Articulation**

Focus

Key question: is tone actively selected in phonological encoding, or inherited from lemma level?

Yes
Wan & Jaeger (1998)

Tone is encoded like segments.

Tone errors are relatively common.

Errors are contextual

Same kinds of speech errors as segments (anticipations, perseverations, etc.)

No
Chen (1999)

Tone is encoded like metrical stress.

Tone errors are exceedingly rare, like stress errors.

Apparent tone errors have alternative analyses

Implicit priming studies - atonal syllables

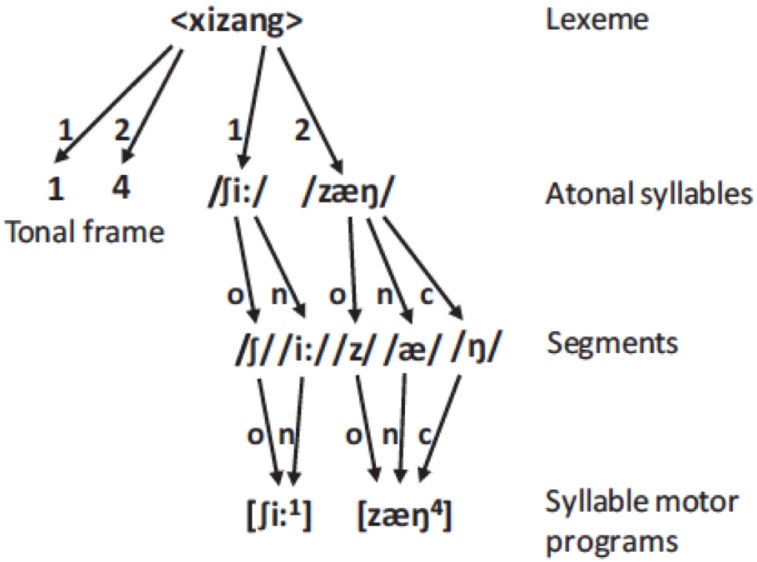
Parallels

Active debate

From Roelofs (2015)

Tone is inherent, not selected

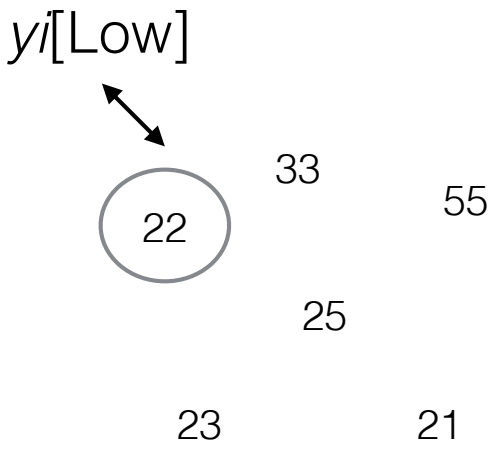
- Kember et al. (2015): tone is more stable, less prone to error
- Tone mapped to prosodic frame, like lexical stress in Weaver++ (Levelt et al. 1999)



Tone is selected in form encoding

- Wan (2006) on Mandarin, Liu & Wang (2008) on Taiwanese: tone errors rather common, compare in many ways to segmental errors
- Tone is selected via an activation dynamics similar to segments

Selecting tone



Today's talk

Data: tone slips in Cantonese from the SFU Speech Error Database (SFUSED)

Frequency: are tone errors rare or not?

Prediction: non-trivial occurrence → actively selected

Interactivity effects: is encoding of tone interactive in that it is not independent of the encoding of other linguistic elements (other sound or word structure)?

Prediction: interactivity → actively selected

Methods

Data source: Simon Fraser University Speech Error Database (SFUSED)

Data collection

- Collected offline from audio recordings (roughly 32 hours of unscripted speech)
- Collected by 4 native speakers of Cantonese, 2 listened to each recording
- Basis for selection: high production quality, balance of men/women, natural conversations with long stretches of unscripted speech.

