May 24, 2019 UW Phonology Lab

# 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)

# Phonological generalizations in speech errors

**Syllable position effect** (Boomer and Laver 1968, Fromkin 1971)
Sounds tend to slip in the same positions as they occur in source words, e.g., onsets slip with onsets, codas with codas, etc.

**Phonological similarity effect** (Cutler 1980, Dell and Reich 1981)
Intended and intruder sounds tend to be phonologically similar, e.g., substitutions *p>f* more common than *p>r*

**Repeated phoneme effect** (MacKay 1970, Dell 1984)
Sound errors that share a context in intended and source words are common, e.g., *heft lemisphere (left hemisphere)*

**Single phoneme effect** (Nooteboom 1969, Shattuck Hufnagel 1983)
Large majority of sound errors are single segments (70-90%), not sequences or features

**Phonological (phonotactic) regularity effect** (Wells 1951, Stemberger 1983)
Speech errors tend to be phonologically regular, i.e., obey phonotactics.

# Converging views

***Caveat:*** phonological analysis is a different enterprise from analyzing on-line language production processes

## Viability of phonological segments

- Phonology: distributions and processes depend on phonological segments
- Language production: segments are a fundamental unit in speech planning (Fromkin 1971, Dell 2002)

## Importance of syllables

- Phonology: syllables critical to both segmental and suprasegmental analyses (Itô 1989, Blevins 1995)
- Language production: segments are encoded with syllable constituency and whole syllables are retrieved, especially in Chinese languages (Chen 2000)

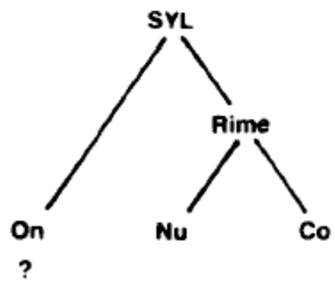
## Sensitivity to similarity structure

- Phonology: graded notion of similarity (function of shared features) crucial for harmony and disharmony phenomena (Frisch 1996)
- Language production: segmental similarity also formalized as a function of feedback from shared features (Goldrick 2004)

# Phonological grammar in production

**Tactic frames for phonotactics** (Shattuck Hufnagel 1979, Dell 1986)
Phonological encoding guided by syllable templates (and word frames and sentence frames) accounts for productive capacity, phonotactics generally

### Syllable frame



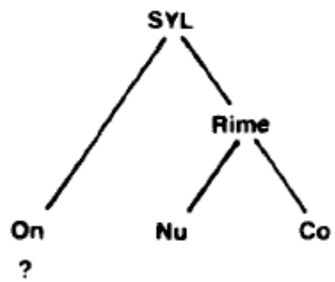

### Lexical insertion of segments

[k]/Onset
[p]/Onset
[ɑ]/Nuc
[b]/Onset
[ɪ]/Nuc
[æ]/Nuc
[t]/Onset
[tr]/Onset
[b]/Coda
[r]/Coda

# Phonological grammar in production

**Tactic frames for phonotactics** (Shattuck Hufnagel 1979, Dell 1986)
Phonological encoding guided by syllable templates (and word frames and sentence frames) accounts for productive capacity, phonotactics generally

### Syllable frame




### Lexical insertion of segments

[k]/Onset
[p]/Onset
[ɑ]/Nuc
[b]/Onset
[t]/Onset
[I]/Nuc
[æ]/Nuc
[tr]/Onset
[b]/Coda
[r]/Coda

### Additional roles for grammar:

**Underspecification in language production** (Stemberger 1991)
Segments may be underspecified in phonological encoding to account for their dominance in speech errors (cf. feature specification and phonological activity)

**Segment-to-frame association** (Levelt and Wheeldon 1994)
Activated segments are aligned with a metrical frame using left-to-right template mapping (cf. template mapping in auto-segmental phonology)

**Markedness effects in speech production** (Goldrick and Rapp 2007, Goldrick and Daland 2009)
Speech errors are shaped by markedness (toward unmarked structure) in constraint-based optimization models.

