Grammar is both categorical and gradient

Andries W. Coetzee

Abstract. In this paper, I discuss the results of word-likeness experiments conducted with speakers with Hebrew and English. The results of the experiments show that language users use their grammar in a categorical and a gradient manner. In word-likeness rating tasks, subjects make the categorical distinction between grammatical and ungrammatical – they assign all grammatical forms equally high ratings and all ungrammatical forms equally low ratings. However, in comparative word-likeness tasks, subjects are required to compare different grammatical or ungrammatical forms with each other. In these experiments, they make more fine-grained, gradient well-formedness distinctions between different grammatical or ungrammatical forms. This poses a challenge to standard derivational models of generative grammar. These models can easily account for the categorical distinction between grammatical and ungrammatical, but they have more difficulty with the gradient well-formedness distinctions. I show that the inherent comparative character of an OT grammar enables it to model both kinds of behaviors in a straightforward manner.

There is a growing body of literature showing that phonological grammar influences phonological performance. We know that grammar plays a role in phoneme identification (Massaro & Cohen, 1983; Moreton, 2002), the segmentation of speech into words (Kirk, 2001; Suomi et al., 1997), lexical decision (Berent et al., 2001ab), word-likeness ratings (Berent et al., 2001ab; Frisch & Zawaydeh, 2001), etc. Once we accept that performance reflects the influence of grammar, we can use performance data as a window on what grammar looks like. In this paper, I discuss performance data showing that grammar is both categorical and gradient. Grammar must be able to distinguish between grammatical (possible) words and ungrammatical (impossible) words. However, grammar must also be able to make gradient well-formedness distinctions within these two sets. In the set of grammatical forms, there are some forms that are “more” and some that are “less” grammatical. Similarly, there are “more” and “less” ungrammatical forms.

This challenges standard generative models that were designed to do only the former. I will show that Optimality Theory is ideally suited to model this kind of data. The reason is that OT is by its very nature comparative. The architecture of an OT grammar therefore already allows it to compare forms for their relative well-formedness.

This paper is structured as follows: Section §1 is a general discussion of the relationship between grammar and word-likeness judgments. Section §2 discusses the results of word-likeness experiments performed by Berent and colleagues on Hebrew (Berent & Shimron, 1997; Berent et al., 1991b). These experiments show that Hebrew speakers categorically distinguish between grammatical and ungrammatical forms in some task conditions, but that they also make gradient well-formedness distinctions in other task conditions. Section §3 discusses word-likeness experiments that I conducted with English listeners. These experiments confirm the results of Berent et al.’s Hebrew experiments. In section §4, I suggest a straightforward way in which we can account for these results within OT. Section §5 contains a conclusion and summary.
1. Grammar and word-likeness judgments

It is well known that language users have strong intuitions about what counts as a possible word of their language. Although [bl]k is not a word of English, it is perfectly well-formed according the phonotactics of English. *[kn]k, on the other hand, violates a phonotactic constraint – English does not tolerate [#kn-] word-initially. If these two forms were presented to English speakers in a word-likeness rating task, [bl]k will receive higher ratings than *[kn]k. This can be interpreted as evidence for the influence of grammar on word-likeness ratings.

However, there is a confound that sheds doubt on this interpretation. N nonce words that contain phoneme sequences that occur frequently in the lexicon are rated as more word-like than nonce words with less frequent phoneme sequences (Bailey & Hahn, 1998; Coleman & Pierrehumbert, 1997). Low ratings for nonce words with phonotactically illegal (and therefore non-occurring) sequences can then be interpreted as the logical extreme of such a frequency bias – a better rating for [bl]k might simply reflect the fact that [#bl-] has a higher frequency than [#kn-] in the English lexicon.

But there is also experimental evidence showing that grammar does contribute independently from lexical statistics to phonological processing. In a study of word-likeness ratings in Arabic, Frisch and Zawaydeh (2001) used non-words all of which contained unattested consonant sequences. Half of their stimuli contained consonant sequences that were absent from the lexicon because they violated a phonotactic constraint of Arabic (they contained contiguous homorganic consonants in violation of the Obligatory Contour Principle). The other half contained sequences that were accidentally absent from the lexicon (i.e. there are no phonotactic constraints against them). Both of these kinds of tokens contain non-occurring sequences and are therefore equal in terms of their lexical statistics. However, they found that tokens that contained OCP-violating sequences received lower word-likeness ratings than other tokens. Since both of these token types contain non-occurring sequences, this difference cannot originate in lexical statistics. They ascribe the difference to grammar.\(^1\)

Now that we have established that grammar plays a role in word-likeness judgments, we can look at word-likeness judgments for information on the structure of grammar – this is the topic of the next two sections. I discuss two kinds of word-likeness judgment tasks. Language users employ the information provided by grammar differently in the two tasks, and consequently treat the same (kind of) token differently.

In a “word-likeness rating” experiment, subjects are presented with one nonce word token at a time, and they have to assign each token a rating from some rating scale. In a “comparative word-likeness” experiment, subjects are presented with more than one token at a time, and they have to order the tokens according to their word-likeness.

In the experiments that I discuss below, we find evidence for the categorical nature of grammar in the word-likeness rating experiments. All nonce words that are well-formed according to the grammar received relatively high ratings. Consequently, these tokens are not distinguished form each other in terms of their assigned ratings. Similarly, all nonce

---

\(^1\) Moreton (2002) shows that principled and accidental gaps are processed differently.
words that are phonotactically ill-formed received very low ratings and were not distinguished from each other. Although the subjects had a rating scale with several discrete values, they treated the task as an “accept or reject” task, using basically only two values on the scale. Language users can therefore use the information provided by grammar to make a categorical distinction between grammatical and ungrammatical.

In the comparative word-likeness experiments, we find evidence for the gradient nature of grammar. In these experiments, subjects are sometimes required to compare two grammatical nonce words or two ungrammatical nonce words with each other, and to select the one that is more word-like. In the word-likeness rating task, they might have assigned two grammatical nonce words equally high ratings. But now that option is not available, and we find the following: Although two nonce words might both be grammatical, it is possible that the one contains a more marked structure and is therefore less well-formed. When forced to choose between two such forms, language users prefer the more well-formed token. The same happens when they are forced to compare two ungrammatical forms. Two forms might both be ungrammatical because they contain marked structures not tolerated in the language. However, one of the forms might contain a more marked structure. When forced to choose between two such nonce words, language users prefer the one that is less ill-formed. In addition to the categorical grammatical/ungrammatical distinction, language users can also make finer gradient distinctions in terms well-formedness.