Training

- Phonetic training: trained and tested on transcribing Cantonese speech
- Listening tests: three 30 minute recordings in English, prescreened for errors, given feedback on submitted and missed errors

Verification and classification

- Verification: data analyst checks if submitted errors meet the definition of a speech error (unintended, non habitual deviation from speech plan)
- Classification: valid errors assigned one of 60 variables that cross-classify the error into traditional taxonomies of errors, including those of Asian languages (Stemberger 1993, Dell 1986, Wan & Jaeger 1998).
- Data quality: better sample coverage than other studies (error about every 45 seconds)

More on methods: ***Alderete & Davies 2018, Language and Speech***

Error types in SFUSED Cantonese

Objective: use large database of Cantonese speech to probe encoding of tone.

Error type	Example	Count
Sound substitution	mai23 → bai23 ‘rice’	1,153
Sound addition	uk55 → luk55 ‘house’	110
Sound deletion	si22jip22 → si22ji_22 ‘career’	90
Tone substitution	hei33kek22 → hei23kek22 ‘drama’	435
Complex sound errors	jyn21tsyn21 → jyn21dzyn33 ‘completely’	316
Phonetic errors	sy55 → si-y55 ‘book’	70
Morphological errors	baːt33gwaː33geŋ33 → baːt33gwaː33ø	26
Lexical errors	jiŋ55man25 ‘English’ (lei22man25 ‘Italian’)	245

Observation: tone slips are not rare at all in Cantonese.

Error types in SFUSED Cantonese

Objective: use large database of Cantonese speech to probe encoding of tone.

Error type	Example	Count
Sound substitution	mai23 → bai23 ‘rice’	1,153
Sound addition	uk55 → luk55 ‘house’	110
Sound deletion	si22jip22 → si22ji_22 ‘career’	90
Tone substitution	hei33kek22 → hei23kek22 ‘drama’	435
Complex sound errors	jyn21tsyn21 → jyn21dzyn33 ‘completely’	316
Phonetic errors	sy55 → si-y55 ‘book’	70
Morphological errors	baːt33gwaːɜ33geŋ33 → baːt33gwaːɜ33ø	26
Lexical errors	/jiŋ55man25 ‘English’ (lei22man25 ‘Italian’)	245

Second most common type

20.4% all sound errors

Observation: tone slips are not rare at all in Cantonese.

Re-examining Chen (1999): turns out that this study has a relatively small number of sound errors in general, but tone errors are not at all uncommon as a percentage of sound errors: roughly 15% of sound errors, cf. 13% from Wan and Jaeger (1998)

Majority of tone errors are contextual

Observation: the majority of tone slips (76%) are contextual in the sense that there is a nearby syllable with the intruder tone (four syllable envelope).

Anticipatory activation

gam25jim23 /**dou33** jan21 **ge33** ‘affect other people’
(Intended: **dou25**)
一個凝聚力,咁亦都感染 /到 人^嘅

Interpretation: if tone is selected in phonological encoding, we expect tone slips to be anticipatory or perseveratory, just like segments (also majority are contextual in SFUSED Cantonese).

Interactivity

Interactive spreading effects (e.g., Dell 1986)

- Interactivity effects are a hallmark of phonological encoding.
- Higher incidence of an error due to shared structure; stems from nature of activation dynamics in an interconnected lexical network.

Example: repeated phoneme effect (Dell 1984, MacKay 1970)

Deal Beak has greater chance of *d* → *b* error than ***Deal Bock***
[i] [i] [i] [a]

Rationale for tone

- If tone is selected in phonological encoding, expect the same kinds of interactive spreading effects found for segments and words.
- Wan & Jaeger (1998): greater than chance probability that word substitutions share a tone is a kind of interactivity effect.

Interactivity: Phonological substitutions

Illustration:

source	intended — ***same tone***
… dzau**22**	da:**22** …
	↓
	dza:**22**

Finding: segmental substitutions where intended and source syllables share a tone (green below) are over-represented.

