Oct. 24, 2018 AUDI Series, UBC

Speech errors and phonological patterns

Integrating insights from psycholinguistic and linguistic theory

John Alderete, Simon Fraser University

in collaboration with:

Queenie Chan (SFU), Monica Davies (UBC), Paul Tupper (SFU), Henny Yeung (SFU)

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

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

Speech errors tend to respect grammar

“First law” of tongue slips (due to Wells 1951)
Speech errors respect phonotactic constraints, only produce legal phonological combinations (Boomer & Laver 1968, Nooteboom 1967, Garrett 1980)

Phonological constraints active in repairs (Fromkin 1971: 41)
play the victor → *flay the pictor* (exchange of *p* and *v*, *vl* → *fl*)

Syntactic regularity in speech errors (Garrett 1980, Bock 2011)
- Category constraint (word substitutions respect part of speech labels), producing licit but unintended sentences.
- Sentence blends, role mis-assignments, and spurious agreement relations tend to respect grammar too.

Bock 2011: 332
“The most striking thing about attested syntactic errors is that, like other kinds of speech errors, they occur within a structural matrix that resists modification.”

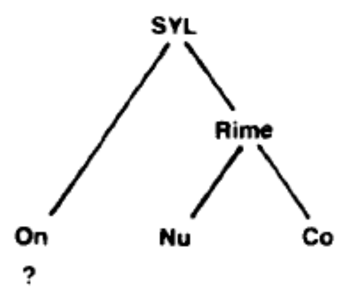
Regularity as a hard constraint

Early models: grammatical regularity not really explained, but the result of a built-in design features

Spreading-interactive model of Dell 1986

- Mental lexicon: activation dynamics for selecting valid linguistic units
- Syllable template: productive capacity for language (simple sentence trees, word trees, and syllables. (a.k.a. ‘structural matrix’)

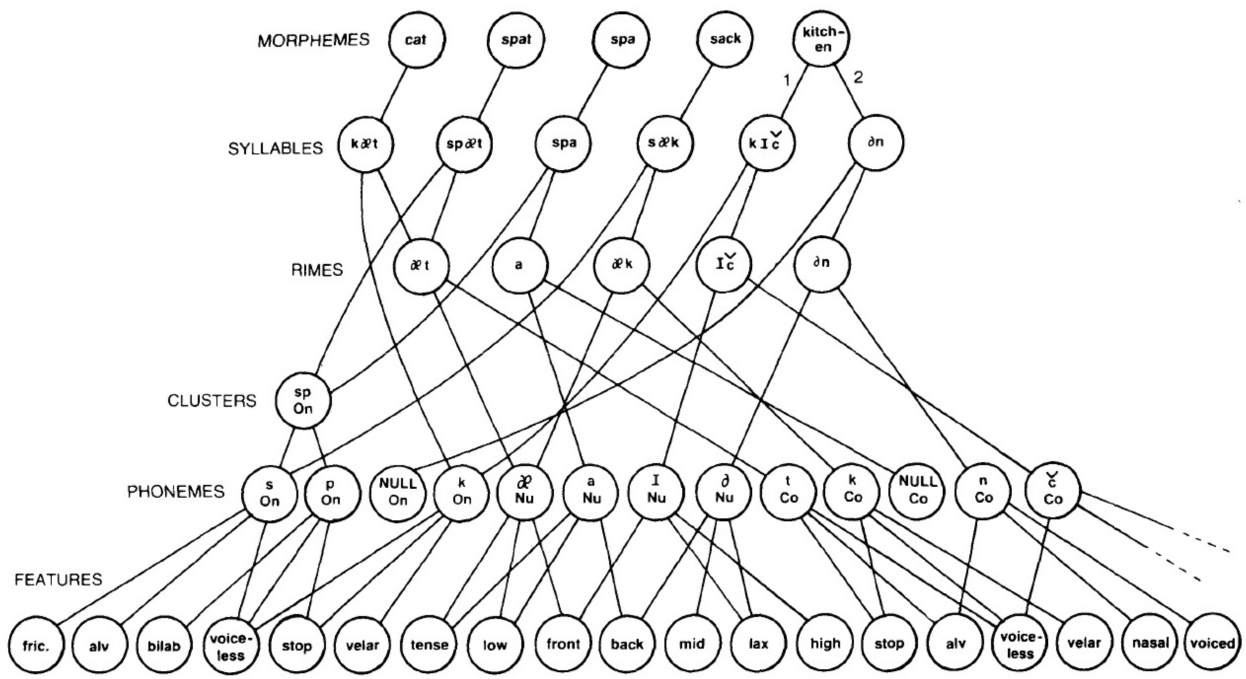
Syllable template



Result: *vl* is not a valid onset for intended word *play* because [*vl*, Onset] is not a node in the mental lexicon.

Clear role for grammar

Mental lexicon (fragment)



But regularity is not a hard constraint

Stemberger (1983): phonological regularity is very high, but speech errors do violate English phonotactics approximately 1% of the time (37 cases out of 6,300 examples)

Speech errors that violate phonotactics (p. 32):

… in the first floor /**dl**orm —- dorm room

I /**sth**ough —- thought I said 'moff'