In the next two sections, I discuss word-likeness experiments in Hebrew (§2) and English (§3) that illustrate these two uses of the information provided by grammar. In section §4, I then develop a version of OT that can account for the two response strategies.

2. Word-likeness ratings in Hebrew

One of the most striking features of Semitic morphology is the limitation on the distribution of identical consonants in verbal roots (Frisch et al., 2004; Gafos, 2003; Greenberg, 1950; McCarthy, 1986, 1994; Pierrehumbert, 1993, etc.). Forms with identical initial consonants are not allowed – *[QiQeS]* ill-formed. On the other hand, forms with identical final consonants are well-formed – *[QiSeS]* is acceptable. I will refer to *[QiQeS]*-forms as “initial-geminates”, to *[QiSeS]*-forms as “final-geminates”, and to forms with no identical consonants *[QiSeM]* as “non-geminates”.

Berent et al. (2001ab, Berent and Shimron, 1997) conducted a series of experiments in which they tested whether this restriction plays a role in how Hebrew speakers rate nonce words. In word-likeness rating tasks, they found that Hebrew speakers rated the two kinds of possible words, *[QiSeM]* and *[QiSeS]*, equally good and both better than the

---

2 This differs from the results of other word-likeness experiments, where subjects used the full scale (Bailey & Hahn, 1998; Coleman & Pierrehumbert, 1997; Pierrehumbert, et al., in press). In these experiments, the ratings reflected lexical statistics – higher ratings were assigned to tokens with higher statistics. In the experiments that I discuss below, the tokens were equated for their lexical statistics so that these statistics could not result in a spread of the ratings. I suspect that this is the reason for the difference in response patterns. However, this needs further research and I will not pursue it here. See also Frisch et al. (2001) for discussion of these two response strategies.
ungrammatical *[QiQeS]-forms. However, in comparative word-likeness tasks, their subjects differentiated between the two kinds of grammatical tokens – they preferred the non-geminates over the final-geminates. Although both of these are grammatical, the final geminates contain a marked structure (geminate consonants) absent from the non-geminates. When forced to choose between them, subjects go for the less marked token. I discuss the experiments of Berent and Shimron (1997) as a representative example of these experiments.

2.1 Word-likeness rating

Berent and Shimron (1997) selected 24 root trios. None of the roots corresponded to an existing Hebrew word. One of the members in each trio had three non-identical consonants (non-geminate). The other members both shared the first two consonants of the non-geminate. One of them doubled the first consonant forming an initial-geminate, and the other doubled the second consonant forming a final-geminate. Each of these trios was conjugated in three verbal forms. Their stimuli therefore included 72 non-geminate nonce words (24 non-geminate roots conjugated in 3 verbal forms), 72 final-geminate and 72 initial-geminate nonce words. The tokens were randomized, and presented in a written word-likeness rating task to 15 native speakers of Hebrew, who had to rate each token on a 5-point scale, with [1] corresponding to a form that is impossible as a word of Hebrew and [5] to a form that is an excellent candidate for a Hebrew word.

Berent and Shimron do not report the average scores assigned to each of the three token types. However, they do report the difference scores – i.e. the difference between the average ratings assigned to each of the three token types. The results of this experiment are summarized in the table in (1), and represented graphically in (2).

(1) Difference scores in word-likeness rating experiment (Berent & Shimron, 1997)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Example</th>
<th>Difference score</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial-geminates and non-geminates</td>
<td>*Q-Q-S</td>
<td>0.881</td>
<td>11.139</td>
<td>46</td>
<td>&lt; 0.000</td>
</tr>
<tr>
<td></td>
<td>Q-S-M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial-geminates and final-geminates</td>
<td>*Q-Q-S</td>
<td>0.801</td>
<td>9.984</td>
<td>46</td>
<td>&lt; 0.000</td>
</tr>
<tr>
<td></td>
<td>Q-S-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final-geminates and non-geminates</td>
<td>Q-S-S</td>
<td>0.081</td>
<td>5</td>
<td></td>
<td>&gt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Q-S-M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 I discuss only the most relevant aspects of the experimental design. Refer to Berent and Shimron (1997) for the full details.

4 These difference scores were computed as follows: (i) Difference Score (Initial–Final) = Mean Score (Final) – Mean Score (Initial). (ii) Difference Score (Initial–No) = Mean Score (No) – Mean Score (Initial). (iii) Difference Score (Final–No) = Mean Score (No) – Mean Score (Final).

5 Berent and Shimron do not report the t-statistic for this comparison. They do, however, report that a p-value of larger than 0.05 was obtained for this comparison using the Tukey HSD test.
2.2 Comparative word-likeness

In this experiment, Berent and Shimron used the same 24 root trios conjugated in the same three verbal patterns as in the word-likeness rating experiment. However, the root trios were kept together. They therefore had 72 non-word trios (24 trios conjugated in 3 verbal forms). Each trio has a non-geminate, a final-geminate and an initial-geminate. The order between the members of each trio was randomized, and the trios themselves were also randomized. These trios were presented in written form to 18 native speakers of Hebrew. Their task was to order each trio in terms of its word-likeness. A score of [3] was assigned to the most word-like member, and a score of [1] to the least word-like member. This setup differs from the word-likeness rating experiment by forcing subjects to choose between the two kinds of possible words.

As with the word-likeness rating experiment, Berent and Shimron report only difference scores. Their results are summarized in the table in (3) and represented graphically in (4).