# Pause: Production models are not grammars

| | Production models | Generative models |
|---|---|---|
| | Spreading-activation models (Dell 1986 et seq.), WEAVER (Levelt et al. 1999), OSCAR (Vousden et al. 2000), Gradient Symbol Processing (Smolensky et al. 2014) | SPE Phonology (Chomsky & Halle 1968), Lexical Phonology (Kiparsky 1985), Optimality Theory (Prince & Smolensky 1994) |
| **Objectives:** | Capture on-line production processes | Correct analysis of sound patterns |
| **Properties:** | **Activation dynamics**<br>Activation flows through network, different outcomes different points in time | **Not dynamical**<br>Rarely have a dynamics*, principally interested in pairing inputs with outputs |
| | **Numerical processing**<br>Behaviour predicted by numerical computations on large bodies of data | **Symbolic computation**<br>Outcomes predicted from manipulation of symbols, |
| | **Frequency effects**<br>Biases towards frequent sounds and sequences intrinsic to lexical organization | **Role of frequency de-emphasized**<br>Structures produced without regard for frequency in the lexicon or in speech*. |
| | **Highly interactive**<br>Production processes take place within a large network of inter-connected elements | **Non-interactive**<br>Mappings analyzed in isolation, largely divorced from lexical networks* |

**Exceptions*: Articulatory Phonology, Exemplar Phonology

# Some leading questions

1. Given the differences between language production and generative models, are all direct roles for phonological grammar valid and appropriate? How reconcile competing explanations?

Take home: some constructs, like syllable frames, need to be reconsidered because alternative explanations exist that draw on processes intrinsic to language production.

2. What is the empirical basis for speech error patterns that support a role for grammar?

Take homes:

- New results suggest a weaker role for phonotactics in speech errors
- New data set for tone suggests a stronger role for tone as a planning unit than previously acknowledged.

# Road Map

**SFUSED (Simon Fraser University Speech Error Database)**
Describe methods for large database of speech errors, compare findings with other large collections.

**Phonological regularity**
Re-examine the impact of phonotactic restrictions in speech errors, consider model implications with a weaker role for grammar.

**Phonological encoding of tone**
Examine a new speech error collection of Cantonese, argue for a more direct role for phonological tone than admitted in contemporary production models.

# 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
- Document speech errors with rich linguistic detail
- Methodologically sound techniques for collecting and analyzing speech errors
- Document speech errors in non-Indo-European languages

# General methods

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

**Offline collection from audio recordings, cf. “online” (on-the-spot) data collection**

- 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

**Multiple data collectors and training, cf. few expert listeners approach**

- Total of 16 data collectors, about a month of training
- 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
- Right Lexeme? ◉ Y ○ N ○ NotAppl
- Form Rule Violation? ○ Y ◉ 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 Clipped? ○ Y ◉ 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
- Record Confirmed By: ☒ ja ☒ ak ☒ ld ☐ on

Identical Neighboring Seg? ○ Y ◉ N
Syllable with error has:
◉ Main Stress ○ Second. Stress ○ No Stre
Phonologically Legal? ◉ Y ○ N
Triggers Resyllabification? ○ Y ◉ 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 |

# **Perceptual biases** (Bock 1996, Pérez 2007)

**Problem**: collection of speech errors is very error-prone and affected by biases that may skew distributions in large data collections.