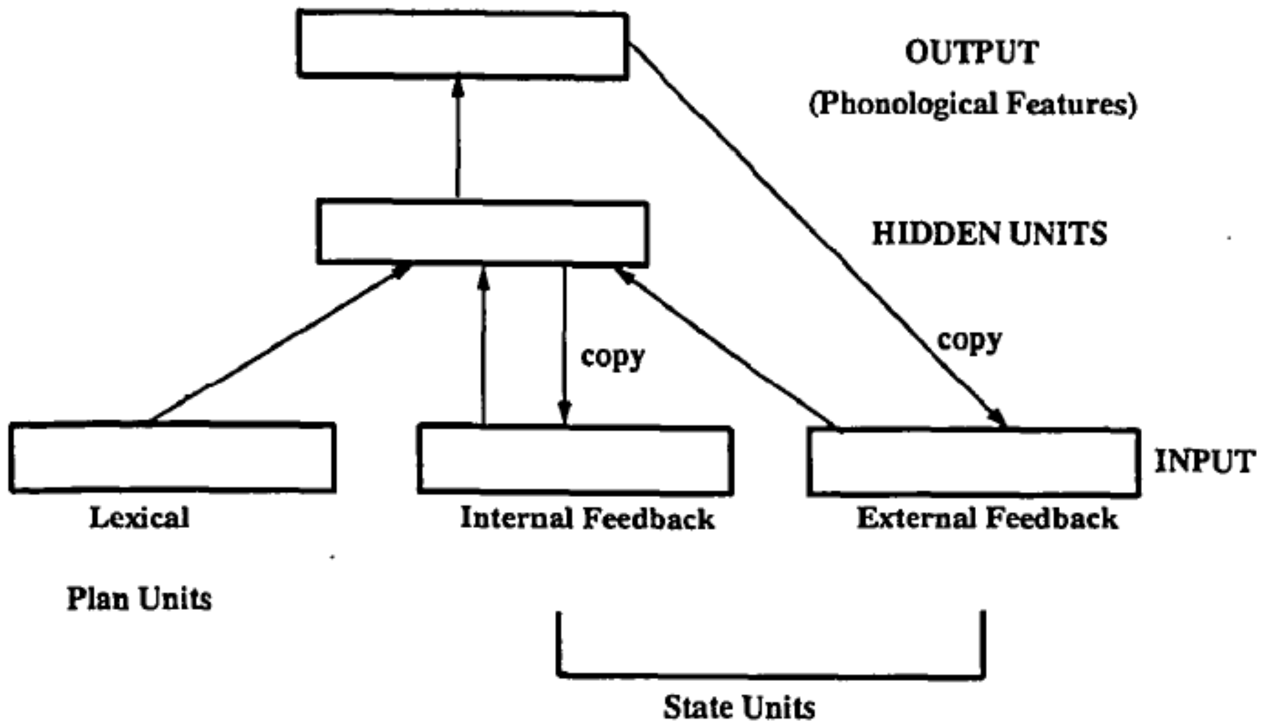
… knowledge of the cooperative /**rp**in—- principle

Problem: require a model that accounts for overwhelming phonological regular, and yet can produce phonotactic violations

Issue (Stemberger 1983: 32-33): "… we cannot conclude that these constraints exist as some sort of rule"; phonotactic constraints could be either ***actively applied*** (c.f. syllable templates) or ***passively applied*** through independently necessary production mechanisms sensitive to frequency.

Regularity without syllable templates

Dell et al. (1993): computational model of phonological encoding that tried to account for phonological regularity as a consequence of underlying production processes. Simple recurrent network (sequential network, recurrent with two state layers, distributed representations), trained on sample of small English words, then tested on a set of phonological patterns typical of speech error collections (regularity, word-onset effect, syllable constituent effect).



Result: given certain parameters (trained on frequent vocabulary, internal and external input), the model produces errors that are **phonotactically regular about 96.3% of the time**.

Competing and complementary explanations

Question: Just how much phonological grammar is there in phonological encoding? (Dell et al. 1993, Fromkin 1971, Goldrick 2002, Romani et al. 2002, Stemberger 1983)

Phonological regularity (speech errors obey phonotactics)
- Linguistics: segments slotted into syllable templates
- Psycholinguistics: learned frequencies of sequences

Segment substitution patterns (sound errors involving segments)
- Linguistics: phonological markedness prefers unmarked segments
- Psycholinguistics: type frequency gives bias for frequent segments

Planning units (what linguistic units manipulated in speech errors)
- Linguistics: features, segments, sub-syllabic units (onsets, rimes), syllables
- Psycholinguistics: primacy of segments and syllables, downplay prosody

Larger objective: integrate insights of both linguistics and psycholinguistics in an empirically grounded model of phonological encoding

Objective: *combine insights from both phonological and psycholinguistic theory in an integrated model of phonological encoding*

SFUSED Simon Fraser University Speech Error Database	Investigate phonological patterning in a large, methodologically rigorous speech error database
Phonological regularity	Investigate phonotactic violations in SFUSED and discuss model implications for syllable templates and frequency based solutions
Consonant substitutions	Investigate competing explanations from cross-linguistic markedness and frequency
Phonological encoding in Cantonese	Investigate the role of explicit representations of tone in SFUSED Cantonese

Objective: combine insights from both phonological and psycholinguistic theory in an integrated model of phonological encoding

SFUSED **Simon Fraser University Speech Error Database**	*Investigate phonological patterning in a large, methodologically rigorous speech error database*
Phonological regularity	Investigate phonotactic violations in SFUSED and discuss model implications for syllable templates and frequency based solutions
Consonant substitutions	Investigate competing explanations from cross-linguistic markedness and frequency
Phonological encoding in Cantonese	Investigate the role of explicit representations of tone and syllables in SFUSED Cantonese

SFU Speech Error Database (SFUSED)

Current languages

SFUSED English (10,104 errors)

SFUSED Cantonese (2,549 errors)

Goals

- Build a multi-purpose database that supports all types of language production research
- Linguistically sophisticated coding of errors
- Methodologically rigorous techniques for collecting and analyzing speech errors
- Examine how the structure of non-Indo-European languages impacts language production processes

Methods

Speech errors: unintended, non-habitual deviation from the speech plan (Dell 1986)

Offline collection from audio recordings

- Errors collected from third party sources, podcasts on variety of topics
- Podcasts selected for having natural unscripted speech, usually Midlands dialect 'Standard American' English, high production quality, no media professionals
- Multiple podcasts (8 currently) with different talkers, approx. 50 hours of each podcast
- Record dialectal and idiolectal features associated with speakers
- Also experimented with "online" collection, i.e., on-the-spot observation

Multiple data collectors and training

- Total of 15 trained data collectors (reduces collector bias)
- Undergraduate students given phonetic training and tested for transcription accuracy
- Introduction to speech errors, definition and illustration of all types
- Listening tests: assigned pre-screened recordings, asked to find errors; learn to detect errors and record idiolectal features by reviewing correct list of errors.
- Trainees that reach a certain level of accuracy and coverage can continue.

Classification separate from data collection

- Data collectors use speech analysis software and detailed protocol for detecting errors in audio recordings, and excluded 'red herrings'
- Submissions: speech errors in spreadsheet format, batch imported into database
- Data analysts (different than collector) verify the error, classify it using the SFUSED fields

SFUSED English interface

SFU SED

Record ID no. 12470
Last Modified 2017-09-27
Total Completed 4712

Major Class Fields

Master Type: Phonological Substitution
Altern. MType:
Level: Sound
Type: Substitution
Direction: NotAppl
Contextual? Y N (N)
Right Lexeme? Y N NotAppl (Y)
Form Rule Violation? Y N (N)

Specific Class Fields

Obvious Malapropism Y N NotAppl
Phonotactic Violation
Gradient Type
Prosody Type
Form Rules Type
Morpho Cats
Onsetless Syllable Source? Y N
Onset CC Source? Y N
Two Term Intervener
Transformation Type
Complex Processes-check all appropriate:
Sub Del Stress Exch
Add Shift Grad Blend

Example Fields

A: I don't allow my dog to get blood transfusions. B: Did you xxx oh //[sw]eaking of ah, /C[w]ist, Christ and science, there's a show on the history channel …

	Intended	Error
Orthographic:	speaking	[sw]eaking
Phonetic:	'spikɪŋ	'swikɪŋ

Word Bounded? Y N (N) Clipped? Y N (N)
Source Different Talker? Y N Corrected?

