A. Background

(1) Is phonology irrelevant to experimentally elicited well-formedness judgments?

(2) Summary of presentation
   - Experiments
   - Results
   - Analysis
   - Conclusion

(3) Experimentally elicited well-formedness judgments in controlled experimental setting.

(4) Experimentally elicited well-formedness judgments:
   (a) Speakers intuit a continuum of well-formedness paralleling phonotactic probability (Coleman & Pierrehumbert, 1997, Frisch et al., 2000).
       Example: /blɪk/ vs. /sfɪk/
   (b) Neighborhood Density is also known to influence well-formedness judgments (Ohala & Ohala, 1995, Greenberg & Jenkins, 1964)
       A word with many neighbors is judged to be more well-formed than a word with few neighbors.
       Example: /dæɡ/ (many neighbors) vs. /pənθ/ (few neighbors).
   (c) Boersma claims that phonological constraints play a role in determining well-formedness (2004)
       Example: i-prototypicality effects

(5) How do these three factors interact?

(6) We conducted an experiment to test whether a phonological constraint (independent of known factors) plays a role in the determination of well-formedness judgments, or if constraints are irrelevant in the determination of well-formedness.

(7) The constraint we tested is *COMPLEX (Prince and Smolensky 1993)
   (a) For languages like Mutsun or Hawaiian, *COMPLEX is high ranked. Only simple (one consonant) onsets are allowed.
   (b) For languages like English or Russian, *COMPLEX is ranked low. Simple and complex (2 or more consonants) onsets are allowed.
B. Experiments

(8) Method
We utilized an established well-formedness task common in psycholinguistic research. Subjects were asked to rate the acceptability of non-words.
Experiment 1: Responses ranged on a scale from 1 (best) to 7 (worst).
Experiment 2: Subjects responded yes or no to the question, “Is this word a possible word of English?”

(9) Subjects
Experiment 1:
20 native English speaking undergraduates from the University of Arizona participated in the experiment for extra credit.
Experiment 2:
30 native English speaking undergraduates from the University of Arizona participated in the experiment for extra credit.
Data from non-native English speakers was excluded.

(10) Materials:
(a) The subjects responded to a random ordering of monosyllabic nonsense forms. 21 total target items were controlled in triplets, (ie /voιk/, /floιk/, /stroιk/) varying only by onset length (1-, 2-, or 3-segment onset). Frequency and neighborhood density were matched across the three conditions.
(b) Target items were randomized with an additional 20 distractors.

(11) List of stimuli

<table>
<thead>
<tr>
<th>C</th>
<th>CC</th>
<th>CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>zιlм</td>
<td>twιlm</td>
<td>skιlм</td>
</tr>
<tr>
<td>gaiθ</td>
<td>staιθ</td>
<td>stιaιθ</td>
</tr>
<tr>
<td>voιk</td>
<td>floιk</td>
<td>stιoιk</td>
</tr>
<tr>
<td>viʃ</td>
<td>kwιʃ</td>
<td>skιʃ</td>
</tr>
<tr>
<td>dʒæntʃ</td>
<td>spæntʃ</td>
<td>stιæntʃ</td>
</tr>
<tr>
<td>basp</td>
<td>plasp</td>
<td>stιasp</td>
</tr>
<tr>
<td>ʃιrυ</td>
<td>klιrυ</td>
<td>stιrυ</td>
</tr>
</tbody>
</table>

(12) Procedure
Subjects heard tokens through headphones and responded by typing 1 through 7 (Experiment 1), and 1 or 2 (Experiment 2). Subjects were instructed that 1= possible English word, and 7= not possible English word, or alternately, for the second experiment, 1= yes, and 2=no.
(13) Possible outcomes
(a) Onset length is not a factor in determining well-formedness
(b) Onset length is a factor
   - Gradient response based on onset length.
   - Categorical response: We could see an effect of *COMPLEX (1 vs. n)

C. Results

(14) Experiment 1 (1-7 judgments): A significant main effect of onset length!
[F(2,38)=4.121; p<0.024]

(15) Mean Responses of Experiment 1

![Mean Responses of Experiment 1](image)

(16) Planned comparisons
(a) Subjects significantly [F(1,19)=6.362; p<0.021] judged non-words with two-segment onsets to be more well-formed with respect to English than non-words with a simple (single consonant) onset.
(b) Subjects significantly [F(1,19)=7.436; p<0.013] judged non-words with three-segment onsets to be more well-formed with respect to English than non-words with a simple (single consonant) onset.
(c) There was no significant [F(1,19)<1; ns] judgment difference for subjects between two- and three-segment onsets.
(17) Experiment 2 (Yes/No judgments): Again a significant main effect of onset length
\[ F(2,58)=16.179; p<0.001 \]

(18) Mean Responses of Experiment 2

(19) Planned comparisons
(a) Subjects significantly \[ F(1,29)=20.014; p<0.001 \] judged non-words with two-segment onsets to be more well-formed with respect to English than non-words with a simple (single consonant) onset.
(b) Subjects significantly \[ F(1,29)=24.969; p<0.001 \] judged non-words with three-segment onsets to be more well-formed with respect to English than non-words with a simple (single consonant) onset.
(c) There was no significant \[ F(1,29)<1; ns \] judgment difference for subjects between two- and three-segment onsets.

