Towards the quantification of *Peggy Babcock*: The role of syllable position in speech errors

Tine Mooshammer, Haskins Labs
Raj Dhillon, Haskins Labs & Yale
Hosung Nam, Haskins Labs

Mark Tiede, Haskins Labs & RLE MIT
Louis Goldstein, Haskins Labs & USC
Elliot Saltzman, Haskins Labs & BU
The title of this talk comes from a famous paper on speech errors by Butterworth and Whittacker. They give this as they call it ‘notorious’ Peggy Babcock example because this sequence elicits many errors as can be seen here: Bagcock, Bagpock or Bagpop as outcomes.

Why this example is relevant for this talk will be shown later.

Our central question is whether there are more errors in the syllable onset or in the syllable coda. Up to now evidence has been found for both. Some studies found more errors in the onset, but that has been attributed to a perceptual masking effect by Cathi Browman.

Others have found more errors in the coda.

To investigate this question further we use EMA data in order to avoid the perceptual bias towards onset errors and to quantify gradient speech errors, as has been found by instrumental phonetic studies.
There exist two classes of models for explaining the occurrence of speech errors. The first group is based on the assumption that phonemes get mis-selected because they are activated at the same time on the phonological level. In traditional models, the phoneme which gets most activation will be articulated, and the competing one not. This leads to categorical phoneme sized errors on the phonetic output level. Within the more recent Cascading Activation model the activation of all co-active phonemes cascades to the motor level and therefore it can account for the occurrence of gradient, non-canonical errors. Both accounts assume some kind of baseline co-activation by repeating parts of words. But this does not explain asymmetries due to syllable position:

More frequent coda errors have been attributed to re-activation within the sequential cuing model. For word sequences such as pin pick with a mismatch in the coda, saying the second word PICK reactivates the whole word PIN because initial consonants index and activate lemmas. When the speaker reaches the /k/ in PICK, the reactivated coda /n/ of PIN competes with the /k/ and an error might occur. This does not happen for the onset mismatch such as PIN TIN because TIN doesn’t reactivate PIN.
The second model, based on frequency locking, assumes that errors occur in order to simplify competing frequencies between the articulators. As can be seen here for COP COP (the upper two panels) each tongue dorsum gesture is associated with a lip gesture. Therefore they are moving in a 1:1 frequency ratio. For TOP COP tongue dorsum and lips are moving at different frequencies. Only every other lip gestures is associated with a tongue dorsum gesture, that is they are moving in a 2:1 ratio. After intrusive gestures occur we get a transition to a 1:1 ratio. How this a approach could account for asymmetries in onset and coda will be discussed later.
The aims of this study are to compare error rates and spatial variability for word repetitions with onset mismatch and coda mismatch. Additionally we introduced some more conditions:

We will look at sequences with mismatch in both, the onset and the coda. These items are of the type of the famous Peggy Babcock example, which was only mentioned in the Butterworth & Whitacker study but not further analysed.

According to the sequential cuing model this elicit should elicit about many errors as the onset mismatch condition.

Furthermore we added sequences with alternating missing onsets, such as TOP OP, and alternating missing codas, such as TOP TA. This will inform us whether the actual articulators have to be moving in a different frequency or whether the abstract frequency ratio mismatch elicits intrusions.

Bisyllabic words were included in order to find out whether the position within a syllable or within a word is relevant, since the medial consonant in words like PICKY is an onset and not a coda.
EXPERIMENT

- 3D EMA recordings
  - 3 sensors on the tongue
  - 2 sensors on the jaw
  - one on the upper and lower lip
  - 4 reference sensors
- nine speakers of AmEngl.
- word repetition to a speeded up metronome
  - trial duration 20 sec.
  - 10 sec stable, then speed up
This table shows the material that was recorded with the nine speakers. As can be seen, not all speakers produced the same number of trials. Apart from the alternating word repetition, the participants also produced the non-alternating simple repetitions.

Get rid of table. You can just say that the nimbler of tokens varied.
For calculating error rates intended gestures were labelled at the maximum constrictions, shown here as red lines for tongue tip maxima during the final /d/ in COD of the word sequence COD COB. These time-points were used for extracting the unconstrained alternating Lip Aperture amplitudes, again during /d/. The same was done during /b/ for the intended LA aperture gesture and the unintended TT movements.

We then calculated the interquartile split mean between the intended and the unconstrained positions. These means, shown as green lines, were used as thresholds to define several error types:

Reduction are defined as intended gestures below the mean value. Here shown in violet for reduced tongue tip positions.

Intrusions are defined as exceeding positions of the unconstrained articulator above the mean value. For example tongue tip positions during /b/ that are larger than the mean value, shown in green here.

Substitutions occur when there is a full intrusion of the unconstrained articulator and at the same time a full reduction of the intended articulator, here shown in orange.

Error rates were then calculated as percentages of number of syllables in one trial.
Analyzing Delta Measure

**Procedure**
(adapted from McMillan & Corley 2010)

1. extract articulatory positions during intended gesture for all sensors (here TD, TB, TT, JAW, UL, LL for /p/ in during *pod cod*)
2. calculate mean position (+) for all sensors
3. calculate Euclidean distances between mean positions and single data points (6 sensors X 2 spatial dimensions --> 1 distance)

→ measure of spatial variability within a trial

Prediction following McMillan & Corley:
delta is larger for alternating words than for the non-alternating controls

Our method for defining intrusions and reductions has been criticized by McMillan and Corley because it is based on an arbitrarily defined threshold, separating a continuum of possible values into artificial categories. Therefore we also adapted their method, the delta value, which is a measure for spatial variability.

