

Phonological regularity, perceptual biases, and the role of grammar in speech error analysis

John Alderete, Simon Fraser University

in collaboration with:

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

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, Nootboom 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.”

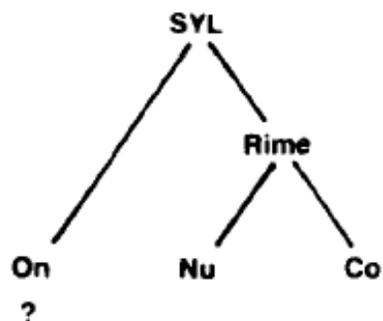
Regularity as a hard constraint

Early models: grammatical regularity not really explained, but the result of a built-in “structural matrix”

Spreading-interactive model of Dell 1986

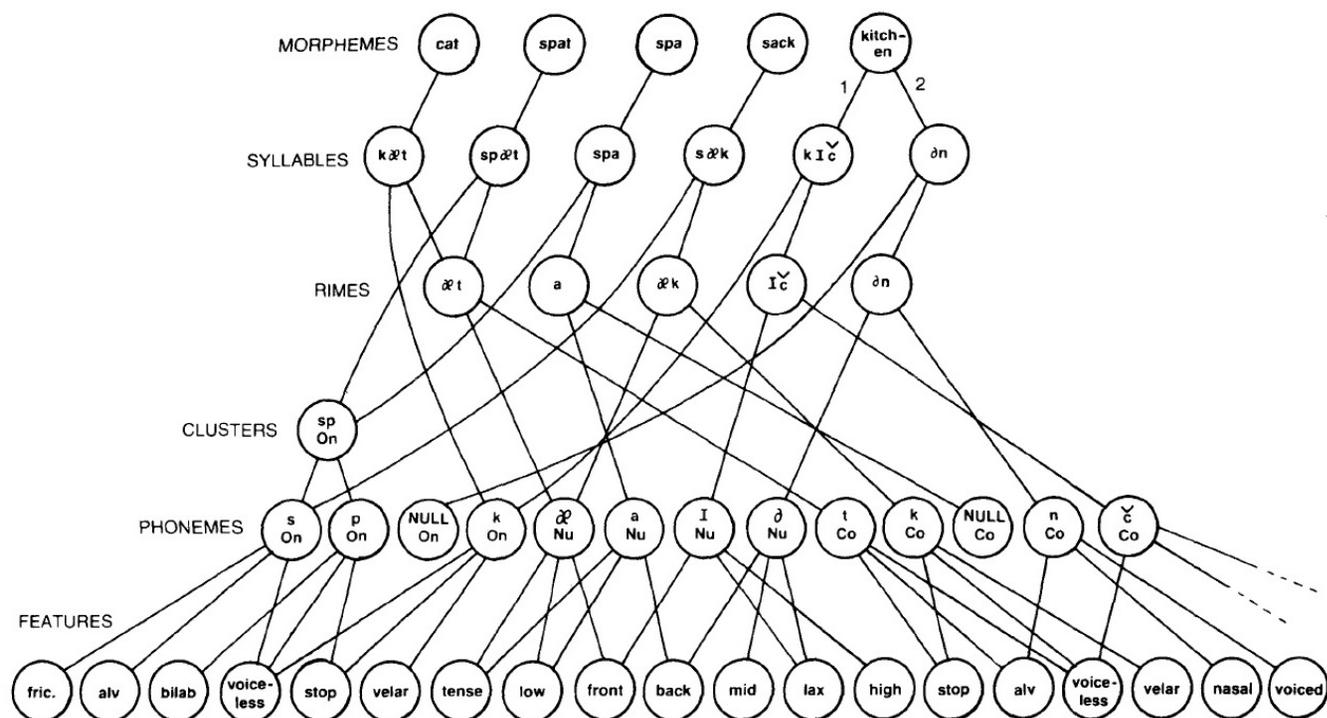
- Mental lexicon: activation dynamics for selecting valid linguistic units
- Tactic frames: productive capacity for language (builds sentence trees, word trees, and syllables. (a.k.a. ‘structural matrix’) ***This is a clear role for grammar!***

Tactic frame



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

Mental lexicon (fragment)



But regularity is not a hard constraint

Stemberger 1983: phonological regularity is very high, but speech errors do violate English phonotactics, approx. 1% of the time (37 violations/6,300 examples); standard of 99% phonological regularity

... in the first floor /**dl**orm — dorm room

I /**sth**ough— thought I said ‘moff’

... knowledge of the cooperative /**rp**in— principle

Problem: need a model that can produce phonotactic violations (rarely); not possible if regularity is a hard constraint.

Issue (Stemberger): the dominance of phonological regularity in speech errors does not entail that speech errors are controlled directly by phonotactic constraints—other independently needed mechanisms sensitive to frequency could be at work.