Word Fields

Lexical Word? N
POS: Verb NotAppl
Open/Closed: Open
Regular/Irregular: Irreg
Error-Intended Semantic Relationship: Same meaning (same lexeme)
Error-Intended Morphological Relationship: Same Lexeme (differ only g. f.)

Sound Fields

	Supplanted Intended: p	Intruder: w	Source Sound
CV Structure:	C	C	
Syllabic Role:	Onset	Onset	NotAppl
Word Position:	Medial	Medial	NotAppl
Whole Syllable:	spi	swi	

Error action is in: Onset

Given Record Fields

Researcher: ak
Found Date: 2017-09-16
File: spy107_2010-03-10
Podcast: spy
Time stamp: 0:11:28
Online/offline: off
Talker Self?
Speaker: gc Sex: m
Spreadsheet: spy107
Confidence Intended
Confidence Transcript.
Personal Info? Y N
Record Complete? Y N (Y)
Record Confirmed By:
[X] ja [X] ak [X] ld [ ] on

Identical Neighboring Seg? Y N (N)
Syllable with error has:
Main Stress (selected) Second. Stress No Stre
Phonologically Legal? Y N (Y)
Triggers Resyllabification? Y N (N)

Markedness Measures

Onset (Initial)	Satisfaction	Violation	No Change
NoHiatus	Satisfaction	Violation	No Change
NoComplexOnset	Satisfaction	Violation	No Change
NoComplexCoda	Satisfaction	Violation	No Change
NoCoda	Satisfaction	Violation	No Change
NoDiphthong	Satisfaction	Violation	No Change

Advantages

Alderete & Davies 2018, Language and Speech

Reliability and data quality

- Audio recording supports data collection separate from verification by another researcher (at least 25% omitted)
- Because use different collectors, can minimize collector bias and measure it if it exists
- Audio recordings help in spotting idiolectal features and phonetic structures

Metrics

- Audio recordings have a duration, with allows measures that are not possible with online collection, e.g., collection metrics
- Supports much better estimates of speech error frequency
- Using capture-recapture methods, we find that speech errors are much more frequent than reported in prior work (an error at least every 48.5 seconds, probably more)

Data discovery

- Audio recordings allow acoustic analysis, probe fine-grained phonetic detail
- Can address frequent cry for “more context” (can be recovered from audio)
- With a time metric, can investigate time-based effects like speech rate

Better sample of true population of speech errors

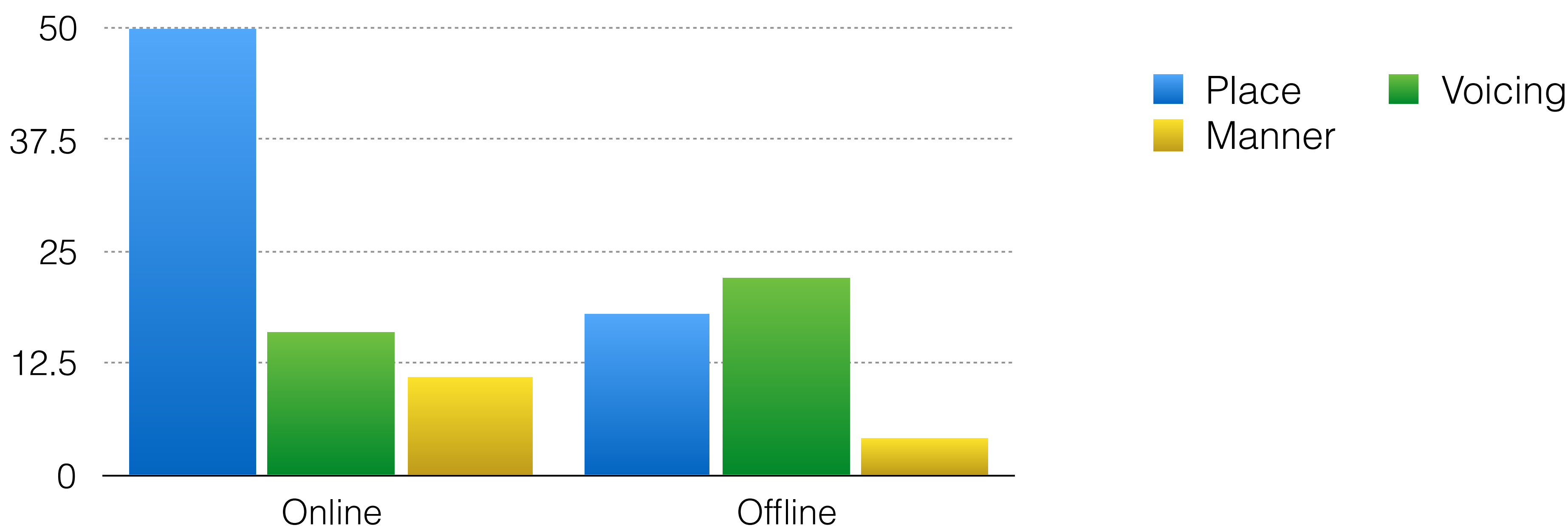
- Sample has much higher coverage, likely three to four times better
- Better sample: less ‘easy to hear’ and more ‘hard to hear’ speech errors
- Collect more errors that occur in fast speech

Better sample: robust to perceptual biases

Perceptual bias: errors in mis-pronunciation in place of articulation are easier to detect than voicing (Cole et al. 1978, Stemberger 1992, see Pérez et al. 2007).

Test: compare data collection "online" (on-the-spot observation) and "offline" (from audio recordings, most of SFUSED data), balanced for experience levels.

Finding: online data collection reflects pattern expected by perceptual bias (many more errors in place), but offline is not skewed by bias.



Summary of differences

Alderete & Davies 2018, Language and Speech

Sound errors

- Online errors have more corrected errors than offline errors.
- Online has a stronger repeated phoneme effect than offline errors.*
- Online errors have a stronger lexical bias than offline errors.(*)
- Online errors have a weaker word-onset effect than offline errors.*
- Online errors are more likely to be contextual than offline errors.*
- Online errors have more perseverations and exchanges than offline errors.*
- Online sound substitutions are more symmetric and more concentrated in a small number of substitutions than offline errors, which are more diffuse and asymmetrical.*

Word errors

- Online errors have less additions and deletions and more blends than offline errors.*
- Online word substitutions are much more likely to be in nouns than offline errors, which are more diffuse across lexical and function categories.*
- Online errors tend to respect the category constraint more than offline errors.