**Content bias**: errors are easier to detect if they affect meaning

**Attention bias**: lower level errors (phonetic or phonological) are often more difficult to detect and therefore require greater attention, substitution harder detect than exchange (e.g., ***left lemisphere*** vs. ***heft lemisphere***)

**Word onset**: errors are less noticeable if they occur at the end of words than the beginnings of words

**Predictability**: errors also easier to detect when they occur in highly predictable environments (e.g., *… go smoke a* ***cikarette***) or primed with words associated with the error word

**Bias for discrete symbolic categories**: especially for sound errors, biased toward assigning sounds to discrete phonological categories

**Compensation for coarticulation**: phonetic environments may enhance certain contrasts and lead to selection of some discrete sounds over others

**Feature biases**: sound errors with changes in some features are easier to detect that others, e.g., place easier to detect than voicing

**Data collector/talker bias**: collectors differ in the rate of detection and types of errors (see perceptual biases), and collectors may be limited to specific talkers with unique error patterns.

**Theoretical bias**: purchase of a theory or specific hypothesis may affect the errors collected

# Mitigating biases

**Offline with audio recordings**

- Reduces perceptual biases and constraints on attention because collector can replay, slow down, plan data collection in ways that supports better data collection.
- Audio recording allows data collection from multiple collectors (typically two)
- Audio recordings help in spotting idiolectal features, casual speech phonology, and phonetic structures

**Multiple data collectors**

- Talker bias reduced because many different talkers in multiple podcast series
- Collector bias reduced because of extensive training
- Use of many collectors also minimizes collector bias (reduce to individuals)

**Classification separate from verification**

- Audio recording supports data collection separate from verification by another researcher (at least 25% omitted)

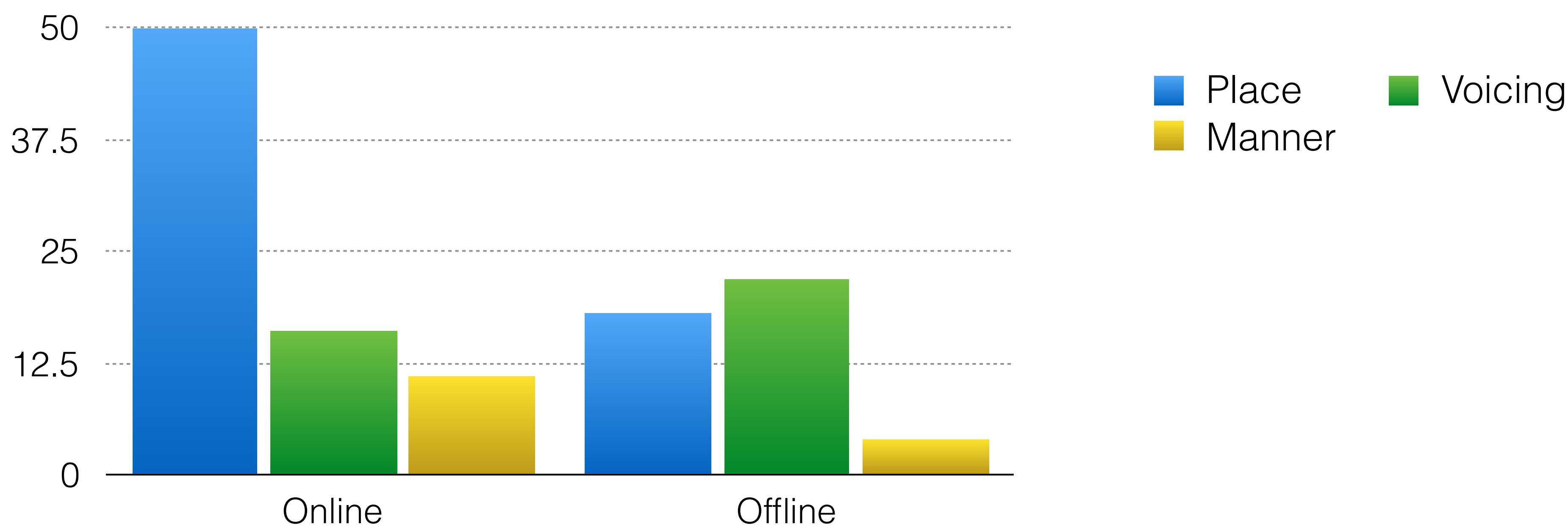
# Better sample: robust to perceptual biases