Dell et al. 1993: regularity without tactic frames

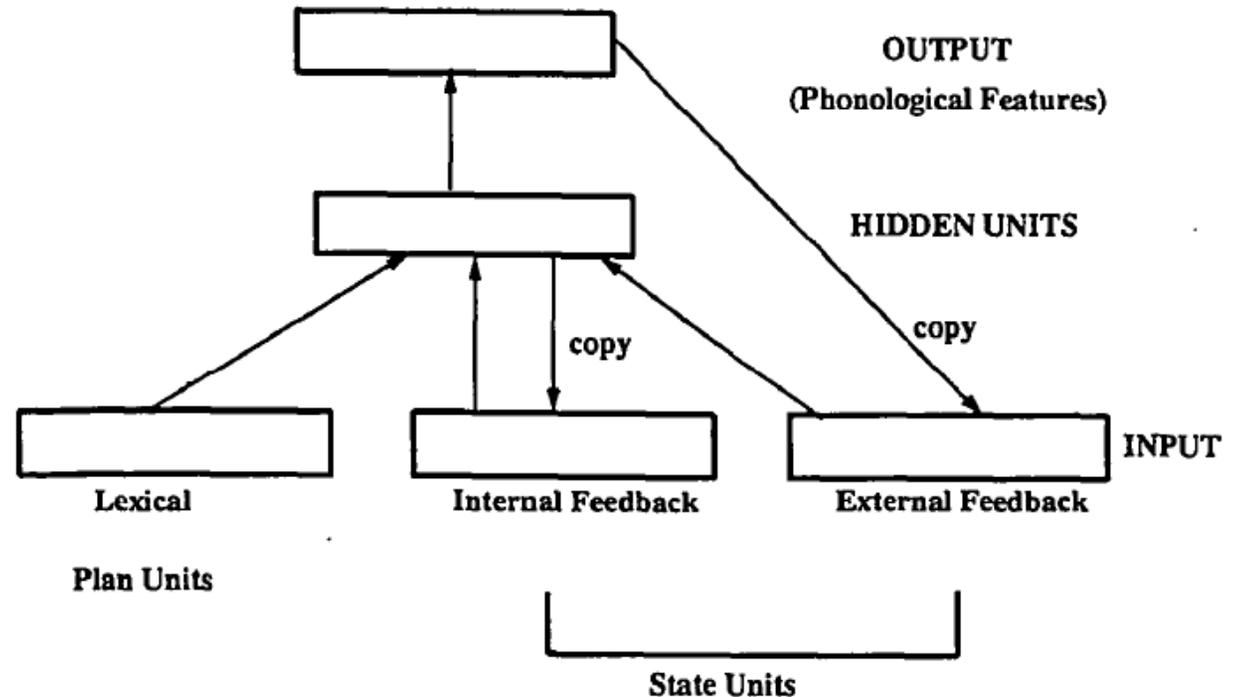
Dell et al. (1993): simple recurrent network proposed as a model of phonological encoding. Trained on a sample of English words and tested against a set of phonological benchmarks characteristic of speech error patterns (i.e., phonological regularity, CV substitutions, syllabic constituent effect, word-onset asymmetry)

Network features:

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)



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**

Upshot: regularity seems to be achievable without tactic frames (But a little below Stemberger's standard of 99%)

Questions

Just how much phonological grammar is there in phonological encoding?

The success of Dell et al's SRN suggests that some phonological structures (e.g., syllable templates) may not need to be formal mechanisms in model of language production processes.

Is cross-linguistic markedness a factor in the structure of speech errors, and if so, how is it incorporate into phonological encoding?

- Markedness is an important ingredient to contemporary phonological grammar (Constraints and Repairs, Optimality Theory, HG, MaxEnt)
- Markedness has also be argued to be a factor in the structure of speech errors (Blumstein 1973, Goldrick 2002, Goldbrick and Rapp 2007, Romani and Calabrese 1998), or not (e.g., Shattuck-Hufnagel and Klatt 1979)
- If markedness is a factor in the structure of speech errors, how is it included in model implementations?

How does methodology affect description of phonological regularity?

- Perceptual biases may reduce the rate of phonologically illicit errors (Dell et al. 1993) and other phonological effects (Alderete & Davies 2016)
- If speech errors are collected in such a way that these biases are reduced, does that effect phonological regularity and other phonological factors?

Focus and approach

SFU Speech Error Database: principal dataset

Describe a methodology for collecting speech errors and demonstrate that it is more reliable and robust to perceptual bias than prior work. SFUSED methods for data collection/analysis.

Re-assess phonological regularity in English speech errors

Just how phonologically regular are speech errors? (How common are phonotactic violations?)

Examine dimensions of cross-linguistic markedness to see if it has a major impact on the structure of speech errors.