* = significant association from chi square test

Take home: *methods of collecting and analyzing speech errors have significant consequences for empirical patterns, which clearly inform phonological patterns.*

Objective: combine insights from both phonological and psycholinguistic theory in an integrated model of phonological encoding

SFUSED Simon Fraser University Speech Error Database	Investigate phonological patterning in a large, methodologically rigorous speech error database
Phonological regularity	*Investigate phonotactic violations in SFUSED and discuss model implications for syllable templates and frequency based solutions*
Consonant substitutions	Investigate competing explanations from cross-linguistic markedness and frequency
Phonological encoding in Cantonese	Investigate the role of explicit representations of tone in SFUSED Cantonese

Motivation for study of phonotactics

Sound errors respect phonotactics
(Wells 1951, Boomer & Laver 1968, Nooteboom 1967, Garrett 1980)

Phonotactics might just be learned sequences
(Dell et al. 1993)

Problem for SRN: 99% regularity, cf. ~96% regularity

Perceptual biases: against perceiving errors with phonotactic violations; listeners may regularize them or simply fail to hear them. (Cutler 1982, Shattuck Hufnagel 1983)

Methods: English phonotactics

Objective: define a system of phonotactics for English and use them to assess phonotactic violations in SFUSED English.

Onset **Peak** **Coda**
(s)(C1)(C2) X4 (X5) (C6)(C7)(C8)(C9)

Guiding assumption: a word is phonotactically licit if it can be syllabified within a well-formed syllable of English (Kahn 1976, Giegerich 1993, Jensen 1993)

Conditions:

All C positions are optional.
Banned C1: *ŋ ʒ* , Banned Codas: *h, j, w.*
Onset clusters: obstruent + sonorant
Appendix + C, C always a voiceless stop, *sf* rare/loans
Banned onset clusters: vd fric/affricate + sonorant, labial + *w*, coronal nonstrident + *l,*
θw ʃjV ʃw ʃl sr sh gw stw skl
Onglide *j*: part of peak because of limited distribution, but cannot occur in *CCju* cluster.
Coda clusters X5+C6: falling sonority (r > l > nasals > obstruents) and *s* + *p t k; lg* is banned.
C7-9 are appendices limited to coronal obstruents
Nasal + obstruent clusters agree in place and the obstruent is voiceless.
Tense vowels and diphthongs are bimoraic (fill X4 and X5), lax vowels are short fill X4.
Stressed and final syllables are bimoraic (lax vowels occur in closed syllables) and all syllables maximally trimoraic (syllables tense vowels only have simple codas)

Results: illustrating phonotactic violations

Alderete & Tupper 2018, WIREs Cognitive Science

Substitutions

1500 … by the maps at the ^selection /[**ʃkr**in] (screen)
5739 … they shoot, /[**ʒ**u] shoot The Thick of It … (you)

Illicit onsets/appendices

1245 … Their HOV /[l**aɪŋ**] xxx lane is like one driver (lane)
10,780 … well it /absorb[ʒ] it, it's now giving it off (absorbed)

Illicit codas/rimes

Additions

49 … get the Ferrari down a /[**flj**u] xxx few ^floors? (few)
5599 … talking a ^dream, what that ^dream /[**mr**]eans … (means)

Illicit onsets

1526 The ^person /[k**eɪmp**] ^up to the desk.

Illicit rime

Deletions

3954 … Lisa, /**Sr**eech and Lisa. (Screech)
8943 … I think you're a /h**u[ŋə**]= hunk-a-rama.

Illicit onsets

Sequential/Word Blends

7211 … because we /[spɪ**lkf**] xxx we, we speak film

Illicit codas

7120 Top ten /**thw**ays to make me cry (things, ways)

Illicit onset

(SFUSED record ID # on left)

Results by error type

Observations: % of phonotactic violation differs by type, but overall % of irregularity much higher than 1% found in Stemberger's corpus.

Error type	Example	*N*	Violations	% of *N*
Substitutions	pleep for *sleep*	1,376	44	3.20
Additions	bluy for *buy*	358	33	9.22
Deletions	pay *for play*	169	3	1.78
Exchanges	heft lemisphere for *left hemisphere*	37	2	5.41
Shifts	*splare backforests* for *spare blackforests*	7	0	0.0
Sequential Blends	Tennedy for *Ted Kennedy*	57	4	7.02
Word Blends	tab for *taxi/cab*	72	4	5.56
Totals		2,076	90	**4.34**

Perceptual bias: missed phonotactic violations

Conjecture: low counts of phonotactic violations due to perceptual biases against them (Cutler 1982, Shattuck Hufnagel 1983)

Probe: Alderete and Davis (2018) used balanced sample of online vs. offline errors and found a significant association between methodology and regularity ($\chi(1)2=7.902$, $P=0.0049$).

	Offline	Online
Phonotactic Violations	17 (3.19%)	8 (0.95%)
No Violations	516 (96.81%)	831 (99.05%)

Perceptual bias: all sound errors

Conjecture: low counts of phonotactic violations due to perceptual biases against them (Cutler 1982, Shattuck Hufnagel 1983)

Probe: counting all sound errors and blends, % of phonotactic violations higher ($X2 = 16.9618$, $p < .05$); note effect does not depend on what counts as a violation.

	Offline	Online
Phonotactic Violations	76 (5.5%)	11 (1.6%)
No Violations	1,326 (94.5%)	660 (98.4%)

Overwhelmingly regular, but above chance?

Question: the lower standard of phonological regularity raises the question of whether it is significantly above chance levels.

Permutation test (see Dell & Reich 1981)

1. Randomly permute segments from a list of intruder segments (given from error corpus), holding constant the phonological context (i.e., segments selected independent of context)
2. Use multiple trails to obtain a distribution of the percentage of regular errors under the independence assumption.
3. Test to see if there is sufficient evidence to reject independence hypothesis.

Example: C1 slot in substitutions

1. Take all intruders in C1 slot of <u>C1</u>C2 complex onsets in the error data, randomly permute them in the same slots they occur in errors.
2. Permute segments of C1 many times, get an average rate of phonologically regular words
3. Compare with actual errors in C1 of C1C2 errors to see if deviate significantly.

