Morphological Information and Fine Phonetic Detail in English Irregular Verbs

Michelle Sims, Benjamin V. Tucker, Harald Baayen

University of Alberta

Mental Lexicon Conference - Windsor, Ontario - 1 July 2010
Introduction
Speech Variation
Irregular English Verbs
Speech Reduction Theories
Methodology
Predictors
Results & Discussion
Morphology
Frequency
Gang Size
Imageability
Discussion
Interpretation
Conclusions
Speech Variation

everyday speech is not careful but instead contains a great amount of phonetic reduction and variation

this acoustic signal is the fingerprint of speech processing
Irregular English Verbs

difference between the past and present tense is the alternation of a single phonetic segment (vowel)

sing / sang
run / ran
hold / held

morphological effects on vocalic variation in duration
Speech Reduction Theories

Lexical Confusability: the denser the gang, the less likelihood for reduction (Scarborough 2004)

Smooth Signal Hypothesis: lexical frequency will negatively correlate with vocalic duration (Aylett & Turk 2004)

Paradigmatic Signal Enhancement Hypothesis: the more common the vocalic alternation, the longer the vowel (Kuperman et al 2007)
Methodology

Buckeye Corpus of Spoken English

Irregular English Verbs

48 items
monosyllabic
vocalic alternation between past/present forms

7410 tokens total
Methodology: Predictors

linear mixed effects regression model

dependent variable: vocalic duration

independent variables: morphology, log frequency, log gang size, mean imageability rating

random effects factor: word lemma
Methodology: Predictors

Independent variable: log gang size

Large gangs:
/o/ to /u/
blow to blew
throw to threw
know to knew

Small gangs:
/ɔ/ to /ɛ/
fall to fell
- longer durations in the past tense
- higher frequency correlates with longer vocalic duration in the past tense
- no significant effect in the present tense
- bigger gangs correlate with shorter vocalic durations in the past tense
- no significant effect in the present tense
- more imageable words correlate with longer vocal durations in the past tense
- no significant effect in the present tense
Discussion

It appears that speech production is subject to morphology, frequency, gang size and imageability...

though their effects do not follow from current theories.
Discussion

Smooth Signal Hypothesis

predicted the more frequent a word is, the less information it carries & the less information a word carries, the shorter it will be

This Data

past tense vowels are longer when word frequency is high
Discussion

Lexical Confusability predicted that words belonging to denser neighborhoods will have longer durations

This Data

past tense vowels are shorter when the gang they belong to is denser
Discussion

Paradigmatic Signal Enhancement predicted that words following more common paradigms will have longer durations (increased acoustic salience)

This Data

past tense vowels are shorter when the vocalic alternation pattern they follow is more common
Discussion

Imageability predicted that more imageable words have longer naming latencies.

This Data

past tense vowels are longer when the word is more imageable
Discussion: Interpretation

competition with regularization

1. less pressure to regularize if in a large gang

2. hyperarticulation of high frequency words is a consequence of resisting regularization
Conclusions

Morphology plays a significant role in acoustic realization.

Lexical frequency and paradigmatic alternation also play a role in acoustic realization, though their effects are in surprising directions.

mnsims@ualberta.ca
Conclusions

Further Investigations:
- regular verbs
- vowel quality
- prosodic features
- probabilistic syntactic factors
- sociolinguistic features
- other measurements of vowel reduction
Irregulars
Regulars
Irregulars & Regulars

(a) Duration (ms) vs. Tense
(b) Duration (ms) vs. Regularity
(c) Duration (ms) vs. log Frequency
(d) Duration (ms) vs. Image Rating
Vowel Quality

Lax

Tense

Density

log Duration

Density

log Duration
Vowel Quality

<table>
<thead>
<tr>
<th>Tense by Tenseness</th>
<th>past</th>
<th>present</th>
</tr>
</thead>
<tbody>
<tr>
<td>lax</td>
<td>749</td>
<td>3301</td>
</tr>
<tr>
<td>tense</td>
<td>1929</td>
<td>4596</td>
</tr>
</tbody>
</table>
Methodology: Predictors

independent variable: imageability

Prado & Ullman (2009) found a verb’s imageability to co-determine its naming latency

mean ratings of imageability were gathered through a short study
Methodology: Predictors

study to gather imageability ratings:

112 participants
87 female, 25 males
17-63 years (71.4% 18-21 years)
native English speakers

rate 171 verbs (48 irregular, 123 regular) on how readily the verb produces a mental image

1 → 5
least imageable          most imageable
Methodology: Corpus

Buckeye Corpus of Spoken English

300,000 words

40 speakers
  20 male/20 female
  20 < 30 years/ 20 > 40 years

residents of Columbus, OH

speech from informal interviews