Preview of SFUSED Cantonese (*Results preliminary*)

Progress report on recent data on tone slips in Cantonese

SFU Speech Error Database (SFUSED)



Current languages

SFUSED English (10,104 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

SFUSED English interface



Record ID no. 12470

Last Modified 2017-09-27

Total Completed 4712

Major Class Fields

Master Type

Altern. MType

Level

Type

Direction

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

Orthographic:

Phonetic:

Error

Word Bounded? Y N

Clipped? Y N

Source Different Talker? Y N

Corrected?

Word Fields

POS:

Lexical Word?

Open/Closed

Regular/Irregular

Error-Intended Semantic Relationship

Error-Intended Morphological Relationship

Sound Fields

Supplanted Intended:

Intruder:

Source Sound

CV Structure:

Syllabic Role:

Word Position:

Whole Syllable:

Error action is in:

Given Record Fields

Researcher

Found Date

File

Podcast

Time stamp

Online/offline

Talker Self?

Speaker Sex

Spreadsheet

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

Key methodological decisions

Offline collection from audio recordings (see Chen 1999 on Mandarin speech error database)

- errors collected from podcasts on different topics
- podcasts selected for having natural unscripted speech, usually Western Canada and U.S. (Midlands dialect 'Standard American')
- multiple podcasts (8 currently) with different talkers, approx. 50 hours of each podcast
- record dialectal and idiolectal features associated with speakers (because habitual, so not an error); listeners develop expectations about individuals

Multiple data collectors

- reduces collector bias, allows it to be studied (collector ID associated with all records)
- total of 13 data collectors

Training regime

- undergraduate students with introduction to formal linguistics, phonetics and phonology
- given phonetic training in transcription and tested for transcription accuracy
- introduction to speech errors, definition and illustration of all types
- training through listening tests: assigned pre-screened recordings, asked to find errors; learn 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 established protocol for finding errors in audio recordings, submit errors in spreadsheet format
- data analysts (must be different than collector) verify the error, classify it using the SFUSED fields

Advantages of methodology

Summary of findings from Alderete & Davies 2016

Reliability and data quality

- audio recording supports data collection separate from verification by another researcher; typical 25% of proposed errors don't meet standards
- with 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, which allows measures that are not possible with online collection, e.g., collection metrics ("minutes per error")
- 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" (which can be recovered)
- 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
- less 'easy to hear' and more 'hard to hear' speech errors, reduce impact of perceptual biases
- collect more errors that occur in fast speech

Offline: less 'easy to hear' errors

Online data collection (most prior research): requires on-the-spot observation, only collect errors with high degree of confidence.

Offline (SFUSED): collection from audio recordings, can replay and listen to slow speech

Result: speech error patterns in SFUSED more diffuse, less concentrated in highly salient errors like blends or exchanges.

Exchanges in SFUSED Ex. *We can just wrap mine in a /torn /korkilla* (corn tortilla, 1495)
Early data collection had 1,100 errors collected online. Sample balanced for experience.

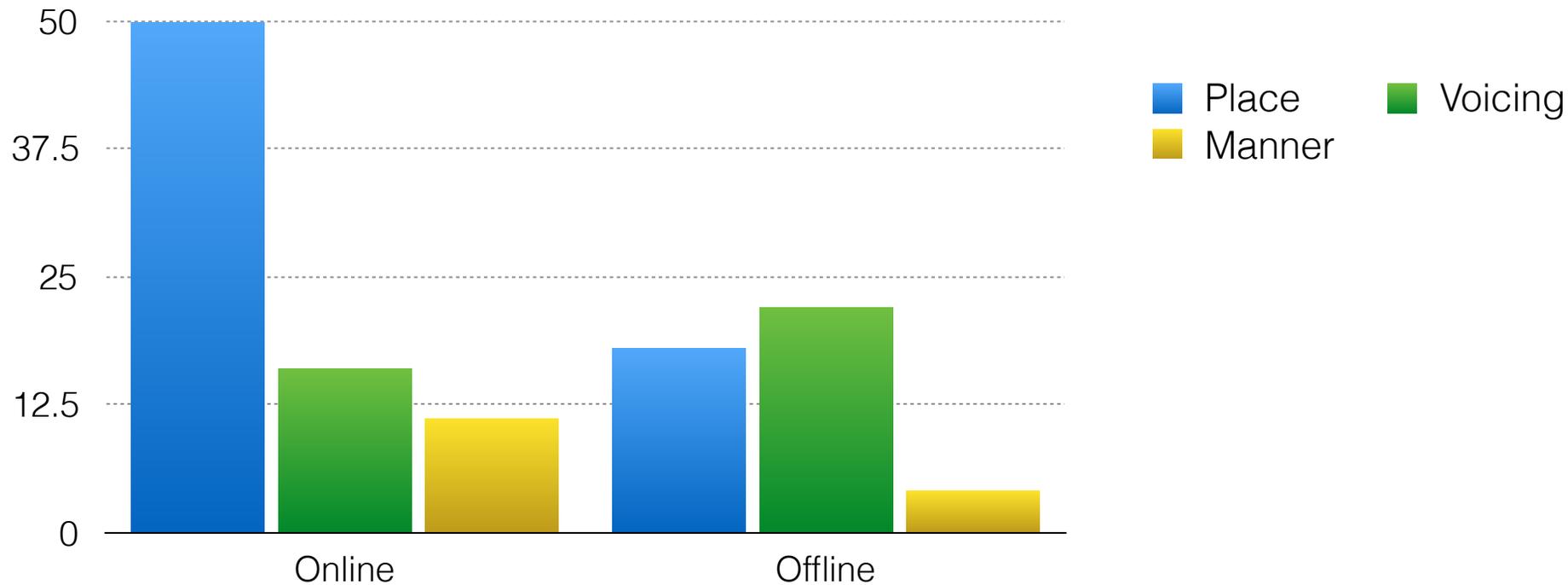
| | Offline | Online |
|-----------|------------------|------------------|
| Morphemes | | 6 |
| Phrases | | 1 |
| Sounds | 1 | 25 |
| Words | 1 | 15 |
| Totals | 2 (0.38% of 533) | 47 (5.6% of 839) |