Results: some, not all, above chance

Alderete & Tupper 2018, WIREs Cognitive Science

Finding: in both substitution and addition errors, significantly above chance in non-initial positions (C2 of cluster), but not above chance initially (C1 of cluster)

Type	Context	Example	*N*	Actual	Random	Significant?
Substitutions	_C of $\underline{C}C_{\text{Onset}}$	*blue → plue*	37	81%	78%	No (p=0.38)
	C_ of $C\underline{C}_{\text{Onset}}$	*dream → dweam*	36	100%	83%	Yes (p=1e-6)
Additions	_C into $\underline{C}C_{\text{Onset}}$	*last → flast*	29	62%	64%	No (p=0.77)
	C_ into $C\underline{C}_{\text{Onset}}$	*bad → brad*	75	87%	79%	Yes (p=0.005)

Interpretation:

- Non-initial contexts require analysis; 'structural matrix' of some kind to predict above chance level.
- C1 errors are dominated by errors that occur word-initially, so could be an effect of the word-onset bias (Wilshire 1999)

Discussion: comparison with SRN

New standard: 93-95% phonologically regular, cf. 99% of Stemberger 1983, (94.5% regularity reported in offline data still probably affected by perceptual bias)

Goodness of fit: Dell et al. 1993 simple recurrent network tested a variety of parameters that compare well with these findings.

- Models trained on frequent vocabulary and with both internal and external representations: **96.3% regularity**
- Range for other assumptions about input: **89-95% regularity**
- Many of the errors with phonotactic violations resemble the phonotactically illicit errors we have found, with illegal clusters and initials.
- SRN also predicts **more errors word-initially**, because no prior probability to predict future sounds

Limitations

- Model trained only on three segment words, so no polysyllabic words
- Didn't really allow for additions, which account for a lot irregularity (perhaps 1/3)
- Phonotactics slightly different than one used here (likely less stringent).
- Didn't account for prosody (stress in errors) and other structures.

Take home: with the new standard, tactic frames (cf. syllable templates) are not obviously necessary to the analysis of phonotactic regularity in speech errors.

Discussion: grammar at the granular level

Conclusion: seems sensible to conclude that global phonotactics, broad classifications of ‘phonologically grammatical’ and ‘phonologically ungrammatical’ can be produced with a model that lacks syllable templates and node matching.

Problem: contemporary phonological theory does much more than classify forms as ‘grammatical’ and ‘ungrammatical’. Most modern phonological analyses posit constraints that target specific types of phonological structure in particular positions in a form.

Classical Optimality Theory (Prince and Smolensky 1993/2004)
Harmonic Grammar (Legendre, Miyata and Smolensky 1990)
MaxEnt Grammar (Hayes and Wilson 2008)

Granular constraints: Onset, NoCoda, *Labial, *h]#

Question: is there any evidence for that granular constraints such as these shape speech errors?

Objective: combine insights from both phonological and psycholinguistic theory in an integrated model of phonological encoding

SFUSED Simon Fraser University Speech Error Database	Investigate phonological patterning in a large, methodologically rigorous speech error database
Phonological regularity	Investigate phonotactic violations in SFUSED and discuss model implications for syllable templates and frequency based solutions
Consonant substitutions	*Investigate competing explanations from cross-linguistic markedness and frequency*
Phonological encoding in Cantonese	Investigate the role of explicit representations of tone in SFUSED Cantonese

Background: segmental markedness

Bias for marked → unmarked mappings in speech errors at segmental level

Experimentally induced speech errors

- Kupin 1982: disyllabic tongue twisters, unmarked forms preferred
- Goldrick 2002: implicit learning paradigm, examined substitutions where markedness and frequency make different predictions.
  Example: [t] is unmarked relative to [s], but less frequent
  [s] → [t] > [t] → [s] supports markedness account

Aphasic speech

- Blumenstein 1973: single feature consonant substitutions favour marked → unmarked mappings (just Broca's and Wernicke's aphasics, not conduction aphasics)
- Romani et al. 2002: markedness superior to frequency in aphasic consonant substitutions
- Goldrick and Rapp 2007: brain-damaged subject with deficit in post-lexical phonological processes, more accurate with coronals /t d/ (93%) than dorsals /k g/ (86%)

Against markedness as a factor

Some studies have found no effect of markedness, and segment substitutions reflect baseline frequencies ('availability'): Shattuck-Hufnagel & Klatt 1979, Stemberger 1991

Focus: single consonant substitutions

Example: ... in each pixel /**r**un ^row at a time.
(Intended: *one*, *w* → *r)*

- Prior research has focused on this type of error, so can compare findings.
- Most common type of error in all speech error corpora in normal speakers, so able to get sufficient data
- Can study segmental markedness by considering differential effect of markedness on feature structure.
- Cross-linguistic markedness makes predictions about marked and unmarked values of features: e.g., voiced marked relative to unvoiced

Method: single feature consonant substitutions

Procedure: take a consonant confusion matrix (*N= 1,506*, SFUSED English)

Sound Intruder

Sound Suppl..	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʒ	tʃ	dʒ	m	n	ŋ	l	r	w	j	h
p		27	10		10		10	1							2		2	1						
b	14	1	1	9	1	1	1	8			1						7			3		2		
t	7	2		15	13	1	2	1	2	2	2	3	3		10			7		6	1		1	1
d	1	8	3		1	6			2	2	3	3				6		5		2	3			
k	10	2	17	2		16					2	1	2		3	1			1	1	1			2
g		3		2	15		1	1			1		1					1						
f	8	3	2			1		6	3		9	1								1	1			1
v		9	1		1	1	11			2		2				2	1	1		1	3	3		
θ	2		3	2	2		1			2	13	1	1		1					1				
ð		1	3	3		1		2	2		1	1			1	1	1			4				1
s		1	8	1	4	1	4	1	12	1		7	55		5			2		1				1
z		1		2		1		2	3	1	8		4	1		2	1		1	2				
ʃ				1		1	1		1		14	1		1	6						1			
ʒ											1													
tʃ			4		4						5		5	3		1					1		1	
dʒ			1	7	1	3					1			1	5							1		
m	2	8					1	1		1		1						24		3		6		
n			3	12	1						1	1					16		4	8	5		1	1
ŋ																		1						
l		1			1	3				2			1				4	3	2		21	6	2	
r	1			1				1		1					1	1	3			41		12	2	
w	1	2	1	2			1	1									4			8	15			
j		1				2								1		1				2	3			
h		1	4	1	4	3	1		1		3				1					2				

Test: single feature consonant substitutions

Procedure: take a consonant confusion matrix (*N= 1,506*, SFUSED English)