(3) Mean ratings in the word-likeness ratings experiment (Berent & Shimron (1997))

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Example</th>
<th>Difference score</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial-gemination and no-gemination</td>
<td>*Q-Q-S</td>
<td>1.122</td>
<td>18.55</td>
<td>46</td>
<td>&lt; 0.000</td>
</tr>
<tr>
<td></td>
<td>Q-S-M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial-gemination and final-gemination</td>
<td>*Q-Q-S</td>
<td>0.682</td>
<td>11.28</td>
<td>46</td>
<td>&lt; 0.000</td>
</tr>
<tr>
<td></td>
<td>Q-S-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final-gemination and no-gemination</td>
<td>Q-S-S</td>
<td>0.44</td>
<td>_</td>
<td>_</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Q-S-M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6 Berent and Shimron do not report the t-score for this comparison. They do report that a p-value of smaller than 0.05 was obtained for this comparison using the Tukey HSD test.
As before, their subjects preferred the grammatical non-geminate and final-geminate forms over the ungrammatical initial-geminates. However, unlike in the word-likeness rating experiment, the subjects did distinguish between the two kinds of grammatical forms – they preferred the non-geminates over the final geminates.\footnote{Since we know that lexical statistics influences word-likeness judgments (see §1), we have to check whether the results of these experiments can be explained by lexical statistics. Berent and Shimron (1997) do not report on the lexical statistics of their tokens. However, Berent \textit{et al.} (2001a) conducted very similar experiments and found virtually the same results, and they do report the lexical statistics of their tokens. They extracted the 1412 productive verbal roots from a Hebrew dictionary and calculated statistics over these roots. For each token, they calculated the summed bigram frequency (SBF) as follows: SBF (C\textsubscript{1} C\textsubscript{2} C\textsubscript{3}) = F(C\textsubscript{1} C\textsubscript{2} X) + F(X C\textsubscript{2} C\textsubscript{3}) + F(C\textsubscript{1} X C\textsubscript{3}). They found that the average SBF of their final-geminate forms (11.58) was higher than that of their non-geminate forms (9.33). If the lexical statistics were dictating their subjects’ responses, final-geminates should have been preferred over non-geminates.}

This experiment shows that language users can also use the information provided by grammar to make gradient distinctions between different grammatical forms. Although both non-geminates and final-geminates are grammatical, final-geminates are somehow less well-formed than the non-geminates. I consider reasons for the well-formedness difference between these kinds of forms in §4.1.

### 3. Word-likeness ratings in English

English restricts the consonants that can co-occur in the onset and coda of a syllable (Fudge, 1969; Davis, 1984). I focus on only one aspect of this restriction here – words of the form [sCvC] are tolerated if both C’s are [t], but not if both are [k] or [p] – state is a word, but *skake and *spape are not even possible words (Browne, 1981; Davis, 1984, 1991; Fudge, 1969). We therefore have a situation very similar to the Hebrew example
above – we have a phonotactic constraint that can be used to divide a set of non-words between grammatical forms ([sTvT]) and ungrammatical forms (*[sKvK] and *[sPvP]). However, the situation is also different from the Hebrew example – in English we have two kinds of ungrammatical forms (*[sKvK] and *[sPvP]) rather than two kinds of grammatical forms as we had in Hebrew (non-geminates and final-geminates).

This makes for an interesting way in which to replicate some of the results of Berent *et al.*, and to extend upon their results. If English speakers react the same way as Berent *et al.*’s Hebrew speakers did, then we would expect them to rate all grammatical forms ([sTvT]) high, and all ungrammatical forms (*[sKvK] and *[sPvP]) low in a word-likeness rating experiment – i.e. we do not expect to see a difference between the two kinds of ungrammatical forms. However, if there is a well-formedness difference between *[sPvP]- and *[sKvK]-forms, we do expect to see evidence for this difference in a comparative word-likeness experiment. I performed a series of experiments with speakers of American English to test *inter alia* these predictions. For a full discussion of the experimental design and for other related results, see Coetzee (to appear, 2004). I discuss only the relevant aspects of the experimental design and results here.

### 3.1 Word-likeness rating

In this experiment, twenty subjects were presented with non-words of the form [sCvC] where the two C’s were two [t]’s, [k]’s or [p]’s. Tokens were presented auditorily in randomized order. The subjects had to rate the non-words for their word-likeness on a 5-point scale, where [5] corresponded to a token that sounded perfect as a possible word, and [1] to a form that did not sound like it could be a word of English. Tokens were selected in three conditions: (i) T~K: 5 non-words of the form [sTvT] and 5 of the form *[sKvK], (ii) T~P: 5 non-words of the form [sTvT] and 5 of the form [sPvP], (ii) K~P: 5 non-words of the form *[sKvK] and 5 of the form [sPvP].

This allows for a comparison of the grammatical forms ([sTvT]) with each of the two ungrammatical forms (*[sKvK] and *[sPvP]), and for a comparison of the two kinds of ungrammatical forms with each other. If the English subjects respond in the same way as the Hebrew subjects, then we would expect the grammatical [sTvT]-forms to be rated better than the ungrammatical forms, but we would expect no difference between the two kinds of ungrammatical forms. These expectations were confirmed by the results of the experiment. The results are summarized in the table in (5), and portrayed graphically in (6).

The difference between the grammatical forms and the two kinds of ungrammatical forms is significant by subjects and items. Like the Hebrew subjects, the English subjects rated grammatical forms better than ungrammatical forms. However, there is no significant different in the scores of the two kinds of ungrammatical forms. This extends on the Berent *et al.*’s results. In their experiment, we saw that subjects do not distinguish between different kinds of grammatical forms in a word-likeness rating task. The results of this English experiment show the same for ungrammatical forms.

---

8 In the token selection, I controlled for lexical neighborhood density and transitional probability (Luce & Pisoni, 1998). Within conditions, the tokens did not differ in terms of these statistics, so that lexical statistics should not cause a preference for any of the two kinds of tokens. Post-hoc regression analyses on the response data also did not reveal any significant influence of the lexical statistics. For the specific tokens and their lexical statistics, see Coetzee (2004: 453, 471).
Mean ratings in the three conditions in word-likeness rating experiment

<table>
<thead>
<tr>
<th></th>
<th>K~P</th>
<th>T~K</th>
<th>T~P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sKvK</td>
<td>2.52</td>
<td>3.64</td>
<td>3.65</td>
</tr>
<tr>
<td>sPvP</td>
<td>2.41</td>
<td>2.43</td>
<td>2.41</td>
</tr>
<tr>
<td>Mean</td>
<td>2.52</td>
<td>3.64</td>
<td>3.65</td>
</tr>
</tbody>
</table>

By subject

\( t(19) = 0.88, \ p = 0.19 \)  
\( t(19) = 5.81, \ p < 0.000 \)  
\( t(19) = 5.39, \ p < 0.000 \)

By item

\( t(8) = 1.29, \ p = 0.12 \)  
\( t(8) = 13.61, \ p < 0.000 \)  
\( t(8) = 13.40, \ p < 0.000 \)

Mean ratings in the three conditions in word-likeness rating experiment

3.2 Comparative word-likeness

The same twenty subjects participated in this experiment. They were presented with pairs of \([CVCC]\) non-words, and it was their task to select the member of each pair that they deemed most word-like. There were three kinds of pairs: (i) T~K: 15 pairs of the form \([STVT]~*[SKvK]\), (ii) T~P: 15 pairs of the form \([STVT]~*[SPvP]\), (ii) K~P: 15 pairs of the form \([SKvK]~*[SPvP]\). This design allows us to compare the grammatical \([STVT]\)-forms with each of the two kinds of ungrammatical forms (*[SKvK] and *[SPvP]). It also allows for a comparison between the two kinds of ungrammatical forms. Each token pair was presented twice, so that each subject responded 30 times in each condition.