Offline: more 'hard to hear' errors

Finding: errors in mis-pronunciation are easier to detect in place than voicing (Cole et al. 1978).

Online: counts reflect this perceptual bias

Offline: counts don't reflect this bias; more voicing errors detected than place.



Summary: online vs. offline (Alderete & Davies 2016)

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

How does methodology affect data composition?

How phonologically regular are speech errors in SFUSED English?

English phonotactics

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)

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

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 fjV fw fl 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)

Examples of phonotactic violations

Substitutions

Illicit onsets/appendices

1500 ... by the maps at the ^selection /**[ʃkrin]** (screen)

5739 ... they shoot, /**[ʒu]** shoot The Thick of It ... (you)

Illicit codas/rimes

1245 ... Their HOV /**[laɪŋ]** xxx lane is like one driver (lane)

5898 Vin Diesel got kicked off of /Rei**[ŋ]**deer Games ... (reindeer)

Nonnative sounds

5964 ... first of all, Katrina /**[kly]**= clearly defined (clearly)

Additions

Illicit onsets, appendix + onset

49 ... get the Ferrari down a /**[flju]** xxx few ^floors? (few)

1278 I don't like the ^/**vr**iral ^marketing. (viral)

5599 ... talking a ^dream, what that ^dream /**[mr]**eans ... (means)

Illicit codas/rimes

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

(SFUSED record ID # on left)

More examples

Deletions

3954 ... Lisa, /**S**reech and Lisa. (Screech)

8943 ... I think you're a /**h**u[**ŋ**ə]= hunk-a-rama.

Exchanges

4581 ... the children in the trailer for /**M**oon[ra**ŋ**] /**K**eez= Moonrise Kingdom.

Sequential Blends

4453 ... A diary is a /**[sb]**ook xxx a very special book.

5278 ... you can't quite /**[pjɪrt]** xxx put your finger on.

7211 ... because we /**[spɪlkf]** xxx we, we speak film

Word Blends

870 /**[sastæ]** makes me frisky. (pasta, sauce)

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

7270 ... /so[**m-bw**ʌn-di] xxx uh in the ... (someone, somebody)

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

Perceptual bias: missed phonotactic violations

Conjecture: prior research points out that there is probably a perceptual bias against phonotactic violations (e.g., Cutler 1982). Listeners may regularize them or simply fail to hear them.

Probe: Alderete and Davis 2016 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: Dell et al. 1993 point out that there is probably a perceptual bias against phonotactic violations. Listeners may regularize them or simply fail to hear them.

Probe: counting all sound errors and blends, % of phonotactic violations higher ($X^2 = 16.9618$, $p < .05$); note effect does not depend on what counts as a violation.

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

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: **96.5% 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.

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 likely 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: other potential roles of grammar?

Markedness

- Markedness is “the stuff” of most grammars in contemporary phonology. If we could find a role for markedness, this would be a clear role.
- Focus is on cross-linguistic markedness, not language particular markedness relations, because latter is hard to separate from frequency.

Frequency

- Frequency structure is increasingly a part of formal grammar, e.g., weights in Harmonic Grammar and MaxEnt grammar.
- Often overlaps markedness (marked is less frequent), but not always.
- Language production: frequency is a standardly assumed output bias in language production (Dell 1986); may be difficult to separate from its use in grammar

Feature specification (‘anti-frequency’)

- Feature specification is a core assumption in linguistic grammar, and has also been argued to account for speech error facts (Stemberger 1991)
- Specified sounds (because contrastive) override unspecified (because predictable): e.g., palatal bias in consonant substitutions.