Sound Intruder

Sound Suppl..	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʒ	tʃ	dʒ	m	n	ŋ	l	r	w	j	h
p		27	10		10		10	1							2		2	1						
b	14	1	1	9	1	1	1	8			1						7			3		2		
t	7	2		15	13	1	2	1	2	2	2	3	3		10			7		6	1		1	1
d	1	8	3		1	6			2	2	3	3				6		5		2	3			
k	10	2	17	2		16					2	1	2		3	1			1	1	1			2
g		3		2	15		1	1			1		1					1						
f	8	3	2			1		6	3		9	1								1	1			1
v		9	1		1	1	11			2		2				2	1	1		1	3	3		
θ	2		3	2	2		1			2	13	1	1		1					1				
ð		1	3	3		1		2	2		1	1			1	1	1			4				1
s		1	8	1	4	1	4																	
z		1		2		1																		
ʃ				1		1	1																	
ʒ																								
tʃ			4		4																			
dʒ			1	7	1	3																		
m	2	8					1																	
n			3	12	1																			
ŋ																								
l		1			1	3																		
r	1			1																				
w	1	2	1	2			1																	
j		1				2																		
h		1	4	1	4	3	1																	

Test: examine consonant pairs that differ in a single feature, adjust for baseline frequencies

Mapping	Count	Baseline
p → b	27	*p* produced 50 times in 1000
b → p	14	*b* produced 29 times in 1000

Baseline frequencies: estimating relative risk

Tupper and Alderete 2017
(See also Stemberger 2007)

	Event	**General Population**
Condition 1	a	b
Condition 2	c	d

mutually exclusive

$$RR = \frac{a/b}{c/d}.$$

Baseline frequencies: estimating relative risk

Tupper and Alderete 2017

	Event	General Population
Condition 1	a	b
Condition 2	c	d

mutually exclusive

$$RR = \frac{a/b}{c/d}.$$

Null hypothesis: RR = 1

	Voicing [s z]	Token Frequency
z → s	8	16
s → z	7	99

$RR(sz) = 7.07$

Baseline frequencies: estimating relative risk

Tupper and Alderete 2017

	Event	General Population
Condition 1	a	b
Condition 2	c	d

mutually exclusive

$$RR = \frac{a/b}{c/d}.$$

Null hypothesis: RR = 1

	Voicing [s z]	Token Frequency
z → s	8	16
s → z	7	99

$RR(\mathrm{sz}) = 7.07$

Test results: are the observed differences in probability of two events significant (not due to chance)? And if so, what direction (favour marked or unmarked structure?)

95% confidence interval (testing null hypothesis that log(RR) = 0):

$$\log RR \in \left(\log\frac{a/b}{c/d} - 1.96\sqrt{\frac{1}{a}+\frac{1}{c}}, \log\frac{a/b}{c/d} + 1.96\sqrt{\frac{1}{a}+\frac{1}{c}}\right)$$

Agresti 1996

Example: log(*RR*)= 1.956, 95% confident log(*RR) =/* zero, can reject null hypothesis. Direction (sign): favours unmarked segment [s].

Results: [voice], [anterior], [continuant], [nasal]

Voicing

Unmarked	Marked	Direction	Significant?
p	b	marked	N
t	d	marked	N
k	g	unmarked	N
f	v	unmarked	N
s	z	unmarked	Y

Continuancy

Unmarked	Marked	Direction	Significant?
p	f	unmarked	N
b	v	unmarked	N
t	s	unmarked	Y
d	s	unmarked	N

Anteriority

Unmarked	Marked	Direction	Significant?
s	ʃ	unmarked	N
t	tʃ	unmarked	Y
d	dʒ	unmarked	Y

Nasality

Unmarked	Marked	Direction	Significant?
b	m	marked	N
d	n	unmarked	N

Finding: 4 of 14 consonant pairs reached 95% significance, all in the direction predicted by markedness (some pairs not reported due to insufficient data)

Results: place features

Coronal - Labial

Unmarked	Marked	Direction	Significant?
t	p	unmarked	Y
d	b	unmarked	Y
n	m	unmarked	Y
s	f	unmarked	Y

Labial - Dorsal

Unmarked	Marked	Direction	Significant?
p	k	marked	N
b	g	unmarked	N

Coronal - Dorsal

Unmarked	Marked	Direction	Significant?
d	g	marked	N
t	k	unmarked	Y

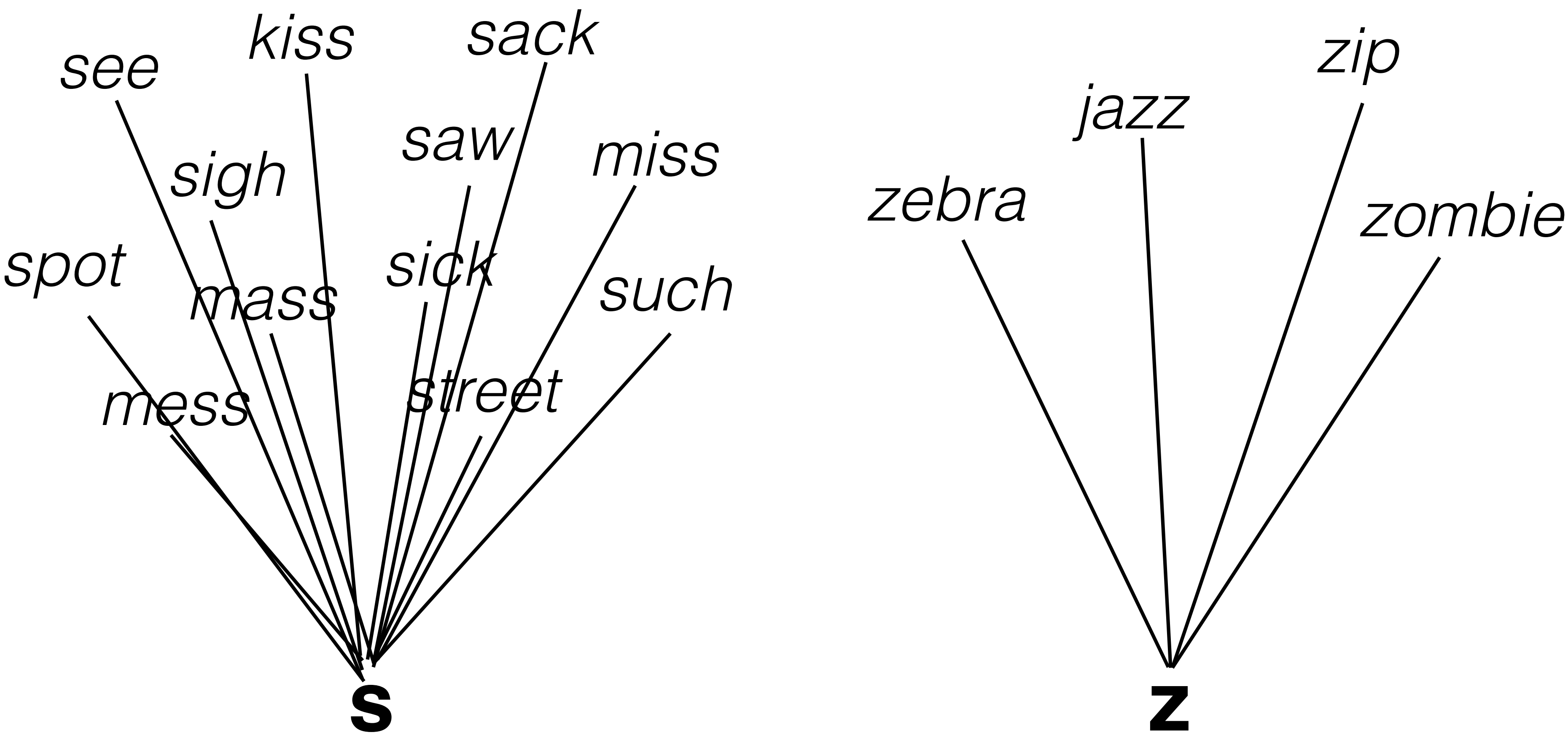
Finding: majority of place-changing substitutions significant, especially those involving coronals.