The horizontal and vertical position of all moving sensors is extracted during the extrema of the intended gesture, here during the /p/s when the speaker was repeating POD COD, shown as red stars. Then the mean of all positions during /p/ was calculated.

Euclidean distances between the mean positions and all instances of for example /p/ are then calculated. These distances correspond to the delta measure.

As was found by McMillan and Corley deltas are larger for alternating words than for non–alternating controls, which are shown below for the /p/s in POD POD. They attribute the larger deltas in alternating trials to co-activation of phonemes cascading to the phonetic level.
Let’s first turn the questions whether we get more errors in the onset or in the coda.
As can be seen here coda mismatch as in POD POG induces much more substitutions, reductions and intrusions than the onset mismatch. The increase in reductions could be due to the optional substitution of final /t/ with glottal stop in American English. Excluding items with final /t/ indeed reduces the rate to 3.6%.
It also interesting to note that the most frequent error type for both mismatch conditions was intrusion confirming Marianne’s and Louis’ results.
### Results: Onset vs. Coda

**Spatial variability: delta measure**
- Significantly more variability for alternating word sequences (e.g. top cop, top tock) than for non-alternating (e.g. top top, tock tock) (McMillan & Corley 2010)
- No main effect of onset or coda mismatch
- Significant interaction: Larger increase of variability for coda mismatch

#### Summary for top cop vs. top tock
- Both measures, error rate + delta, show higher values for coda mismatch

The delta measure also increases more for coda mismatch than for onset mismatch when compared to the corresponding deltas for non-alternating control trials.
Therefore both, the error rates and deltas confirm the Segment Cuing Model which predicts more errors for coda mismatch because repeating the same onsets triggers competition in the coda.
Next we look at items with double mismatch, for example BABCOCK or as used here POP TOT. The Sequential Cuing model would predict a similar behavior as for onset mismatch. However, the error rates here show a huge increase for the coda errors in the double mismatch condition compared to the coda errors in the single mismatch condition. There is almost no difference for the onsets between single and double mismatch.
Results: The case of Ms. Babcock

Spatial variability: delta

- Non-alternating controls show smaller variability for double mismatch than for single mismatch.
- Single mismatch control trials: two alternating articulators (e.g., top top).
- Double mismatch control trials: one moving articulator (e.g., pop pop).
- Higher delta values for the coda than for the onset.
- In coda position, larger increase of delta for double mismatch than for single mismatch.

The deltas of the non-alternating condition differ between single mismatch and double mismatch, probably because of the number of articulators involved. For the single mismatch control trials, for example, top top, two articulators are moving whereas in the double mismatch control condition, for example POP POP, only one articulator is moving. Therefore, we get lower deltas for pop pop items than for top top items. For both, single mismatch and double mismatch, the deltas increase more in coda position than in onset position, but this increase is much larger for double mismatch coda position, as shown by the arrows. This is contrary to what would be predicted by the Sequential Cuing Model.
more intrusions and variability in the **coda** than in the onset

* cannot be explained by **Sequential Cuing Model** (Sevald and Dell 1994) because in the double mismatch condition (`pop tot`) different onsets trigger more errors and increased spatial variability in the coda than same onsets in the single mismatch condition

* **Frequency Locking:**
  * in the coupled oscillator model of syllable structure (Goldstein et al., 2006; Nam et al, 2009), onsets are more strongly coupled with the vowel than codas
  * for the **coda mismatch** case, the onset articulator moves with the higher frequency and the coda articulator with the lower frequency
  * due to the stronger coupling, the onset frequency exerts a greater pull on the low frequency articulator of the coda
  * because the coupling in the coda is weaker, it also gives way to this pull more often, leading to more frequent intrusions than in the onsets
  * `pop tot`???

within AP/Task dynamics, onsets are more strongly coupled with the vowel than coda. -->

In the coupled oscillator model of syllable structure (Goldstein et al., 2006; Nam et al, 2009), onsets are more strongly coupled with the vowel than coda
Summary and Discussion

- highest error rate and variability increase for **double mismatch** sequences

  - Argyro Katsika, Stefanie Shattuck-Hufnagel, Christine Mooshammer, Mark Tiede and Louis Goldstein
  Effects of prosodic grouping and phonemic content on error patterns in the production of repeated syllable strings

- the starting frequency ratio is not 1:2 but much more complicated
- alternation between **short** and **long** intervals for each articulator
- phase shift between the two articulators
- break-down
- **1:1 ratio**
- more complicated starting ratio might also explain why it is so much harder
CONCLUSIONS

- More errors and more variability in the coda than in the onset
- This cannot be explained by the SCM, because of the high number of errors we found for the double mismatch condition
- **Frequency locking**, combined with coupling differences in onsets and codas, can explain higher error rates in the coda and higher error rates for double mismatch condition
- **BUT** not the locus of increased error rates (coda of double mismatch items) until the frequency and analysis of this transition is better understood
Acknowledgments

Thanks to:

- NIH for support via NIDCD DC 008780
- Labeller: Hansook Hanie
- Advice: Jelena Krivokapic, Stefanie Shattuck-Hufnagel
- Speaker participants
- You for your attention