Based on the results of Berent et al. and of the word-likeness rating experiment discussed in §3.1, we expect the subjects to prefer \([STVT]\)-forms over *[SKvK]- and *[SPvP]-forms. But what do we expect with regard the two kinds of ungrammatical forms? If there is a

---

\[9\] \( t \)-test is paired sample of two means, \( p \)-values are one-tailed.

\[10\] \( t \)-test is (non-paired) two sample of means, \( p \)-values are one-tailed.

\[11\] I controlled for lexical statistics in this experiment in the same way as in the word-likeness rating experiment. See footnote 8 above.
well-formedness difference between *[sKvK] and *[sPvP], then we expect to find evidence of this difference in the comparative word-likeness experiment, since subjects are forced to choose between these two kinds of tokens. However, it is not straightforward to determine whether there is a well-formedness difference between these two kinds of tokens. We do know that they are both ungrammatical. But are they equally ungrammatical? There is evidence that English restricts the co-occurrence of labials more severely than the co-occurrence of dorsals. (i) English tolerates words of the form [sKvG] but not of the form *[sPvB] – e.g. skag but *spab. (ii) English tolerates words of the form [sKvXK] where [X] stands for a nasal or a liquid, however, words of the forms *[sPvXP] are not tolerated – e.g. skulk, skunk, but *spulp, *spump. This leads us to suspect the following: Although both *[sKvK] and *[sPvP] are ungrammatical, *[sPvP] is less well-formed. We therefore expect *[sKvK] to be preferred over *[sPvP] in this experiment.

These predictions were borne out in the experiment. The table in (7) summarizes the results by condition, and the results are represented graphically in (8).

(7) Percentage that token type was selected in comparative word-likeness experiment

<table>
<thead>
<tr>
<th>K~P</th>
<th>T~K</th>
<th>T~P</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*[sKvK]</td>
<td>*[sPvP]</td>
<td>*sTvT</td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td><strong>By subject</strong></td>
<td>(t(19) = 1.12, \ p = 0.14)</td>
<td>(t(19) = 4.54, \ p &lt; 0.000)</td>
</tr>
<tr>
<td><strong>By item</strong></td>
<td>(t(14) = 1.92, \ p = 0.037)</td>
<td>(t(14) = 15.58, \ p &lt; 0.000)</td>
</tr>
</tbody>
</table>

(8) Percentage that token type was selected as more word-like in each of the three experimental conditions

---

12 *t*-tests are paired sample of two means, *p*-values are one-tailed.
These results show that the grammatical [sTvT]-forms were preferred over the two types of ungrammatical forms. This replicates the findings of Berent et al. for Hebrew and was also expected based on the results of the word-likeness rating experiment discussed above. However, the results of this experiment also extend on those of Berent et al. The English listeners preferred the *[sKvK]-forms over *[sPvP]-forms, although neither of these are possible words of English. This shows that language users can make finer distinctions within the set of ungrammatical forms in terms of well-formedness.

Taken together, the results of Berent et al. on Hebrew and the results on English discussed here, show the following: Language users can use the information provided by grammar in a categorical and in a gradient manner. In task conditions that do not require explicit comparison between forms, language users make only the categorical distinction between grammatical and ungrammatical. However, if task conditions require an explicit comparison between forms, then language users make finer gradient distinctions within the sets of grammatical and ungrammatical forms. We therefore need a theory of grammar that can do both of these things. In the next section, I show how an OT grammar can be used to do just this.

4. An Optimality Theoretic account of categorical and gradient behavior

In this section, I will show that an OT grammar is ideally suited to model both kinds of behavior observed in the experiments discussed above. The following is what happens in a word-likeness rating task: When presented with a non-word, the language user asks the question: Is there any possible input that my grammar will map onto this non-word token? An affirmative answer to this question means that the token is a possible word and therefore grammatical. A negative answer means that it is not a possible word and therefore ungrammatical.

The evaluation proceeds differently in a comparative word-likeness task. Now the language user is presented with more than one non-word. The first step is to determine the input for each non-word that would result in the most harmonic mapping onto the non-word. (Crucially, it is not necessary that the mapping from the assumed input onto the non-word be possible in the language. All that is necessary is that it be the input that would result in the most harmonic mapping onto the non-word.) The language user then compares the input–output-mappings for each of the non-word tokens.

4.1 Hebrew

In Hebrew, non-geminates ([Q-S-M]) and final-geminates ([Q-S-S]) are grammatical, while initial-geminates (*[Q-Q-S]) are ungrammatical. In this section, I first develop an OT account for this. I then show how this OT grammar can be used to explain the results of Berent et al.’s experiments discussed above.

---

13 The difference reached significance by items but not by subjects. I do not currently have an explanation for this difference between the subjects and items analysis.

14 I sketch only the outlines of an account here. See Coetzee (2004:348-380) for more detail.
I assume that final-geminates are derived from bi-consonantal roots (Gafos, 1999; McCarthy, 1986; Ussishkin, 1999; etc.). There are then three things we have to explain:

I follow McCarthy (1986) in assuming that there is an OCP constraint against identical, contiguous root consonants. The specific instantiation of the OCP is defined in (9).