	Shared	Not Shared
Expected	105.35	526.65
Observed	149	483

$$X(1) = 21.703,\ p < 0.00001$$

Tone of syllable w/source

Tone of syllable w/intended

	22	33	55	23	25	21
22	46	18	30	14	22	12
33	25	19	21	11	13	12
55	20	31	48	7	21	18
23	13	5	8	5	10	4
25	34	16	14	11	17	15
21	19	12	21	11	15	14

Details:
-interacts with tone type
-factor in tone frequency
-[22] and [55] show strong effect, others do not

Interactivity: Word substitutions

	intended		error	
Illustration:	dzoŋ**22**	→	dzau**22**	*share same tone*

Limitation: insufficient data to investigate interactivity for individual tones

Findings:

- Word substitutions in monosyllable words (*n=45*) have a great than chance probability of sharing a tone, as in Mandarin (Wan & Jaeger 1998)
- Disyllabic words harder to interpret, but in the same direction.

Lexical substitutions in mono-syllabic words

	Shared	Not Shared
Expected	7.5	37.50
Observed	13	32

$X(1) = 4.84, p < 0.0278$

Interactivity: Phonological similarity

Phonological similarity (e.g., Shattuck-Hufnagel & Klatt 1979)
Phonological similar sounds slip more often than dissimilar sounds.

Example: more slips of /p/ and /f/ (both voiceless labials) than /p/ and /r/.

Phonological similarity and phonological encoding
Phonological similarity is generally assumed to result from feedback from features to segments in phonological encoding (e.g., Dell 1986).

> Similarity effect is also a hallmark of phonological encoding (or articulation, cf. inner speech).

Prediction
If tone is actively selected in phonological encoding, expect more slips with similar tones than dissimilar tones.

Similarity effect, cont'd

Finding: there is a significant correlation between similarity and confusability in tone confusion matrix. The more similar, the more likely two tones are, the more likely to swap.

Example: 70 substitutions with 22/33, only 13 of 22/55

$r = 0.562$, $p = 0.0437$ (simulated, 5000 permutations in a Mantel test)

Intruder tone

Intended tone	22	33	55	23	25	21
22		37	7	25	18	26
33	33		7	16	16	6
55	6	17		0	13	2
23	16	9	7		18	11
25	20	20	20	15		1
21	32	5	2	14	0	

How similarity calculated?
-no obvious feature system
-phonetic distance, using Chao system

Summary of findings

1. Tone errors are not rare in Cantonese

2. Most tone errors are contextual

3. Encoding of tone is interactive

    Word substitutions

    Phonological substitutions

    Similarity effects

Summary of findings

Parallels with Segments

1. Tone errors are not rare in Cantonese

Segmental common type of speech error in most corpora

2. Most tone errors are contextual

Most segmental errors are contextual (Nooteboom 1969)

3. Encoding of tone is interactive

Word substitutions

Malapropisms (Fay and Cutler 1977, cf. Wan & Jaeger 1998)

Phonological substitutions

Repeated phoneme effect (Dell 1984, Mackay 1970)

Similarity effects

Phonological similarity effect (Shattuck-Hufnagel 1979)

Summary of findings

	Parallels with Segments
1. Tone errors are not rare in Cantonese	Segmental common type of speech error in most corpora
2. Most tone errors are contextual	Most segmental errors are contextual (Nooteboom 1969)
3. Encoding of tone is interactive	
Word substitutions	Malapropisms (Fay and Cutler 1977, cf. Wan & Jaeger 1998)
Phonological substitutions	Repeated phoneme effect (Dell 1984, Mackay 1970)
Similarity effects	Phonological similarity effect (Shattuck-Hufnagel 1979)

Conclusion: encoding tone is part of phonological encoding, and requires a selection mechanism similar to that for segments.

