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

Spring 1991: Paris 3 report card

INSTRUCTOR NARRATIVE EVALUATIONS

FOR JOHN ALDERETE

Education Abroad Program - France

SPRING 1991

Course: LIG-117Y, Intro to Ling 9 units Grade: A

Instructor: Leroy

COURSE EVALUATION: Participation, presence and quality of work were entirely satisfactory. The questions and the comments by this student in class made evident his sound knowledge of the French language as well as a sound understanding of the analytic procedures discussed during the course of the year.

Course: LIG-137, Intro to Semiology 4.5 units Grade: A-

Instructor: Kristeva

Tone languages

Contrastive tone - talkers must select tone for every syllable

Cantonese: [ji22]
Task: correctly select tone [22] from inventory of six tones.

Questions about nature of tone speech planning

- Are tones selected at all, or are they features of morphemes?
- If selected, are they selected separately from segments, or as part of a sequence, e.g., a rhyme?
- How is the selection of tone ordered with respect to the selection of consonants?

Principal goal for today

Contrast two positions: early and interactive encoding of tone, or late tone processing with non-interactive encoding.

"Early encoding"
Non-independent selection
Part of phonological encoding
(e.g., Wan & Jaeger 1998)
~Tone encoding like segments

Evidence:
-abundant contextual errors
-interactivity effects
-tone sandhi

"Late encoding"
Independent encoding
No part of phonological encoding
(e.g., Chen 1999, Roelofs 2015)
~Tone encoding like stress

Evidence:
-almost no tone errors
-tone alone not enough for implicit priming effect

Contextual tone errors

Contextual errors: standard analysis is that an intruder sound is due to elevated activation levels of the same sound in a near-by word.

Example:

… 下呢個 /jy21ho**22** … ^/van**22**用# 呢個...

intruder tone

source tone

Intended: jy21**ho21** 如何

Result of activation dynamics: anticipatory activation of [22] downstream leads to higher overall activation of intended tone on *ho*, leading to [21] > [22] error.

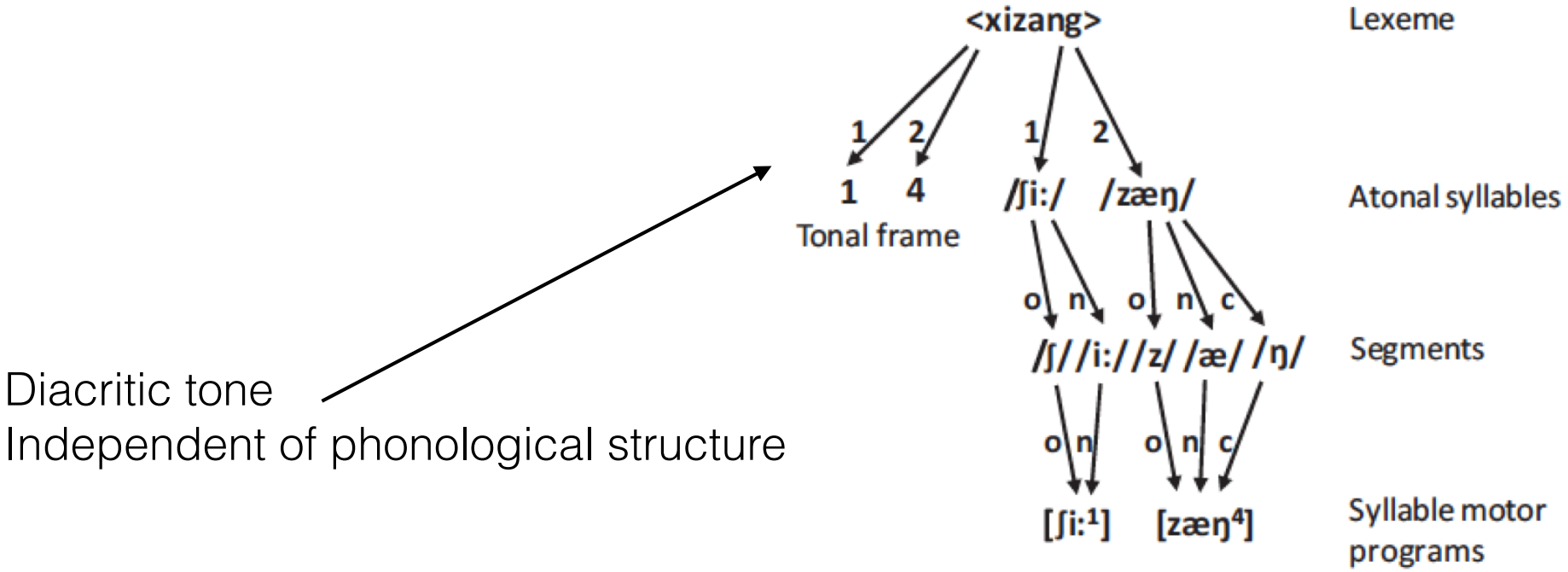
Is tone part of phonological encoding?