***How does segmental markedness
shape speech errors?***

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], also 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

Test: single feature consonant substitutions

Procedure: take a consonant confusion matrix ($N= 1,506$)

| Sound Suppl.. | Sound Intruder | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------------|----------------|----|----|----|----|----|----|---|----|---|----|---|----|---|----|----|----|---|----|---|----|----|----|---|---|--|
| | 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 | 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 | 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 ($N= 1,506$)

| 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 | | | | |
| ð | | 1 | 3 | 3 | | 1 | | | 2 | 2 | | 1 | 1 | | | 1 | 1 | 1 | | | | | | | 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 | | | | | | | | | | | | | | | | | | | | | | |
| n | | | | 3 | 12 | 1 | | | | | | | | | | | | | | | | | | | |
| ŋ | | | | | | | | | | | | | | | | | | | | | | | | | |
| l | | | 1 | | | 1 | 3 | | | | | | | | | | | | | | | | | | |
| r | | 1 | | | 1 | | | | | | | | | | | | | | | | | | | | |
| w | | 1 | 2 | 1 | 2 | | | | | | | | | | | | | | | | | | | | |
| 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* produce 50 times in 1000

b → p 14 *b* produced 29 times in 1000

Baseline frequencies: estimating relative risk

| | Event | General Population |
|-------------|-------|--------------------|
| Condition 1 | a | b |
| Condition 2 | c | d |

mutually exclusive

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

Baseline frequencies: estimating relative risk

| | Event | General Population |
|-------------|-------|--------------------|
| Condition 1 | a | b |
| Condition 2 | c | d |

mutually exclusive

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

Stemberger 2007
data from sfusedE

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

$$RR(sz) = 7.07$$

Baseline frequencies: estimating relative risk

Stemberger 2007
data from sfusedE

| | Event | General Population |
|-------------|-------|--------------------|
| Condition 1 | a | b |
| Condition 2 | c | d |

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

mutually exclusive

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

$$RR(sz) = 7.07$$

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

Baseline frequencies: estimating relative risk

Stemberger 2007
data from sfusedE

| | Event | General Population |
|-------------|-------|--------------------|
| Condition 1 | a | b |
| Condition 2 | c | d |

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

mutually exclusive

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

$$RR(sz) = 7.07$$

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 $\log(RR) = 0$):

$$\log RR \in \left(\log \frac{a/b}{c/d} - 1.96 \sqrt{\frac{1}{a} + \frac{1}{c}}, \log \frac{a/b}{c/d} + 1.96 \sqrt{\frac{1}{a} + \frac{1}{c}} \right) \quad \text{Agresti 1996}$$

Example: $\log(RR(sz)) = 1.956$, 95% confident $\log(RR) \neq$ 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.

Take home: weak support for a role for markedness in consonant substitutions.

Take homes

New standard for regularity: errors of phonological encoding are phonotactically regular about 94-95% of the time (cf. prior standard of 99%).

Methodology: this new standard is documented using methods less prone to bias; so lower standard is likely the result of perceptual bias.

Grammar in phonological regularity? Syllable structure templates do not seem to be necessary to overall phonological regularity. Models of phonological encoding without them seem consistent with the fact.

Cross-linguistic markedness? An examination of the markedness of consonant substitutions did not reveal a strong role for markedness independent of frequency.

***How do the linguistic structures
of Cantonese inform language
production processes?***

Overview: SFUSED Cantonese

Size

2,535 speech errors of all types (sound errors, word errors, syntactic errors)

Scope and research questions

- How does the reduced syllabary in Cantonese affect phonological encoding (cf. Chen 2000 on syllable level errors in Mandarin)?
- How is native phonotactics obeyed in Cantonese errors, is there greater degree of phonotactic violations? (initial approximation of 6% irregularity, like English)
- Does frequency and/or markedness shape speech errors in Cantonese?
- How is tone processed in language production? Early integration (in phonological encoding) or later process similar to prosody.