(9) **OCP**: No contiguous, identical consonants in the surface realization of a root.\(^\text{15}\)

If we assume that morphological and phonological association are not necessarily isomorphic, the OCP can be used to explain why /Q^R^-S^R^-S^R/ → *[Q^R^-S^R^-S^R] is ungrammatical but /Q^R^-S^R/ → [Q^R^-S^R^-S^R] not.\(^\text{16}\) In *[Q^R^-S^R^-S^R], the sequence [S^R^-S^R] results in a violation of the OCP. However, since the second [S] in [Q^R^-S^R^-S^R] is not morphologically related to the root, it does not violate the OCP. The mapping /Q^R^-S^R/ → [Q^R^-S^R^-S^R] does violate **Uniformity** (because /S/ has two surface phonological correspondents). Hebrew therefore opts to violate **Uniformity** rather than the OCP, giving us evidence for the ranking OCP >> **Uniformity**. This explains why final-geminates are derived from bi-consonantal roots rather than tri-consonantal roots.

In accordance with “Richness of the Base” (Smolensky, 1996), we cannot exclude certain forms from the input. Our grammar must therefore be able to map even /Q^R^-S^R^-S^R/ onto a grammatical output. Since no such roots actually exist in Hebrew, we do not know what Hebrew would do with such inputs. We do know, however, that they will not be mapped faithfully onto the surface. I suggest the following as one possible solution: One of the identical root consonants is deleted, and the root is then treated like a bi-consonantal root – i.e. the last remaining root consonant doubles. Violation of the OCP is avoided by consonant deletion, giving us evidence for the ranking OCP >> **Max-C**.\(^\text{17}\)

McCarthy and Prince (1990) and Gafos (1999, 2003) argue that the verbal stem in Semitic languages must end on a consonant. I follow them in assuming the existence a constraint **Final-C**.\(^\text{17}\)

(10) **Final-C**: The verbal stem must end on a consonant.

This is the constraint that is responsible for forcing bi-consonantal roots to map onto tri-consonantal stems. To understand why, we have to consider the vowels that form part of the verb. There are many conjugational classes in Hebrew that are expressed by a bi-vocalic melody. Consider the so-called *piṭēl* as example. This conjugational class is characterized by the vocalic melody [i-e]. The stem of a Hebrew verb includes both the

---

15. Note that this constraint is defined with regard to the surface realization of the root. It therefore does not place a restriction on possible inputs. In the definition of this constraint, adjacency should be interpreted on the consonantal tier.

16. I use a superscripted R to represent morphological relation to the root, and a subscripted i to represent phonological relatedness.

17. “Stem” should be understood here as the structure to which inflectional affixes are attached.
root consonants and the vocalic melody – that is, inflectional affixes are attached to the unit comprised of the root plus the vocalic melody. Without augmenting a bi-consonantal root, the stem will end on a vowel – /Q\textsuperscript{R}-S\textsuperscript{R}, i.e./ then maps onto [[Q\textsuperscript{R}iS\textsuperscript{R}]].\textsuperscript{18, 19} \textsc{final-c} therefore forces the augmentation of bi-consonantal roots by addition of a consonant.\textsuperscript{20}

Finally, we need to account for the fact that it is always the final root consonant that spreads. Both /Q\textsuperscript{R}-S\textsuperscript{R}/ → [[Q\textsuperscript{R}-S\textsuperscript{R}-S\textsuperscript{i}] and /Q\textsuperscript{R}-S\textsuperscript{R}/ → *[[Q\textsuperscript{R}-Q\textsuperscript{R}-S\textsuperscript{R}]] satisfy \textsc{final-c}. Why then is [[Q\textsuperscript{R}-S\textsuperscript{R}-S\textsuperscript{i}]] grammatical but *[[Q\textsuperscript{R}-Q\textsuperscript{R}-S\textsuperscript{R}]] is ungrammatical? The difference between these structures is in the alignment of the root and the stem. In the grammatical [[Q\textsuperscript{R}-S\textsuperscript{R}-S\textsuperscript{i}]], the root and stem are perfectly aligned at their left edges. However, in the ungrammatical *[[Q\textsuperscript{R}-Q\textsuperscript{R}-S\textsuperscript{R}]], the root and stem are misaligned at their left edges. I argue that it is an alignment constraint that forces the final consonant to spread.

(11) \textbf{align-l}: The left edge of the root and the left edge of the stem must be aligned.

We now have all the constraints and crucial rankings that we need to account for the distribution of contiguous identical consonants in Hebrew verbs. The tableau in (12) shows how this grammar will deal with the different possible root inputs.

In a root with three non-identical consonants, the faithful candidate does not violate any constraints. Unsurprisingly, the faithful candidate (a) is therefore optimal. Now consider the bi-consonantal root input. Faithful (d) fatally violates \textsc{final-c}. Both (e) and (f) satisfy \textsc{final-c} by doubling one of the root consonants, earning them violations of \textsc{uniformity}. However, (f) spreads the initial root consonant so that the root and stem are misaligned in this candidate. This earns it a fatal violation of \textsc{align-l}.\textsuperscript{21} Finally, consider roots that contain contiguous identical consonants (either in initial or final position). The faithful candidates of both of these inputs, (g) and (l), fatally violate the OCP. This violation is avoided by deleting one of the identical consonants. The root is then treated just like a bi-consonantal root – in order to satisfy \textsc{final-c}, the remaining final root consonant doubles.

Now that we have a grammar for Hebrew, we can look at the results of Berent \textit{et al.} again. Consider first the results of their word-likeness rating experiment. In this experiment, Berent \textit{et al.} found that Hebrew speakers rated the two kinds of grammatical non-words ([Q-S-M] and [Q-S-S]) better than the ungrammatical non-words (*[Q-Q-S]), but they did not distinguish between the different grammatical non-words. We can