***Alderete & Davies 2019, Language and Speech***

**Place 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.



# Better sample: less 'easy to hear' errors

***Alderete & Davies 2019, Language and Speech***

**Sound exchanges**

Ex. *We can just wrap mine in a* ***/torn /korkilla*** (*corn tortilla*, 1495)

**Prediction:** attention bias favours more exchanges with online collection

| | Offline | Online |
|---|---|---|
| Morphemes | | 6 |
| Phrases | | 1 |
| Sounds | 1 | 25 |
| Words | 1 | 15 |
| Totals | 2 (0.38% of 533) | 47 (5.6% of 839) |

**% Exchanges elsewhere:**

Stemberger 1982/85: ~6%

Pérez et al. 2007: 35%

Dell and Reich 1981: 54%

# Summary of differences

***Alderete & Davies 2019, 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, and clearly matters for phonological patterns.*

***How does methodology affect data composition?***

***How does methodology affect phonological regularity?***

# Motivation for phonotactics

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

**Phonotactics not a hard constraint**
Speech errors are overall regular, but do admit phonotactic violations: roughly 1% of sound errors in Stemberger corpus

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

***Question: will revised assessment of phonological regularity affect model assumptions?***

# Methods: English phonotactics

***Alderete & Tupper 2018, WIREs Cognitive Science***

**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

**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.

Ph
Vi
No

| | **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 rate of phonological regularity raises the question of whether it is significantly above chance levels.

**Illustration**: /blue/ -> *plue*
**vlue*

- What is the chance rate that an error in C1 position of a CC onset violations phonotactics?
- Does the rate of phonotactic violations in the corpus actually deviate from chance?

**Estimating chance with permutation test** (see Dell & Reich 1981)

1. Randomly permute segments from a list of intruder segments (given from error corpus) by item, holding constant the phonological context (e.g., C1)
2. Use multiple trails to obtain a distribution of the percentage of regular errors under the independence assumption (i.e., intruders and slots for intruders independently selected).
3. Test to see if there is sufficient evidence to reject independence hypothesis.

# Results: complex onsets (mixed)

***Alderete & Tupper 2018, WIREs Cognitive Science***

**Finding**: in both substitution and addition errors into onset positions, violations 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 CC | *blue>plue* | 37 | 81% | 78% | No (p=0.38) |
| | C_ of CC | *dream>dweam* | 36 | 100% | 83% | Yes (p=1e-6) |
| Additions | _C of CC | *last>flast* | 29 | 62% | 64% | No (p-0.77) |
| | C_ of CC | *bad>brad* | 75 | 87% | 79% | Yes (p=0.005) |

**Interpretation:**

- Non-initial contexts require analysis because above change
- C1 errors are dominated by errors that occur word-initially, so could be an effect of the word-onset bias (Wilshire 1999)