Markedness distinct from frequency bias?

Feature	Unmarked	Marked	Direction	Significant?	Frequency bias?
[voice]	s	z	unmarked	Y	Y
[anterior]	t	tʃ	unmarked	Y	Y
[anterior’	d	dʒ	unmarked	Y	Y
[continuant]	t	s	unmarked	Y	N
Place	t	p	unmarked	Y	Y
Place	d	b	unmarked	Y	Y
Place	n	m	unmarked	Y	Y
Place	s	f	unmarked	Y	Y
Place	t	k	unmarked	Y	Y

Confound: while there are many significant results supporting a role for markedness, 8 of the 9 cases could be explained with an output bias for frequent segments (type frequency, interactivity in the lexicon). [t] → [s] is the same mapping Goldrick (2002) found to support the markedness account using experimental methods.

Frequency bias via interactive spreading



Output bias for frequent segments (Dell 1986 et seq.): frequency effect commonly assumed to result from interactive spreading from words with frequent sounds; activation of neighbouring words feeds into high frequency segment, can tip the scales in favor of frequent segment.

Conclusion: weak support for a role for markedness in consonant substitutions; solution based on type frequency is very compelling.

Objective: combine insights from both phonological and psycholinguistic theory in an integrated model of phonological encoding

SFUSED Simon Fraser University Speech Error Database	Investigate phonological patterning in a large, methodologically rigorous speech error database
Phonological regularity	Investigate phonotactic violations in SFUSED and discuss model implications for syllable templates and frequency based solutions
Consonant substitutions	Investigate competing explanations from cross-linguistic markedness and frequency
Phonological encoding in Cantonese	*Investigate the role of explicit representations of tone in SFUSED Cantonese*

Motivation for linguistic representations

Planning units: phonological categories used to assemble a speech plan; speech errors tend to involve established phonological structures.

Segments	**Primacy of segments**: single segment sound errors are the most common type of error, and some segment errors like exchanges have no good alternative analysis.
Onset/Rime	Sub-syllable **CC and VC sequences** also relatively common
Features	**Features Paradox**: errors involving just features are exceedingly rare, but features underlie the similarity effect (similar sounds slip)
Syllables	**Syllable Paradox**: errors involving whole syllables are also exceedingly rare (in English at least), but syllable roles shape error patterns because sounds tend to slip in similar positions.
Prosody	**Inertness of prosody**: stress errors are also extremely rare, and may not even result from phonological encoding

Common assumption: constructing a speech plan is fundamentally a matter of selecting segments (and perhaps sub-syllabic units); other structure, like features and syllables are stored but not actively selected (Fromkin 1971, Shattuck-Hufnagel 1979, Dell 1986, 2014). Metrical structure is not stored or selected, but may be referenced via diacritics (Levelt et al. 1999)

Active debate: is tone part of phonological encoding?

	Yes! Wan & Jaeger 1998, Gandour 1977, Shen 1993, Wan 2006, Liu & Wang 2008	**No!** Chen 1999, Roelofs 2015, Kember et al. 2015
Parallels:	***Tone is like segments***, can be mis-selected, and therefore tone must be represented in phonological encoding, like segments.	***Tone is like metrical structure***. It is diacritically represented in encoding and implemented later by articulatory processes. It cannot be mis-selected.
Evidence:	Tone slips are **relatively common**, and exhibit normal patterns of contextual errors, i.e., perseveration, anticipation, and exchanges.	Tone slips are **extremely uncommon**, and the rare cases that exist have alternative analyses.
	Tone is incorporate in the phonological organization of the lexicon, so must be part of encoding.	Also, tone does not have an implicit priming effect, so perhaps not represented in speech plan.

Tone slips in SFUSED Cantonese

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

Alderete, Chan, and Yeung 2018

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 → baːt33gwaː33∅	26
Lexical errors	/jiŋ55man25 'English' (lei22man25 'Italian')	245

Second most common type

Observation: tone slips are not rare at all in Cantonese, a language with six lexical tones.

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.

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.

Interactive spreading effects for tone

Prediction: contemporary theorizing about phonological encoding predicts that selection of different elements interact with each other in that shared elements increase the chance of a mis-selection.

Interactive spreading for tone

Lexical errors: greater than chance probability that intended and error word share the same tone.

Sound substitutions: greater than chance probability that sound intruders come from a source syllable that has the same tone as the error syllable.

Phonological similarity: tone slips are more common with the intended and error tone are similar.