\begin{itemize}
  \item \textsuperscript{18} Stem boundaries are marked by vertical lines |.
  \item \textsuperscript{19} Other alternatives are ruled out by other high ranking constraints. [[i.Q\textsuperscript{R}eS\textsuperscript{R}]] and [[Q\textsuperscript{R}i.eS\textsuperscript{R}]] are both ruled out by \textsc{onset}. [[Q\textsuperscript{R}eS\textsuperscript{R}']] and [[Q\textsuperscript{R}iS\textsuperscript{R}']] are ruled out by a constraint requiring faithful parsing of the vocalic melody – some version of \textsc{max-v} (Gafos, 2003; Ussishkin, 1999).
  \item \textsuperscript{20} This constraint can be satisfied by epenthesis of an additional consonant (violating \textsc{dep-c}), or by the spreading of a stem consonant (violating \textsc{uniformity}). Since Hebrew opts for the latter, it means that we have the ranking \textsc{dep-c} >> \textsc{uniformity}. Of course, the fact that \textsc{final-c} is satisfied at the expense of \textsc{uniformity}, also gives evidence for the ranking \textsc{final-c} >> \textsc{uniformity}.
  \item \textsuperscript{21} There are two candidates that satisfy \textsc{final-c} and \textsc{align-l}, namely candidate that spreads the initial root consonant to the right and a candidate that spreads the final consonant to the left, i.e. [Q\textsuperscript{R}-Q-S\textsuperscript{R}] and [Q\textsuperscript{R}-S-S\textsuperscript{R}]. These candidates are ruled out by a high ranking constraint requiring the surface correspondents of the root to be contiguous, i.e. \textsc{contiguity} indexed to the root.
\end{itemize}
explain this as follows: When the subjects are presented with a single non-word at a time to rate, they determine whether there is at least one input that their grammar can map grammatically onto that non-word. This is equivalent to asking whether the non-word is a possible word or not. If there is an input that would map grammatically onto the non-word, the subjects assign it a high score. On the other hand, if there is no such input, subjects assign the non-word a low score.

(12) Basic verbal grammar of Hebrew

Now that we have a grammar for Hebrew, we can look at the results of Berent et al. again. Consider first the results of their word-likeness rating experiment. In this experiment, Berent et al. found that Hebrew speakers rated the two kinds of grammatical non-words ([Q-S-M] and [Q-S-S]) better than the ungrammatical non-words (*[Q-Q-S]), but they did not distinguish between the different grammatical non-words. We can explain this as follows: When the subjects are presented with a single non-word at a time to rate, they determine whether there is at least one input that their grammar can map grammatically onto that non-word. This is equivalent to asking whether the non-word is a possible word or not. If there is an input that would map grammatically onto the non-word, the subjects assign it a high score. On the other hand, if there is no such input, subjects assign the non-word a low score.

As the tableau in (12) shows, there is some input that will map onto a non-geminate – namely an input identical to the output. /Q-S-M/ → [Q-S-M] is grammatical, and therefore non-geminate non-words are identified as possible words and assigned high scores. Tableau (12) also shows that there are inputs that will be mapped grammatically

<table>
<thead>
<tr>
<th>Root structure</th>
<th>OCP</th>
<th>AL-L</th>
<th>FIN-C</th>
<th>UNIFORM</th>
<th>MAX-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>No identical:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Q^R-S^R-M^R/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. L Q^R-S^R-M^R</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Q^R-S^R-S_j</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Q_r Q^R-S^R</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bi-consonantal:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Q^R-S^R/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. L Q^R-S^R-S_j</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Q_r Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final identical:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Q^R-S^R-S^R/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Q^R-S^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. L Q^R-S^R-S_j</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k. Q_r Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial identical:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Q^R-Q^R-S^R/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l. Q^R-Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m. Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n. L Q^R-S^R-S_j</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o. Q_r Q^R-S^R</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
onto final-geminates. In fact, there are three such inputs, namely /Q-S/, /Q-S-S/ and /Q-Q-S/. Also final-geminates are then identified as possible words and assigned high ratings. The situation with initial-geminates is different. None of the logically possible inputs will be mapped grammatically onto a form such as *[Q-Q-S]. This is because both /Q-Q-S/ and /Q-S/ are mapped onto final-geminates. Initial-geminates are then identified as impossible words, and assigned low scores.

Now consider the comparative word-likeness experiment. The experimental setup is such that subjects cannot treat all possible words equally – they are forced to choose between them. Berent et al. found that their subjects imposed the following word-likeness hierarchy on the non-word tokens: non-geminates > final-geminates > initial-geminates. I propose that language users do this as follows: For each of the non-words, they find the input that would most harmonically map onto the non-word. We can identify these inputs in the tableau in (12). For the non-geminate [Q-S-M], this input is /Q-S-M/. For the final-geminate [Q-S-S] there are four inputs to consider, /Q-S-M/, /Q-S/, /Q-S-S/ and /Q-Q-S/. Of these three /Q-S/ results in the most harmonic mapping. /Q-S/ → [Q-S-S] violates only UNIFORMITY, while all other inputs violate both UNIFORMITY and MAX-C. For the initial-geminate *[Q-Q-S], we again have to consider all four inputs. Now the input /Q-Q-S/ is the best option22 – the faithful candidate for this input, (12)(I), violates only the OCP. All other *[Q-Q-S]-candidates violate ALIGN-L, which like the OCP is undominated, plus some lower ranked constraints. Once the input that results in the most harmonic mapping for each non-word has been determined, the language user compares the mappings in a tableau. This comparison is shown in (13).

(13) Comparing the different non-words in Hebrew

<table>
<thead>
<tr>
<th>No</th>
<th>Final</th>
<th>Initial</th>
</tr>
</thead>
</table>

The tableau in (13) is not an ordinary OT tableau – it does not compare different output candidates for the same input, but rather different input–output-mappings.23 Rather than the usual pointing hand to indicate the winning candidate, I use Arabic numerals to indicate the order that the grammar imposes on these forms. The non-geminate mapping

22 The mapping does not have to be grammatical. It simply needs to be the best possible mapping.

23 See Berent and Shimron (1997) for a similar proposal, and Coetzee (2004) for an explicit formalization of an OT model that can do these kinds of comparisons.

Keller and Sorace (2005:1516) claim that OT cannot compare candidates that are derived from different inputs. It is true that this kind of comparison is not usually done in production oriented OT. However, it is untrue that OT cannot do this. Nothing in the way that EVAL works depends on the origin of the candidates being compared – i.e. it can compare forms that do not share the same input. Even in classic OT, there is acknowledgement that EVAL can do this. In “lexicon optimization”, EVAL compares different possible inputs for a single output (Prince and Smolensky, 1993). See Coetzee (2004) for a detailed discussion of this characteristic of EVAL.
violates none of the constraints, while the final-geminate mapping violates UNIFORMITY. When these two mappings are compared, the non-geminate mapping is therefore the more well-formed option. Although both of these represent mappings that are possible in Hebrew, they are not equally well-formed. The non-geminate form is perfectly unmarked and perfectly faithful, while the final-geminate is not perfectly faithful.24 This corresponds to the way in which Berent et al.’s subjects responded. When required to compare non-geminates and final-geminates, they preferred the non-geminate over the final geminates.

The initial-geminate form violates the OCP. Because of the ranking OCP >> UNIFORMITY, the final-geminates are better than the initial-geminates, and this is again how they were rated by the subjects in the experiment.