(20) One might suspect that the effect is a result of our materials, but…
(a) Bailey and Hahn (2002) found an effect of phonotactic probability and neighborhood density...
(b) We replicated Bailey and Hahn (2002) and found the same effect of neighborhood density and phonotactic probability
(c) Using a multiple regression we found an independent effect of onset complexity onset complexity, \( r^2 = 0.2146, p<0.001 \)
neighborhood density, \( r^2 = 0.2644, p<0.001 \)
probability, \( r^2 = 0.3762, p<0.001 \)

(21) Summary of results
(a) The phonological variable of onset length plays a role.
(b) The effect is 1 segment vs. \( n \) segments. No effect of 2 vs. 3 segments.
(c) Effect is in the opposite direction one might expect.
D. Analysis

(22) *SIMPLE?  
Runs afoul of typological data.

(23) We will discuss and dismiss some potential explanations, arguing for one analysis which best captures the results of the experiments.

(a) Markedness analysis  Doesn’t work  
(b) Ranking analysis  Doesn’t work  
(c) Unmarkedness analysis  Works!

(24) Markedness analysis:  
The results are actually just a function of the markedness of the initial segment of the onset. Notice the initial segment across the conditions:

(25) Initial Segments by Condition

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>CC</th>
<th>CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>z</td>
<td>ìlm</td>
<td>twîlm</td>
<td>skîlm</td>
</tr>
<tr>
<td>g</td>
<td>arθ</td>
<td>staθ</td>
<td>s.taθ</td>
</tr>
<tr>
<td>v</td>
<td>o.i̯k</td>
<td>f.loi̯k</td>
<td>s.toi̯k</td>
</tr>
<tr>
<td>v</td>
<td>i̯ʃ</td>
<td>k i̯ʃ</td>
<td>s.ki̯ʃ</td>
</tr>
<tr>
<td>dʒ</td>
<td>ðəntʃ</td>
<td>s.pəntʃ</td>
<td>s.təntʃ</td>
</tr>
<tr>
<td>b</td>
<td>asp</td>
<td>p lasp</td>
<td>s.tas̯p</td>
</tr>
<tr>
<td>j</td>
<td>arv</td>
<td>k larv</td>
<td>s.tarv</td>
</tr>
</tbody>
</table>

(26) However, there is not a correlation between the probability of the initial consonant and the well-formedness judgments of our subjects. \[r^2=-0.105; p=0.660\]

(27) Scatter-plot of Frequency by Response
(28) **Ranking analysis:**
   The effect can be captured through a stochastic or probabilistic Optimality Theory, in which the subject’s experience can manipulate rankings in their grammar.

(29) Boersma & Hayes (2001): the role played by various constraints is a function of the speaker’s experience with forms that violate the constraints (Gradual Learning Algorithm).

(30) Ranking Comparison

(a) 

(b) 

(31) Possible rankings:
   Regardless of how *COMPLEX is ranked simple onsets are never bad!

(32) Frequency of occurrence: The frequency of forms experienced in life is said to correlate with performance (judgments).
   (a) This holds true with phonotactic probability (Frisch 2000, Coleman & Pierrehumbert 1997).
   (b) This correlation doesn’t appear to hold at the level of onset complexity.

(33) Listeners have less real world experience with complex onsets (onset counts are from the Brown Corpus)
   words beginning with one segment onset: 574470
   words beginning with greater than one segment onset: 100298
(34) **Unmarkedness analysis:**
Judgments are a reflection of how *unambiguously* English a string is in terms of constraint rankings.

(35) **Grammar Comparison**

<table>
<thead>
<tr>
<th>Input</th>
<th>English</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>/voi.k/</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>/flo.i.k/</td>
<td>✔</td>
<td>*!</td>
</tr>
<tr>
<td>/sto.i.k/</td>
<td>✔</td>
<td>*!</td>
</tr>
</tbody>
</table>

Inputs with a single segment onset are well-formed according to both the English grammar and the ‘other’ grammar. Inputs with multiple segment onsets are well-formed only according to the English grammar.

(36) What is the ‘Other’ grammar?
(a) a particular ‘incorrect grammar’ stage we went through as a child (Levelt et al., 1999)
(b) a product of UG (see for example Itô & Mester 2001)
(c) an online re-ranking of all possible grammars

(37) Speakers are sensitive to patterns that occur in many languages and patterns that occur only in *their* language.

**E. Conclusion**

(38) **Results**
(a) There is an independent effect of onset length in determining the well-formedness of a nonsense word.
(b) A phonological constraint like *COMPLEX plays a role in experimentally elicited well-formedness judgments.
(c) The effect is simple vs. complex: not gradient
(d) The effect is backwards!
(e) At least in some cases, that effect is a function of ranking ambiguity rather than of ranking probability.
*Thanks to Ruby Basham, Laura Freebairn, Diane Ohala, Doug Pulleyblank, Natasha Warner, and Andy Wedel for experimental preparation and helpful discussion.

**F. References**