Yes! Wan & Jaeger 1998, Moser 1991, Shen 1993
Speech errors anticipate and perseverate tone, just like what we find with segments.

No! Chen 1999, see also Roelofs 2015
Tone errors are like stress, exceedingly rare and open to alternative analysis. Should be treated like stress, diacritic in phonological encoding (non-interactive).

Roelofs 2015, WEAVER++ approach

After Chen 1999, assumed that tone is represented with a direct mapping that is the result of a selection process; like stress that is diacritically represented in metrical frame, and implemented later in articulatory processing.



Prediction: like stress, virtually no tone errors.

Motivation: Are tone errors really rare?

Problem it does not seem correct to say that tone errors are like stress errors in that they are exceedingly rare. When tone errors are compared to a baseline of total sound errors, they have a non-negligible frequency in all corpora.

	Chen 1999	Wan & Jaeger 1998	SFUSED Cantonese
Language	Mandarin	Mandarin	Cantonese
# of tones	4	4	6
Total errors	987	785	2,535
Sound errors	150	597	2,203
Tone errors	24	78	418
% Sound errors	15%	76%	87%
Tone errors as % of sound errors	16%	13%	19%

Motivation: 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.

How does the facts of Cantonese speech errors contribute to the early / late encoding debate?

SFU Speech Error Database (SFUSED)

Current languages

SFUSED English (~10,000 errors)

SFUSED Cantonese (2,535 errors)

Goals

- Build a multi-purpose database that documents the rich structure in spontaneous speech errors.
- Examine how the structure of non-Indo-European languages impacts language production processes
- Projected date of release to general public: 2019

Methods

Data sources:

- 50 podcasts from three lifestyle and entertainment podcasts
  Roughly 32 hours of unscripted speech, excluded commercials etc.
- Basis for selection: high production quality, balance of men/women, natural conversations with long stretches of unscripted speech.