A note is on order here about how to interpret the tableau in (13). This tableau shows that non-geminates are more well-formed than final-geminates which are again more well-formed than initial-geminates. I claim that the subjects used this information as follows: When asked to choose between a non-geminate and a final-geminate, the subjects are more likely to select the non-geminate. Because there are many factors in addition to grammar that influences how subjects respond, we cannot expect that they always act according the information provided by grammar. However, we can expect that the information provided by grammar will bias their responses.

Since OT already has the ability to compare forms for their relative well-formedness, it is a small change that has to be made to the theory to explain what the subjects do in a comparative word-likeness experiment. Rather than comparing output candidates for the same input, they compare output candidates that do not share the same input.

4.2 English

English allows words of the form [sTvT], but not of the form *[sKvK] or *[sPvP]. In this section, I first develop an OT account to explain this, and then show how this account explains the results of the word-likeness experiments that I discussed earlier.

I assume a markedness constraint against each of the three kinds of tokens considered here, i.e. *sTvT, *sKvK, and *sPvP. These constraints can be viewed as the local conjunction of the OCP-type constraints against multiple occurrences [t], [k] or [p] in a single syllable, with a constraint against the sequence [s+stop] (Coetzee, 2004:402-420; Coetzee & Pater, 2005).25

---

24 The final-geminate actually violates some markedness constraints that I have not considered here. For instance, in this form the right edges of the root and stem are misaligned. This form therefore violates the R-counterpart to ALIGN-L defined in (11) above. It also has two contiguous identical consonants in the stem even if not in the root. In addition to the root-based OCP defined in (9), there is also a less strict (ranked lower) stem-based OCP (Coetzee, 2004:363-365). The final-geminates violate this stem-based OCP. The fact that the non-geminates were preferred over the final-geminates can also be explained by considering only markedness constraints. This might be preferable since we are dealing with non-words here. Maybe language users do not assume a specific underlying representation for non-words but consider only different surface realizations.

25 I give only a very basic motivation for these constraints here. See Coetzee (2004) for a complete discussion and motivation. See Baertsch and Davis (2003) for a very different approach.
It is well established that languages often place restrictions on the occurrence of multiple identical consonants in certain local domains. The Hebrew data discussed just above is an example of this. Similar restrictions also hold of languages as diverse as Arabic (Frisch et al., 2004; McCarthy, 1994), Japanese (Kawahara et al. to appear), Russian (Padget, 1995), Muna (Coetzee & Pater, 2005), French, Latin and English (Berkley, 2000). In all of these languages, there is evidence that identical (or highly similar) consonants are avoided in some local context. The languages differ in whether these restrictions are gradient or absolute, whether they apply to identical or highly similar consonants, and in what the domain is over which the restrictions apply. In spite of this diversity, it is clear that there must be a general family of constraints that militates against multiple occurrences of identical (or similar) consonants. I propose that the versions of these constraints stated in (14) are relevant to the restriction in English that is our focus here.26

26

(14) *[t…t]σ: Do not allow a syllable with two [t]'s.
   *[k…k]σ: Do not allow a syllable with two [k]'s.
   *[p…p]σ: Do not allow a syllable with two [p]'s.

There is an extensive literature on the markedness of \([s+stop]\) sequences (Broselow, 1991; Davis, 1984; Hayes, 1985:140-149; Kahn, 1980; Lamontagne, 1993: Chapter 6; Selkirk, 1982; etc.). There are two views on why this sequence is marked. One view considers these structures to be true consonant clusters. If this is the case, then they violate the Sonority Sequencing Principle – sonority falls from \([s]\) to the \([stop]\). Under another view, these structures are complex segments rather than consonant clusters. This avoids the violation of the SSP. However, if they are complex segments then they are marked per se. Either way, an \([s+stop]\)-sequence is marked. In (15), I formulate a constraint against such sequences.

(15) *[s+stop]σ: [s] is not allowed to precede a stop in one syllable.

English violates each of the constraints in (14) and (15) individually – this is evidenced by the existence of words such as sty (*\([s+stop]\)σ), toot (*\([t…t]\)σ), cake (*\([k…k]\)σ), and pop (*\([p…p]\)σ). This means that all of these constraints rank very low in English – so low that they have no discernable effect. What English does not tolerate, is the violation of certain combinations of these constraints within a single syllable – specifically, violation of \(*[s+stop]_\sigma\) and any of \(*[k…k]_\sigma\) or \(*[p…p]_\sigma\). OT accounts for the avoidance of the local accumulation of marked structure through local conjunction of markedness constraints (Smolensky, 1995). We can now show that the \(*sCvC\) constraints are just the local conjunction of each of the constraints in (14) with the constraint in (15).

(16) *sTvT: Do not violate \(*[s+stop]_\sigma\) and \(*[t…t]_\sigma\) in the same syllable.
   *sKvK: Do not violate \(*[s+stop]_\sigma\) and \(*[k…k]_\sigma\) in the same syllable.
   *sPvP: Do not violate \(*[s+stop]_\sigma\) and \(*[p…p]_\sigma\) in the same syllable.

26 These constraints are too specific. The domain should probably be defined more broadly – as English also does not allow words of the form \(*[sPv.Pv]\) or \(*[sKv.Kv]\). These constraints should also apply to more than just the voiceless stops – since English also does not tolerate \(*[sLvL]\), \(*[sNvN]\), etc. For more on these kinds of restrictions, see Fudge (1969). For discussion about how to state the domain of these constraints, see Coetzee (2004:420).
We know that English tolerates violation of *sTvT but not of *sKvK or *sPvP. This gives us evidence for the ranking \{*sPvP, *sKvK\} >> Faithfulness >> *sTvT. But it does not tell us how to rank *sKvK and *sPvP relative to each other. However, as I have shown in §3.2, English in general tolerates co-occurrence of dorsals better than co-occurrence of labials. Based on this, we can hypothesize that *sPvP outranks *sKvK.

We now have the following mini-grammar for English: *sPvP >> *sKvK >> Faithfulness >> *sTvT.\(^{27}\) The tableau in (17) shows that this grammar does correctly predict that [sTvT] is a possible word of English while neither *[sKvK] nor *[sPvP] is.