Contributors to SFUSED

Director/analyst/data collector:
John Alderete

Research associates
Paul Tupper (SFU)
Alexei Kochetov (Toronto)
Stefan A. Frisch (USF)
Monica Davies (UBC)
Henny Yeung (SFU)
Queenie Chan (SFU)

Analysts/data collectors
Holly Wilbee (English)
Monica Davies (English)
Olivia Nickel (English)
Queenie Chan (Cantonese)
Macarius Chan (Cantonese)
Heikal Badrulhisham (English)

Data collectors
Jennifer Williams (English)
Julie Park (English)
Rebecca Cho (English)
Bianca Andreone (English)
Dave Warkentin (English)
Crystal Ng (Cantonese)
Gloria Fan E (Cantonese)
Amanda Klassen (English)
Laura Dand (English)
Jane Li (English)
Gavin Tam (English)

Relative rates of errors: segments vs tones

Simply not the case that tone errors are so rare that they can be compared to stress. No study has shown this.

But are they relatively rare, and therefore require a different analysis, distinct from the typical activation dynamics that is used for consonants and vowels? (See Kemper et al. 2015 experimentally induced errors.)

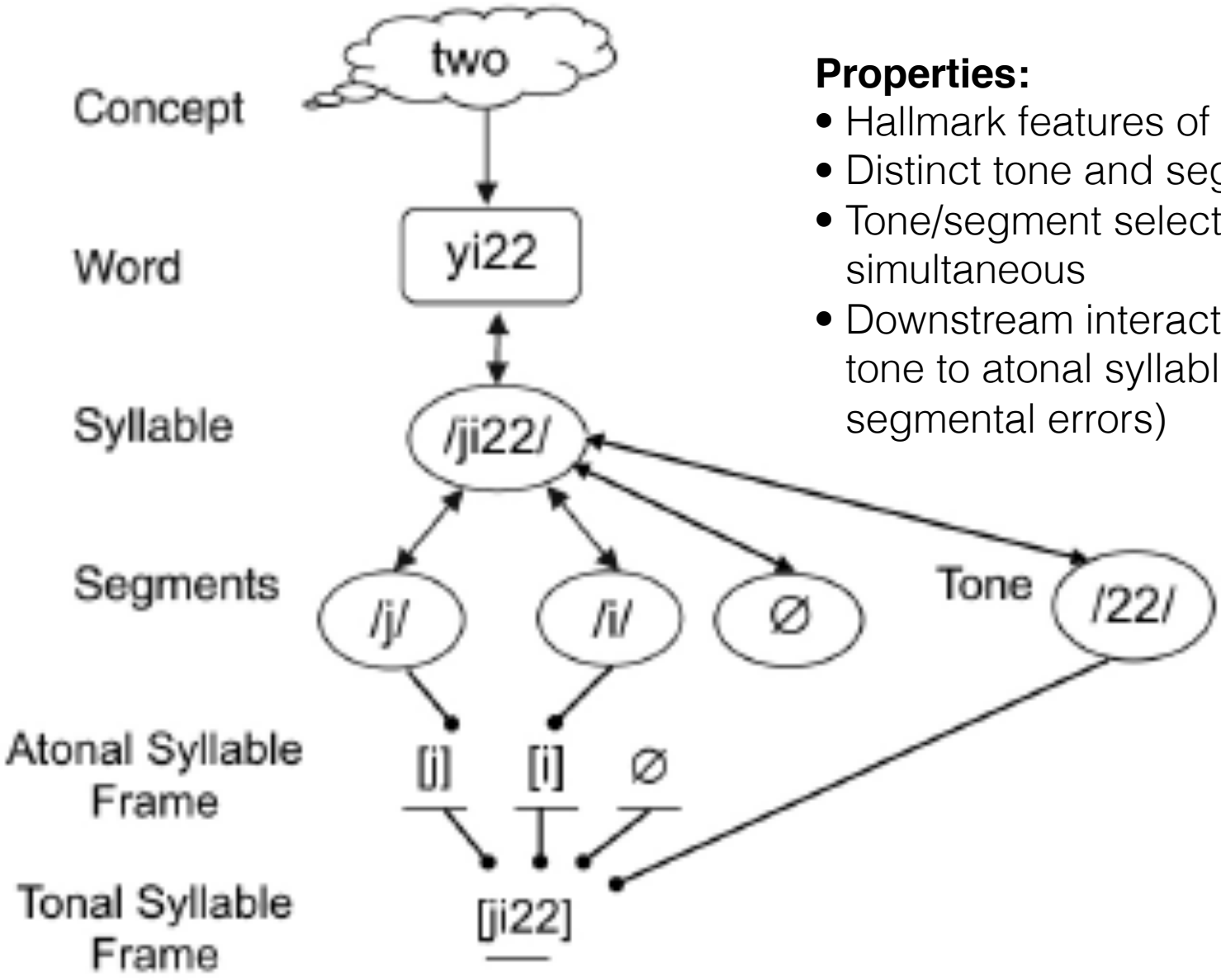
Single unit errors (ignores rare clusters, e.g., CC or VC)