# Model implications

**Review**: Stemberger's (1983) 99% regularity too high, SFUSED English has **94.5% regularity**

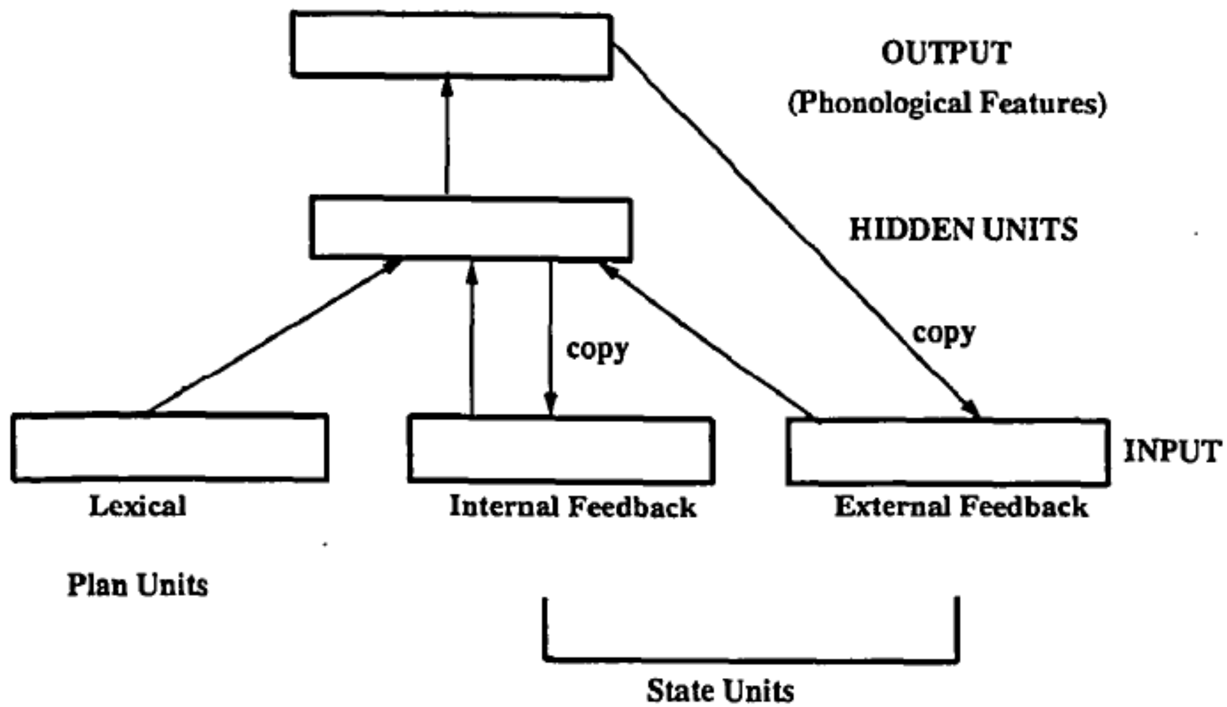
**Dell et al. (1993)**: **A production model without tactic frames**

**Simple Recurrent Network (SRN)**

Sequential: outputs a single segment, then another, in sequence

Recurrent: current segment processed in tandem with knowledge of past segments

Distributed representations: segments are represented as a vector of feature values (cf. distinctive features)




**Results**: Trained on a sample of English words and tested for phonological regularity. Given certain parameters (frequent vocabulary, internal and external input), produces errors that are **phonotactically regular about 96.3%** of the time (range 89-96%).

# Model implications, cont'd

**Review**: phonological regularity much higher word-initially:

Substitutions: 81% (initial), cf. 100% (non-initial)

Additions: 62% (initial), cf. 87% (non-initial)

**Interpretation**: word-onsets are simply more prone to error generally (Wiltshire 1999, cf. Berg & Abd-el-Jawad 1996), so higher regularity can be seen as a reflex of the word-onset effect

**Dell's (1993) SRN**: also shown to exhibit a word-onset effect because first segments lack prior probabilities to predict future sounds.

**Take home:** given the new standard (94-95% regularity) and its skewed distribution (only above chance in non-initial positions), tactic frames are not really necessary. SRN accounts for the observed data with natural production processes that lack syllable frame.