(17) Mini-grammar of English [sCvC]-forms \(^{28}\)

<table>
<thead>
<tr>
<th></th>
<th>*sPvP</th>
<th>*sKvK</th>
<th>Faithfulness</th>
<th>*sTvT</th>
</tr>
</thead>
<tbody>
<tr>
<td>/sTvT/</td>
<td>L</td>
<td>sTvT</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>sTvK</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/sKvK/</td>
<td>sKvK</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>sKvT</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/sPvP/</td>
<td>sPvP</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>sPvT</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This tableau shows that an /sTvT/-input will map faithfully onto itself, and therefore [sTvT] is correctly predicted as a possible word. However, neither /sKvK/ nor /sPvP/ maps faithfully onto themselves. *[sKvK] and *[sPvP] are therefore correctly predicted not to be possible words. We now have all the information we need to explain the results of the English experiments discussed above in §3.2.

Consider first the word-likeness rating experiment. In this experiment, subjects had to rate non-words individually for their word-likeness – i.e. no direct comparison between non-words was required. The results show that the subjects made a categorical distinction between possible and the impossible words – the grammatical [sTvT] non-words received high scores, and the ungrammatical *[sKvK] and *[sPvP] forms received low scores but were not distinguished from each other. The subjects used their grammar in the same manner as the subjects in Berent et al.’s experiment. For each non-word, they determine whether there is an input that will map unto the specific non-word. If such an input exists, the non-word is identified as a possible word, and it receives a good rating. Non-words for which no such input exists are identified as impossible words, and receive low scores.

Now consider the comparative word-likeness experiment. In this experiment, the subjects were required to compare two non-words and to select the one that they deemed most word-like. The results of the experiment show that the subjects rated the non-words as follows: [sTvT] > *[sKvK] > *[sPvP]. They no longer divide the set of non-words into

---

\(^{27}\) Of course, each of the individual constraints from (14) and (15) also rank below Faithfulness.

\(^{28}\) I use the faithfulness constraint \textsc{ident}[\text{place}] here. However, since there are no active alternations in English involving *sPvP and *sKvK we cannot know for sure what the constraint should be. It can be any faithfulness constraint that can be violated to avoid violation of *sKvK and *sPvP.
the two categorical classes of grammatical and ungrammatical. Rather, the non-words are now rated gradiently according to their relative well-formedness. The subjects again responded like the subjects in Berent et al.’s experiment. For each non-word, they first determine the input that would result in the most harmonic mapping onto the non-word. For an [sTvT] non-word, this input is obviously /sTvT/. Any other input that maps onto [sTvT] will violate some faithfulness constraint in addition to *sTvT. Similarly, the relevant input for *[sKvK] will be /sKvK/ – again because any other input will violate a faithfulness constraint in addition to *sKvK. With similar reasoning, we can also show that the relevant input *[sPvP] will be /sPvP/. Once these inputs have been determined, the subjects compare the three input-output-mappings in an OT tableau. This comparison is shown in (18).

(18) Comparing non-words in English

<table>
<thead>
<tr>
<th></th>
<th>*sPvP</th>
<th>*sKvK</th>
<th>Faithfulness</th>
<th>*sTvT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>/sTvT/ → [sTvT]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>/sKvK/ → [sKvK]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>/sPvP/ → [sPvP]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As before, this tableau should be interpreted differently from a standard OT tableau. This is not a production oriented tableau in which different output candidates for the same input is compared. We are not only interested in the best candidate, but rather in how all three candidates are related to each other. Since /sTvT/ → [sTvT] violates the lowest ranking constraints, it is rated best, as indicated by the numeral [1] next to this mapping. Since /sPvP/ → [sPvP] violates the highest ranking constraint, it is rated worst of all, indicated by the [3] next to this mapping. This corresponds to how the subjects rated these tokens in the comparative word-likeness experiment.

5. Conclusion

The view of grammar in standard generative theory can be characterized as follows: For every form encountered by the grammar, the grammar determines whether there is a grammatical derivation that will result in that specific form (i.e. can the form be parsed grammatically). If there is such a derivation, then the form is classified as grammatical, else it is classified as ungrammatical. Grammar is then a function that makes the categorical distinction between grammatical and ungrammatical.

However, I have discussed data from two sets of word-likeness experiments above showing that humans use their grammar in more varied ways. They do have strong intuitions about whether some non-word is a possible word or not (i.e. whether a form is grammatical or not). In task conditions that do not require explicit comparison between forms the subjects in word-likeness experiments correspond according to these intuitions. This shows that the information provided by grammar can be interpreted in terms of the categorical distinction between grammatical and ungrammatical. But in addition to the distinction between possible and impossible, humans can also make more fine grained gradient distinctions between more (im)possible and less (im)possible. They can compare
two grammatical forms, and decide which of the two is more word-like. Similarly, they can compare two ungrammatical forms and decide which is more word-like. In task conditions that explicitly require subjects to make these kinds of comparisons, they respond along these lines.

If language users can use the information provided by grammar to respond both in categorical and in gradient manners, then grammar should be able to provide both categorical and gradient information. This poses a challenge to standard generative grammar which was not designed as a comparative theory of grammar. However, because of its inherent comparative nature, OT is perfectly suited for this task.

We can make the categorical distinction between grammatical and ungrammatical non-words as follows in an OT grammar: For a non-word, if there is an input that would be mapped grammatically onto the non-word, then the non-word is grammatical (is a possible word of the language). On the other hand, if there is no input that would map onto the non-word, then it is ungrammatical (not a possible word of the language). This is equivalent to what standard generative grammar does – asking whether there is a well-formed derivational history for the non-word. In OT, this is easily done using OT tableaux in the standard manner.

Since an OT grammar is by design comparative, it is also a straightforward matter to get information about the finer gradient well-formedness relationship between non-words. This can be done as follows: For every non-word, determine that input that will map most harmonically onto the specific non-word. Then compare the input–output-mapping for the different non-words. In this kind of tableau, EVAL does not compare different output candidates for a single input, but rather input–output-mappings that differ in terms of both input and output.

Because of its comparative design, an OT grammar can be used to account for processing data of the kind discussed above. The fact that an OT grammar can do this even though it was not designed with this kind of application is strong evidence that the basic architectural design of the theory is on the right track.

Having established that processing data reflect the influence of grammar, and having shown that we can account for this influence in OT, we can now use processing data a rich and largely untapped source of information about grammatical competence. This opens up many new and interesting research possibilities.

References