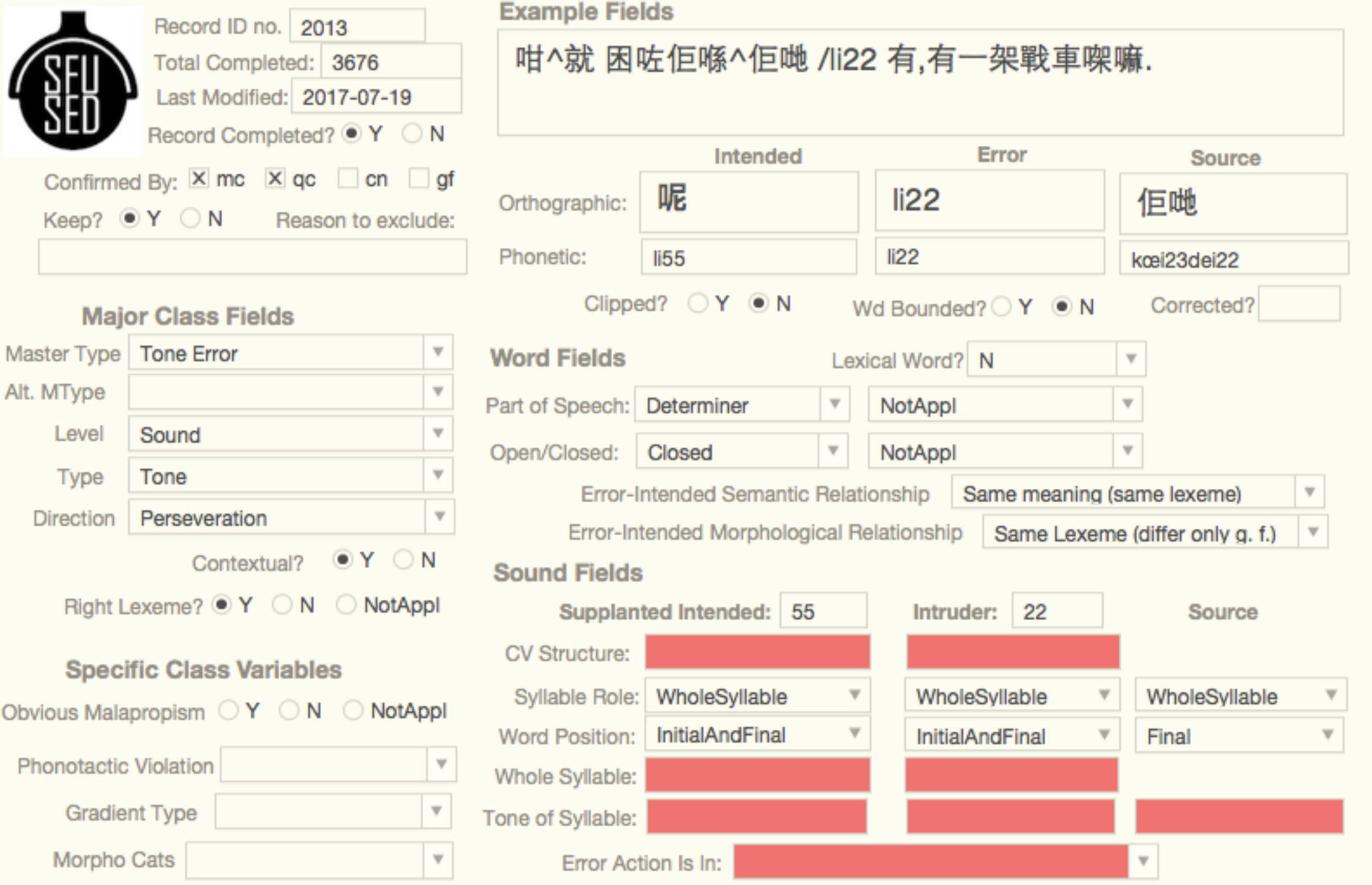
Training

- Data collectors: three undergraduate linguistics students, one graduate student, all bilingual in English and Cantonese
- Phonetic training: bone up on phonetic skills, tested with transcription tests of Cantonese speech
- Introduction to speech errors, an hour collecting in their daily lives
- Listening tests: three 30 minute recordings in English, prescreened for errors, given feedback

Data collection, verification, and classification

- Submissions: each podcast examined by two data collectors, submit errors in spreadsheet, batch imported into database
- Verification: data analyst examines all submitted errors, checks that they meet the definition of a speech error (unintended, non habitual deviation from speech plan (Dell 1986); checked for dialectal and idiolectal variants and causal speech phenomena; many exclusions
- Classification: valid errors assigned one of 60 variables that cross-classify the error into traditional taxonomies of errors (Stemberger 1993, Dell 1986).
- Data quality: better sample coverage than other studies (error about every 45 seconds), and has structure expected of representative sample (e.g., low percentage of exchanges)