Parallels with segments

Malapropisms: substituted words share segments (Fay and Cutler1977)

Repeated phoneme effect: sound errors are more common when the intended and source word share sounds (Dell 1984, MacKay 1970)

Phonological similarity effect: sound errors are more common when two sounds are similar (Shattuck-Hufnagel 1979)

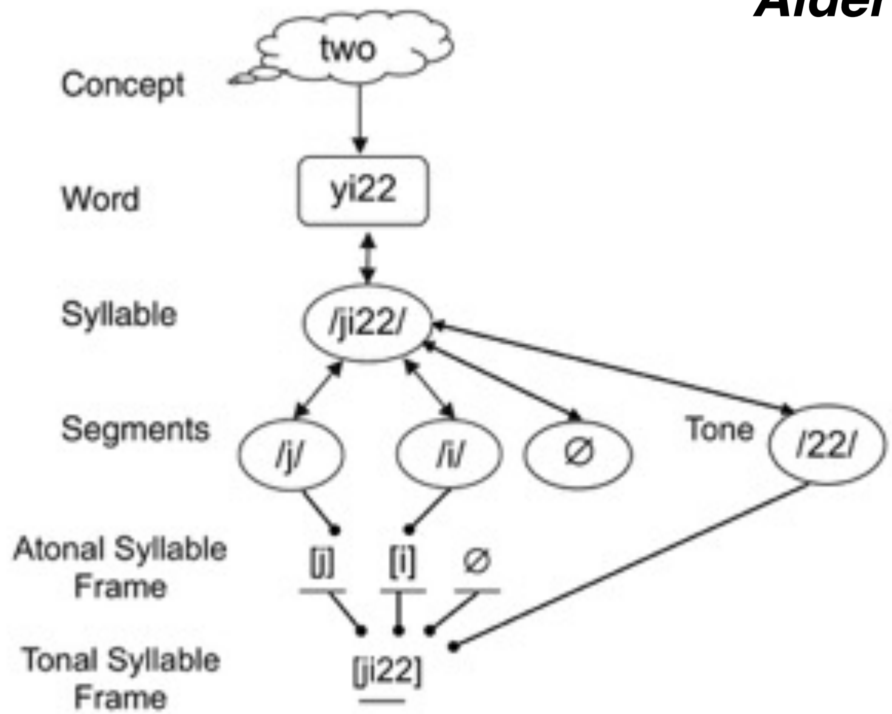
Finding (Alderete et al. 2018): tone is just like segments in that it interacts with other lexical and phonological structures in ways expected if they are actively selected in phonological encoding. Supports Wan and Jaeger's (1998) original claim that tone is phonologically integrated in the lexicon.

Discussion: model implications

Alderete, Chan, and Yeung 2018

Lexical selection

Phonological encoding



Key assumptions:

- Syllable proximate unit to the lemma (O'Seaghdah et al. 2010)
- Selection mechanism for tone
- Tone selection roughly simultaneous with segments
- Downstream interactions mapping tone to syllables

Conclusion: tone is fully represented and actively selected in phonological encoding, accounting for:

- Finding that tone slips are relatively common.
- Tone slips are usually contextual.
- Encoding of tone interacts with encoding of other linguistic elements.

Take home: linguistic representations of tone are crucial to account for key facts of tone processing.

Wrap up: theoretical implications

Necessary ingredients of an adequate model

- Output bias for frequent sounds (processing: type frequency)
- Output bias for frequent sequences (processing: associations in sequences)
- Linguistic structures as planning units (linguistic: phonological tone and syllables are fully specified)

Models that meet these criteria

- SRN of Dell et al. 1993
- Original Dell-net of Dell 1986 (with some overkill)
- Harmonic Grammar with weights for important constraints (Goldrick and Daland 2009), and related Gradient Symbol Processing models (Smolensky, Goldbrick, and Mathis 2014)
- Two step interactive model (Dell et al. 1997)
- Proximate-unit hypothesis inspired models (O'Seaghdha et al. 2010)

Problems raised by the research

Syllable-related markedness (Blumstein 1973, Goldrick and Rapp 2007)
How does markedness and frequency play out in syllable structures, e.g., marked onset clusters, codas, etc. Strong evidence from aphasic research that markedness shapes aphasic speech.

Syllable position constraint (Boomer and Laver 1968, Fromkin 1971)
Onsets slip with onsets, codas with codas, etc. See Dell et al. 1993 results.

Gradience and granular structure
We know that language particular constraints have different weights, or impact phonology differently. How does the different weights impact speech errors. Could higher weighted constraints have a stronger impact.

Word onset effects and contextually (Wilshire 1999)
While Dell's SRN give a very natural analysis of the word-onset effect, research has shown that this effect is limited to contextual errors. This is not predicted in the current model, so somehow competitive inhibition needs to be a prerequisite for

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

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

Why are we still collecting speech errors?

Problem: speech errors ‘in the wild’ are very time-consuming, prone to mistakes in observation and interpretation; often can’t get enough data from a particular pattern to test specific hypothesis.

Stemberger 1992: actually there is considerable overlap in the patterns of errors collected in naturalistic and experimental settings. So speech errors ‘in the wild’ present valid data patterns worthy of analysis.

Some patterns no suitable for experimental study: % of exchanges, lexical bias, non-native segments, phoneme frequency effects, etc.

This research shows that a new approach to data collection (offline, many listeners), has potential for new observations, e.g., phonological regularity

Large databases can be re-purposed and extended, not really true of experiments.

Offline methodology is actually very efficient (see Alderete & Davies 2016 for research costs estimates); can produce a database of 3,000 errors in about the same amount of time it takes to run two experiments.

Idiolectal features are _very important_ in understanding errors (habitual, so not an error), but can only really analyze them after a few hours of listening to a single talker.

Estimating error frequency

Alderete & Davies 2018, Language and Speech

Prior assumption: speech errors are rare in general (**error every 5-6 minutes**), motivates focus on normal language production

Problem: prior estimates of error frequency based on online collection, and many failed to address the fact of missed errors (though all studies concede they miss them).

Capture-recapture: common tool in ecology for estimating a population when exhaustive is impossible or impractical

Take home: speech errors occur much more commonly than enumerated in prior research, at least as often as **48.5 seconds** (upper bound because of non-homogeneity)

Second	A	B	C	AB	AC	BC	ABC	n	m̃	ṽ	SPE
2,100	2	18	3	2	0	3	5	33	16.3	49.3	42.60
1,690	6	5	4	5	0	2	9	31	13.48	44.48	38.00
1,993	2	9	5	1	0	1	5	23	20.08	43.08	46.26
2,385	6	6	5	8	2	1	5	33	11.7	44.70	53.36
4,143	24	9	1	5	1	1	3	44	21.84	65.84	62.93
3,000	9	2	7	3	5	1	2	29	10.63	39.63	75.70
1,800	9	9	3	2	0	1	1	25	29.87	54.87	32.81
2,377	15	2	4	3	2	1	3	30	13.39	43.39	54.78
2,400	18	4	6	1	2	0	7	38	41.93	79.93	30.03