Unit	Error frequency	% of total	Inventory size
C	714	49.58%	19
V	304	21.11%	10
Tone	422	29.31%	6

Claim: many factors are at work

- Consonants may be disproportionally affected by the word-initial bias
- Consonants are often selected for twice in a syllable, as opposed to once.
- Tone is a different selection problem because smaller inventory

How fit with Proximate Unit Hypothesis?



Precedent: O'Seaghda et al. 2010

Properties:

- Hallmark features of Proximate Unit
- Distinct tone and segment selection
- Tone/segment selection roughly simultaneous
- Downstream interactions from mapping tone to atonal syllable (more tone + segmental errors)

Tone errors feed tone sandhi

Observation: tone errors are ‘early’ in the sense that they are the inputs to later tone sandhi rules.

Wan & Jaeger 1999

Mandarin tone sandhi:[21] [21] →**[35]** [21]

Illustration of tone slip

/na35 [jow21 maj51] pɑw51-tʂɨ21/ →

na35 [**jow35 maj21]** pɑw51-tʂɨ21

Where is the place selling newspapers?’

maj51 → maj21 provides context for regular sandhi

Theoretical result:

- If tone slips occur early in phonological encoding, they will feed tone sandhi.
- If they are part of later Articulation processes, don’t expect feeding. Indeed, tone is implemented after tone sandhi processes in Chen’s (1999) and Roelofs 2015 model.

A concern: what about the mergers?

Change in progress

We know that there are several tone mergers going on right now in Cantonese.

Mergers: 23/25, 33/22, 21/22

Would these affect your results?

A serious issue

These mergers are between similar tones, so need a methodology that ensures slips of similar tones are truly errors and not just mergers in the productions of talkers or perceptions of data coders.

Facts

- The general impression of our data analysts is that the principal talkers in the recordings do not have the mergers, and so their productions of these similar tones are distinct. Actually removed some errors that could have been due to mergers
- This impression is supported by the corrected tone errors. 37 tone errors are corrected by the talker (8.5%, cf. 8.2% of corrected phonological substitution errors). Roughly half of these corrections are between two tones that participate in the mergers, e.g., Intended: 22, Error: 33, corrected to 33. Very strong indication that two tones are distinct in the mind's of the talker.
- Expect much higher rates: if higher O/Es for similar tones really due mergers, expect _much_ higher rates that we observe; tone errors happen very very rarely, tone mergers happens several times a sentence for merging talkers (need to refine exact numbers).

Tone encoding completely not encapsulated

Simple vs. complex sound errors:

- Tone errors are the second most common type of error (20% of all phonological errors)
- Complex errors involving both tone and segmental mis-selections are rather common: jyn21 → jy_33 (deletion + tone change)

Observations:

- Tone errors have a greater likelihood of being complex than segmental errors (e.g., deletion + substitution of a segment)
- A little over a third of all tone errors are complex, but one 1/6th of segmental errors are complex
- Of the 316 sub-lexical complex errors, 75% involve tone.

	Simplex	Complex
Phonological errors	1795 (85.52%)	304 (14.48%)
Tone errors	435 (64.83%)	236 (35.17%)

chi2 goodness of fit:
$X(1)2 = 137.35$, $p < .0001$