Illustration: tone error in SFUSED Cantonese



Overview of SFUSED Cantonese

Sublexical errors (90.12%)	*N*	Examples
Phonological errors		
Phonological substitution	1,153	mai23 → **b**ai23 'rice'
Phonological addition	110	uk55 → **l**uk55 'house'
Phonological deletion	90	si22jip22 → si22ji_22 'career'
Phonological exchange	3	li55 di55 → **d**i55 **l**i55 'these'
Phonological shift	1	tsœt55hœi33 → tsœ**i**t55hœ33 'to go out'
Phonological tone substitution	435	hei33kek22 → hei**23**kek22 'drama'
Complex set of processes	316	jyn21tsyn21 → jyn21**dzyn33** 'completely'
Other sublexical errors		
Sequential blends	37	lei23 jiu33 → liu23 'you must'
Phonetic errors	70	sy55 → **s**i-**y**55 'book'
Morphological errors	26	ba:t33gwa:33geŋ33 → ba:t33gwa:33___ 'feng shui mirror'
Word and phrase errors (9.88%)		
Lexical substitutions	85	kœi23 ge33 /**jiŋ55man25** 'his English' (Intended: ji33da:i22lei22man25 'Italian')
Role mis-selections	14	/**ŋo23** wa:22 'I said' (Intended: kœi23 'he')
Word additions	43	gei25 /**jat55** dyn22 jam55ŋok22 'several *one segments of music'
Word deletions	42	lei23 /___ gok33dak55 'you think that' (Intended deleted word: wui23 'will')
Word blends	30	la:m21paŋ21jau23, la:mjan25 → **la:m21paŋ21jan25** 'boyfriend, man'
Word shifts	9	ham22 /**dzo25** jap22 /∅ 'fell into'
Phrasal blend	2	hou25 loi22, hou25 do55 lin21 tsin21 → **hou25loi22lin21tsin21** 'for a long time, many years ago'
Complex set of processes	20	pei33jy21 → bey33 'for example'
Total all errors	2,479	

Cantonese tone

- All syllables have one of six tones, no tone sandhi.
- Tones broken down by register and type.
- Tonal targets assumed in Chen 2000

Register	Type	Tone targets	Example (with tone suffix)
high	level	H	wan55 'warm' (cf. wan25), also: wat5 'twisted'
high	level	M	wan33 'to shut/lock up', also: wa:t3 'to dig
high	rising	MH	wan25 'to look for'
low	level	L	wan22 'to transport', also: wat2 'pit (of fruit)'
low	rising	LM	wan23 'to allow'
low	falling	ML	wan21 'cloud'

Details

- For some speakers, level high [55] in free variation with [53], but not important in our data.
- Level tones have shorter versions (e.g., [5]) in syllables closed by an unreleased stop, but assume these have the same tone structure (a conservative assumption, because raised chance probability of a shared tone).

Tone errors

Observations:

- Tone errors are the second most common type of error (20% of all phonological errors)
- Complex errors (tone and segmental) are rather common.

All tone errors

Single tone substitutions	422 (62.89%)	miŋ21 → miŋ**22** 'understand'
Double tone substitutions	12 (1.79%)	dzik22dou33 → dzik**55**dou**55** 'until'
Tone blends	1 (0.15%)	gam55lin25, gam25tsiŋ21wan22 → gam**45**= 'this year, relationship luck'
Tone + segmental error	236 (35.17%)	juŋ22 → **dz**uŋ**33**

Complex errors

Substitution	199	jyn21tsyn21 → jyn21**dz**yn**33** 'completely'
Deletion	15	jyn21 → jy_**33** 'finish'
Addition	5	ŋo23 → ŋo**ŋ33** 'I'
Exchange	1	duk22dak22 … foŋ55fa:t33 → duk22d**a**:k**22** … foŋ55f**at22** 'unique … way'
Phonetic error	5	jau23 → **je-au25** 'to have'
Substitution and Deletion	6	goŋ25 gan25 → go_**25** _a**ŋ55** 'talking about'
Substitution and Addition	3	ji23geŋ55 → ji**ŋ33**giŋ55 'already'
Deletion and Addition	2	git33gwo25 → gi_**22**gwo**ŋ55** 'result'

Complex vs simple tone errors

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, $p < .0001$

Interim summary

- Selection of segments and tone seems to be independent, and therefore require a distinct selection process.
- But tone encoding not completely encapsulated from segmental encoding; if so, would expect complex errors to be roughly parallel

Patterns of contextual tone errors

Question: how common are contextual tone errors?