Methods

- Sources: audio recordings of unscripted speech from three different podcasts. All errors have audio back-up
- Data collection: four native speakers trained in sfused data collection
- Classification: all data submitted were verified (checked for casual speech and phonetic processes, change of speech plan, idiolectal features) and classified using sfused variables

SFUSED Cantonese interface



Record ID no.
Total Completed:
Last Modified:
Record Completed? Y N

Confirmed By: mc qc cn gf
Keep? Y N Reason to exclude:

Major Class Fields

Master Type
Alt. MType
Level
Type
Direction

Contextual? Y N

Right Lexeme? Y N NotAppl

Specific Class Variables

Obvious Malapropism Y N NotAppl

Phonotactic Violation

Gradient Type

Morpho Cats

Check all Complex Processes:

Sub Del Tone Exch
 Add Shift Grad Blend

Example Fields

譬如阿曾江呀嗰啲呢,應該會 /kei3= xxx 幾廿 xxx 幾廿年,都仲係,仲係嗰個.

| | Intended | Error | Source |
|---------------|--|--|---|
| Orthographic: | <input type="text" value="幾"/> | <input "="" type="text" value="kei33="/> | <input type="text"/> |
| Phonetic: | <input type="text" value="gei35"/> | <input type="text" value="kei33"/> | <input type="text"/> |
| Clipped? | <input checked="" type="radio"/> Y <input type="radio"/> N | Wd Bounded? <input type="radio"/> Y <input checked="" type="radio"/> N | Corrected? <input type="text" value="y"/> |

Word Fields

Lexical Word?
Part of Speech:
Open/Closed:
Error-Intended Semantic Relationship
Error-Intended Morphological Relationship

Sound Fields

| | Supplanted Intended: | Intruder: | Source |
|---------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| CV Structure: | <input type="text" value="C"/> | <input type="text" value="C"/> | <input type="text"/> |
| Syllable Role: | <input type="text" value="Onset"/> | <input type="text" value="Onset"/> | <input type="text" value="NotAppl"/> |
| Word Position: | <input type="text" value="Initial"/> | <input type="text" value="Initial"/> | <input type="text" value="NotAppl"/> |
| Whole Syllable: | <input type="text" value="gei"/> | <input type="text" value="kei"/> | <input type="text"/> |
| Tone of Syllable: | <input type="text" value="35"/> | <input type="text" value="33"/> | <input type="text" value="NotAppl"/> |
| Error Action Is In: | <input type="text" value="Onset"/> | | |

Observation: tone of the error sounds clipped.

Imported Record Fields

Researcher
Found Date
File
Time stamp
Speaker Sex
Podcast
Confidence Intended
Confidence Transcription
Personal Info? Y N

Identical Neighboring Segment?

Y N

Phonologically Illegal?

Y N

Triggers Resyllabification?

Y N

Composition of the database

| | |
|---------------------------|-------|
| Complex Set of Processes | 332 |
| Extreme Reduction | 36 |
| Lexical Substitution | 88 |
| Morphological Error | 24 |
| One of a Kind | 7 |
| Phonetic Error | 60 |
| Phonological Addition | 115 |
| Phonological Deletion | 89 |
| Phonological Shift | 1 |
| Phonological Substitution | 1,178 |
| Role Mis-assignment | 10 |
| Sentence Blend | 2 |
| Sequential Blend | 36 |
| Tone Error | 444 |
| Word Addition | 43 |
| Word Blend | 30 |
| Word Deletion | 42 |
| Word Shift | 9 |

Some observations:

- Sound errors dominate
- Non-trivial number of errors involve combinations of basic operations (additions + deletions)
- Large body of tone errors
- Sizeable body of word errors