***How does phonology contribute to planning units in production?***

# 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 |

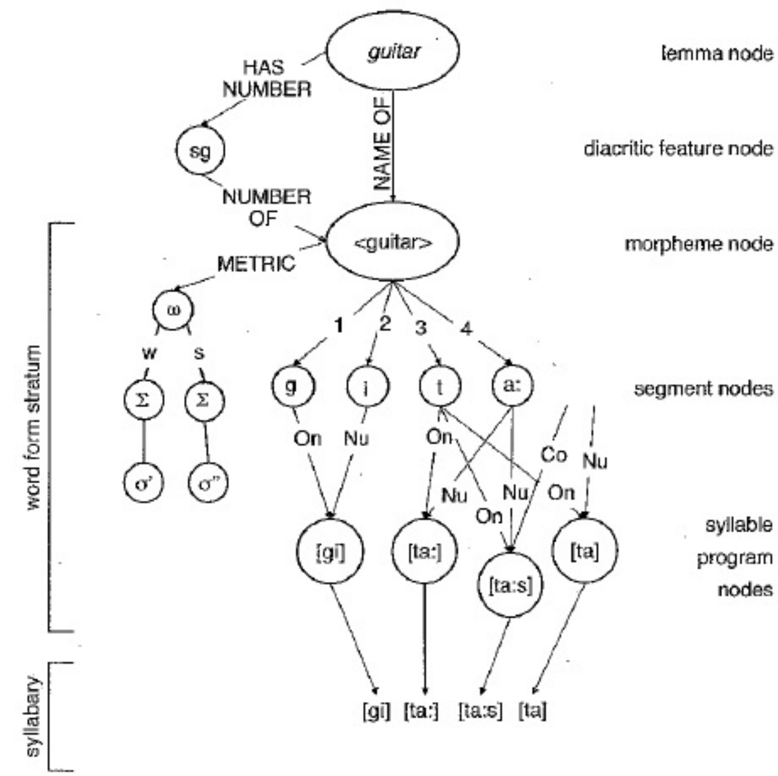
# What about prosody?

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

**Model assumptions** (Fromkin 1971, Shattuck-Hufnagel 1979, Dell 1986, Levelt et al. 1999)

- Constructing a speech plan is fundamentally a matter of selecting segments (and perhaps sub-syllabic units)
- Metrical structure is mapped to a prosodic frame, but may be referenced via diacritics. Explains why stress errors are rare.

**Word-form retrieval in WEAVER++**




**Question**: tone is lexical in many languages, but is a suprasegmental. How is tone processed in phonological encoding, or simply diacritically represented?

# Active debate: is tone part of phonological encoding?

| | **Yes!** Wan & Jaeger 1998, Gandour 1977, Shen 1993, Wan 2006 | **No!** Chen 1999, Roelofs 2015, Kember et al. 2015 |
|---|---|---|
| **Parallels:** | ***Tone is like segments***, can be mis-selected, and therefore tone must be represented linguistically 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 incorporated 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

***Alderete, Chan, and Yeung 2019, Cognition***

**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 |

***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.

# Interactivity

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

- Interactivity effects are a hallmark of processes in 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

source intended — ***same tone***

**Illustration**: … dzau**22** da:**22** …

↓

**dz**a:**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 they are, the more likely two tones 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

***Alderete, Chan, and Yeung 2019, Cognition***

| | ***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 requires full linguistic representations in phonological encoding, instead of diacritics (cf. stress)

# General conclusions

## Methods really matter in speech error research

- The sound patterns we wish to explain are different in different speech error corpora: <u>99% vs. 94.5% regularity</u>
- Models implications need to be studied from solid empirical ground.

## Competing explanations

- Phonological and psycholinguistic theory sometimes have competing accounts: <u>syllable frames vs. sequence frequency effect.</u>
- Look to explanations intrinsic to language production models first, before motivated external constructs.

## Converging explanations

- Phonological and psycholinguistic investigations sometimes converge: <u>tone in phonology, as a planning unit</u>.
- Linguistics still a source for important insights into production processes.

# Problems raised by the research

**Markedness vs. Frequency in sound errors** (Goldrick 2002, Shattuck Hufnagel 1979)
Markedness is an important grammatical construct at the heart of constraint-based grammar. Does markedness shape speech errors (towards unmarked patterns) just like phonology (see Goldbrick & Daland 2009), or could the same effects be predicted by phonological type frequency?

**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.

**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 |