Observation: contextual errors are in the majority.
With 4 syllable window (more stringent that Nooteboom's (1969) standard 7 syllable), 76% are contextual, which compares to the frequency of segmental contextual errors (62%)

Type	7 σ	4 σ	
Perseverations	99 (23.24%)	115 (27%)	tsœt55saŋ55 lin21 ^jyt22 ^jat22 /si**22** 'year, month, day and time of birth' (Intended: si21)
Anticipations	93 (21.83%)	107 (25.12%)	gam25jim23 /dou**33** jan21 ^ge33 'affect other people' (Intended: dou25)
Perseveration + Anticipation	186 (43.66%)	101 (23.71%)	doŋ25bat55^dzy22 ^/hou**22**wan22 'unstoppable luck' (Intended: hou25wan22)
Exchange	1 (0.23%)	1 (0.23%)	dzuŋ22jiu33 → dzuŋ33jiu22 'important'
Non-contextual	47 (11.03)	102 (23.94%)	go33 /ji25saŋ55 dzau22wa:22 'the doctor said' (Intended: ji55saŋ55)

Importance: contextual errors are standardly analyzed as mis-selections caused by higher activation levels in neighbouring form elements in phonological encoding; supports early encoding.

Interactive spreading effects

Analysis of interactive spreading effects (e.g., Dell 1986)
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*

Explanation: because two words share the vowel [i], increases the flow of activation between two words, which increases chances of *d* being realized as *b* (*Beak* increases activation of *b*).

Prediction

- If tone is encoded early, expect the same kinds of interactive spreading effects found for segments and words.
- Wan & Jaeger 1999: greater than chance probability that word substitutions share a tone is a kind of feedback effect.

Interactivity: Phonological substitutions

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

Interactivity: Word substitutions

Findings:

- Word substitutions in monosyllable words (*n=45*) have a great than chance probability of sharing a tone, as in Mandarin (Wan & Jaeger 1999) (GoF: X(1)2 = 4.84, *p* = 0.0278)
- Disyllabic words harder to interpret, but in the same direction.

Monosyllables					
	Shared	Not Shared			
Expected	7.5	37.50			
Observed	13	32			
Disyllables (all observations)	Syllable 1			Syllable 2	
	Shared	Not Shared		Shared	Not Shared
Expected	4.33	21.67		4.33	21.67
Observed	15	11		13	13
Disyllables (no shared characters)	Syllable 1			Syllable 2	
	Shared	Not Shared		Shared	Not Shared
Expected	2.5	12.5		3.17	15.83
Observed	4	11		6	13

Interactivity: Phonological similarity

The fact of phonological similarity (e.g., Shattuck-Hufnagel & Klatt 1979) A cross-linguistically robust fact of segmental errors is the intended and intruder sounds tend to be phonologically similar; substitutions involving a few feature changes are much more common that substitutions with more feature changes.

Caveat: no similarity effect with inner speech (Oppenheim and Dell)

Phonological similarity and phonological encoding
Phonological similarity is generally assumed to result from encoding of features in phonological encoding (e.g., Dell 1986). Segments that shared many features are linked to the same feature nodes in the node network, and these nodes provide feedback to similar sounds, resulting in a higher occurrence of similar sounds in errors.

Predictions

- If tone is encoded phonologically, and tone is linked to a set of features, tones should slip with similar tones.
- If tone is not encoded, there should be no similarity effect.

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
-level tones don't have onset and offset

Independence of tone for selection

Question: is selection of tone independent of segments, or bound to some sub-syllabic unit?

=> *independent*

Independence of tone and segmental errors
Most segmental errors don't involve tone, and most tone errors don't involve segments, so need separate mechanism

Sequential blends
Some blends retain tone on a rhyme that is different from the intended rhyme, e.g., /lei23 jiu33/ > **liu23** 'you must'.

Illicit tone+syllable combinations
Checked syllables ending in /p t k/ are restricted to level tones, but a non-trivial number of tone errors result in a contour syllable on a checked syllable, e.g., fat35jin22 'to discover', which rules out retrieval of stored tone-syllable combinations.