Background: tone in language production

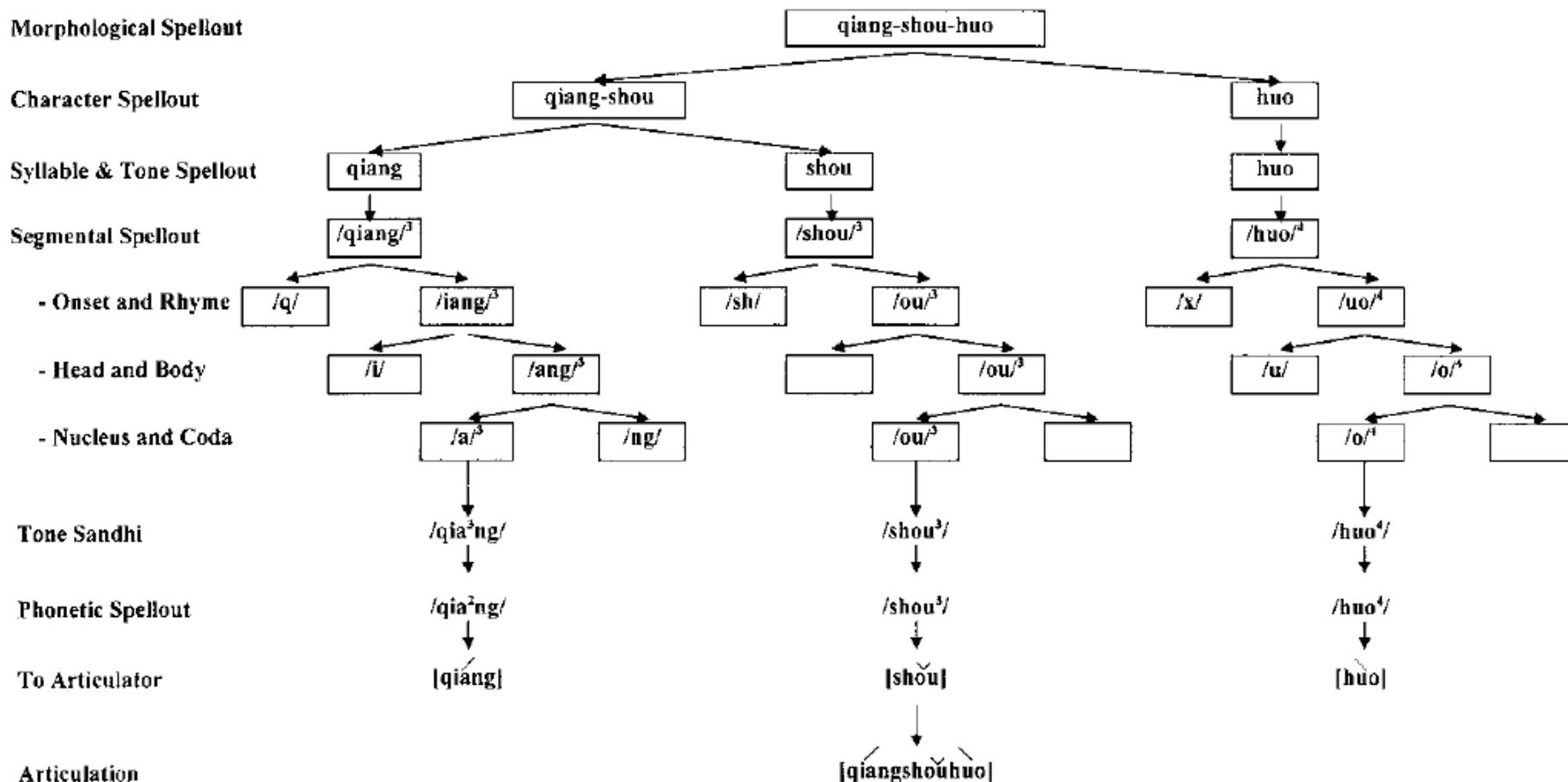
Tone is processed like Cs and Vs (e.g., Gandour 1977)

- Tone is part of the lexical representation of words, so requires encoding in a speech plan.
- Tone involved in the same kinds of errors and segmental structure, e.g., anticipations, preservations of contextual sounds.
- Tone errors in abundance, so require model of encoding tone (see also Moser 1991, Shen 1992)

Tone is not like Cs and Vs, more like stress (Chen 1999)

- Tone errors in Mandarin are actually exceedingly rare, only 24 examples in corpus of 987 slips of tone, most of which ambiguous
- Tone doesn't prime responses in implicit priming experiments (Chen, Chen, Dell 2002), cf. atonal syllables
- Tone is therefore like stress: relatively immutable to errors

Tone processing in Chen 1999



Assumption: tone is not part of early phonological encoding; represented diacritically in early encoding (cf. stress) and then implemented in late production processes (perhaps part of articulation)

Wan and Jaeger 1999

Tone slips are not uncommon in Mandarin

- 78 tone errors documented, out of 785 total speech errors

Tone is relevant to the lexical organization of words

- In 133 lexical substitutions, greater than chance occurrence that the intended and error word contain the same tone (49%, cf. 25%).
- If tone is encoded and activated in lexical retrieval, then it will provide feedback in words that share the same tone, increasing its chance of occurring in an error.

Tone errors feed tone sandhi

The output of a tone error is the input to tone sandhi; not really consistent with diacritic representation of tone errors.

Example (p. 444):

Intended: na35 jow21 maj51 paw51 tsɿ21

Error: na35 (jow21) maj21 paw51 tsɿ21

Sandhi: na35 jow35 maj21 paw51 tsɿ21

Are tone errors really rare?

Argument: 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 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% |

Illustration: tone error in SFUSED Cantonese



Record ID no.
 Total Completed:
 Last Modified:
 Record Completed? Y N

Confirmed By: mc qc cn gf
 Keep? Y N Reason to exclude:

Major Class Fields

Master Type
 Alt. MType
 Level
 Type
 Direction

Contextual? Y N

Right Lexeme? Y N NotAppl

Specific Class Variables

Obvious Malapropism Y N NotAppl
 Phonotactic Violation
 Gradient Type
 Morpho Cats

Example Fields

咁^就 困咗佢嘍^佢哋 /li22 有,有一架戰車㗎嘛。

| | Intended | Error | Source |
|---------------|--|-------------|--|
| Orthographic: | 呢 | li22 | 佢哋 |
| Phonetic: | li55 | li22 | kcei23dei22 |
| Clipped? | <input type="radio"/> Y <input checked="" type="radio"/> N | Wd Bounded? | <input type="radio"/> Y <input checked="" type="radio"/> N |
| Corrected? | <input type="text"/> | | |

Word Fields

Lexical Word?
 Part of Speech:
 Open/Closed:
 Error-Intended Semantic Relationship
 Error-Intended Morphological Relationship

Sound Fields

| | Supplanted Intended: 55 | Intruder: 22 | Source |
|---------------------|--|--|--|
| CV Structure: | <input type="text"/> | <input type="text"/> | |
| Syllable Role: | <input type="text" value="WholeSyllable"/> | <input type="text" value="WholeSyllable"/> | <input type="text" value="WholeSyllable"/> |
| Word Position: | <input type="text" value="InitialAndFinal"/> | <input type="text" value="InitialAndFinal"/> | <input type="text" value="Final"/> |
| Whole Syllable: | <input type="text"/> | <input type="text"/> | |
| Tone of Syllable: | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| Error Action Is In: | <input type="text"/> | | |

Descriptive statistics

Counts

446 tone slips, 28 likely phonetic in nature (failure to reach onset/offset tonal target), 418 phonological tone errors

Direction

| | Count | % |
|-------------------------------|-------|------|
| Anticipations | 103 | 24.7 |
| Perseverations | 113 | 27 |
| Anticipations & Perseveratons | 164 | 39.2 |
| Non-contextual | 38 | 9.1 |

- Mostly contextual errors (though definition is wider than prior studies, 10 characters upstream/downstream)
- Regardless of word envelope, there is a non-trivial set of contextual errors, suggesting mis-selection via activation dynamics

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 than substitutions with more feature changes.

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, 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 in SFUSED Cantonese

Tone confusion matrix (N=418)

| | Intruder | | | | | |
|----------|----------|----|----|----|----|----|
| Intended | 22 | 33 | 55 | 23 | 35 | 21 |
| 22 | | 37 | 7 | 25 | 18 | 26 |
| 33 | 33 | | 7 | 16 | 16 | 6 |
| 55 | 6 | 17 | | | 12 | 2 |
| 23 | 15 | 9 | 7 | | 19 | 11 |
| 35 | 20 | 20 | 20 | 16 | | 1 |
| 21 | 31 | 5 | 2 | 14 | | |

Tone features

-basic six tone system

-use features from Yip 2001 and Barrie 2007 ([upper], [raised], [contour]),
similarity calculated as function of shared feature values

Result

| | 1 Change | 2 Changes | 3 Changes |
|----------|----------|-----------|-----------|
| Observed | 213 | 176 | 29 |
| Expected | 167.2 | 196.5 | 54.3 |

$$\chi^2(2) = 26.46$$

$$P < 0.0001$$

Tone in SFUSED Cantonese: interim summary

Facts supporting early integration of tone encoding with encoding of segments and words

- 1) Large number of tone slips
- 2) Non-negligible number of contextual tones
- 3) Phonological similarity effect on tone errors

Facts as yet undocumented

- 1) Feedback from tone on lexical selection (cf. Wan and Jaeger 1998)
- 2) Feedback from tone on selection of segments or syllables

Future projects with SFUSED Cantonese

Phonological encoding of tone: lots of questions

- correlation analysis for phonological similarity, correct features
- does tone produce feedback in segmental encoding (or even syllables, given important role of syllables in Chinese languages)?
- markedness for tones?
- if tone actively encoded, how aligned with syllables?
- complex processes: segmental and tonal errors together?

Phonological regularity and markedness

- regularity and model predictions
- consonantal substitutions, featural markedness
- syllable structure markedness: additions vs. deletions

Syllable errors and the representation of the syllable

- higher incidence of whole syllable substitutions (see Chen 2000)
- implications for representation of syllables in lexical network (see O'Seaghdha et al. 2010 on proximate units in Mandarin)

Future projects

SFUSED is designed as a multi-purpose database and we actively seek out new project proposals.

New projects can improve the data quality through data cleaning, and also extend the database by introducing new fields.

The database can be imported to Tableau and then explored much faster and more deeply.

Projects can combine questions from linguistic theory (e.g., ‘psychological reality of X’) or experimental paradigms.

Contact: alderete@sfu.ca

Contributors to SFUSED

Director/analyst/data collector:

John Alderete

Research associates

Paul Tupper (SFU)

Henny Yeung (SFU)

Alexei Kochetov (Toronto)

Stefan A. Frisch (USF)

Monica Davies (UBC)

Analysts/data collectors

Holly Wilbee (English)

Monica Davies (English)

Olivia Nickel (English)

Queenie Chan (Cantonese)

Macarius Chan (Cantonese)

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)

Heikal Badrulhisham (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 on a particular pattern.

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.

However, some patterns differ in two datasets, limitations: % 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

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 | \bar{m} | \bar{v} | 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 |

From Alderete and Davis 2016