Summary of results

Evidence for early phonological encoding

- Non-trivial amount of tone errors
- Most tone errors are contextual (require activation dynamics)
- Feedback effects:
  - Word substitutions
  - Phonological substitutions
  - Similarity effects
- Tone errors feed tone sandhi (must precede Articulation)

Evidence that tone and segments selected separately

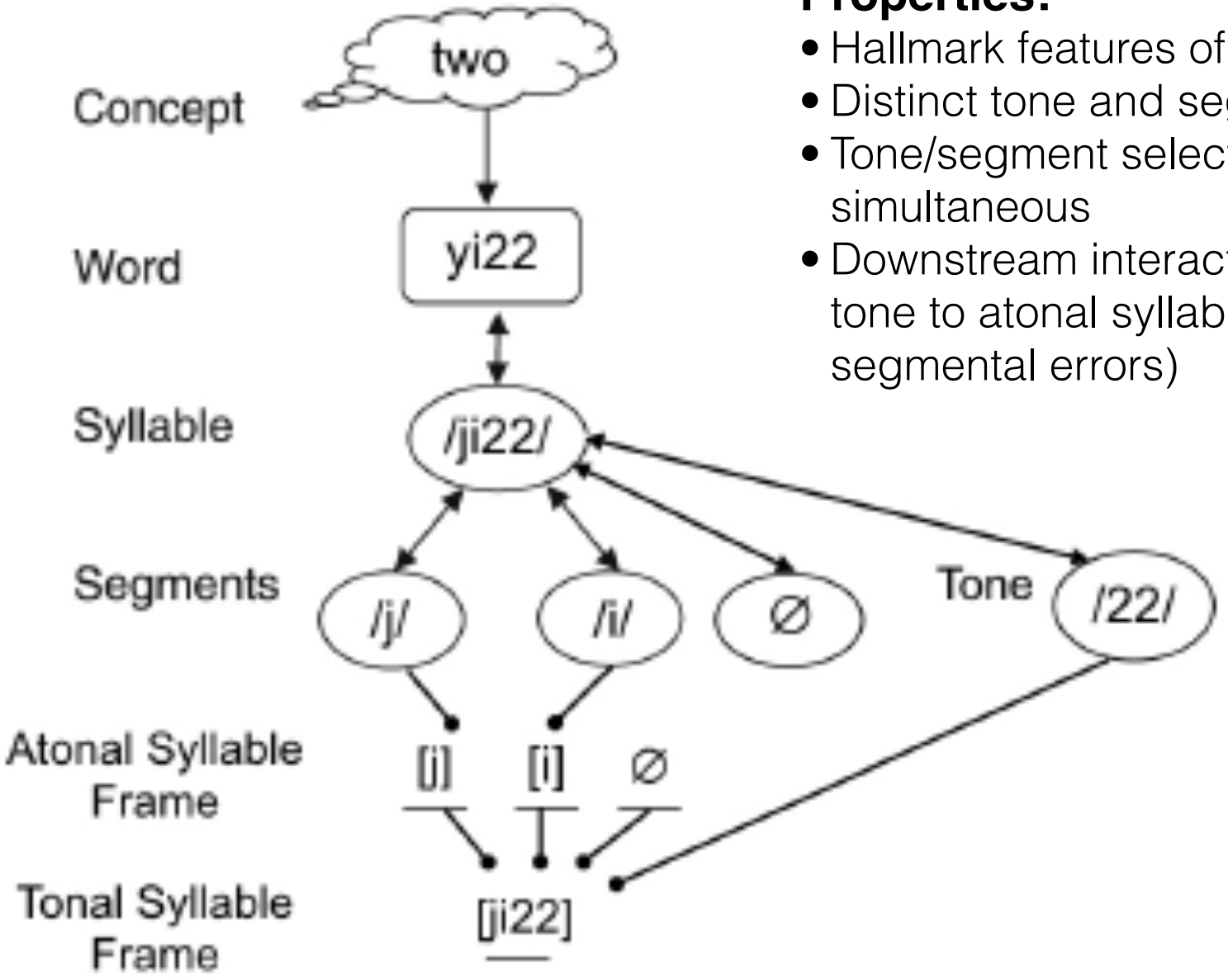
- Dominance of one mis-selection without the other
- Sequential blends where tone is realized on a rhyme other than the intended rhyme, e.g., lei23 jiu33 → liu23 'you must'
- Complex errors skewed towards tone + segmental errors
- Tonotactic gaps: tone slips result in illicit tonal syllables

Sketch of a model

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)



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)

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

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)

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

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).