Phonotactically well-formed onset clusters as processing units in word recognition

Tom Lentz

Utrecht Institute of Linguistics OTS, University Utrecht, Janskerkhof 13a, Utrecht, Netherlands
Tom.Lentz@let.uu.nl

Abstract
Phonotactic well-formedness has an effect on speech processing. This is likely due to an independent sub-lexical representation of phonotactics. Researching that knowledge requires isolating it from indirect effects. A prominent indirect effect comes via lexical neighbourhood. The better phonotactically a word is, the more neighbours it has, the harder it is to recognise it.

The present study examined the sublexical effect for phonotactically good word onsets with auditory priming. Word recognition was facilitated for good clusters, in spite of the larger number of lexical competitors. Word recognition latency is corrected for the effect of lexical neighbourhood, additional effects have their origin in the processing differences of the auditory primes. We found that words with good phonotactic onsets are recognised quicker, but that (destructive) manipulation of the prime onset destroys the benefit of good onsets, and can even revert it.

Index Terms: speech perception, phonotactics, language specificity, phonology, word recognition, auditory priming, speech processing.

1. Introduction

1.1. Separating phonotactic and lexical neighbourhood effects
Recognition of spoken words is shown to be affected by components of these words. There is not only a role for phonemes, but also for larger sublexical chunks, like consonant clusters. Phonotactic well-formedness has a strong influence on word recognition. Vitevitch and Luce [1, 2] (henceforth V&L) show that this is an effect that can be modulated by different tasks, in as far as it is obscured by a stronger and opposite effect of lexical neighbourhood density. As good phonotactically formed clusters are in denser lexical neighbourhoods, the direct benefits of easier processing of good phonotactic combinations are cancelled out by the indirect disadvantage of a larger set of lexical competitors.

V&L modulated the phonotactic effect by using different tasks. The strongest facilitating phonotactic influence was found in a mixed word/non-word same/difference task. The authors explain this with the assumption that the non-words force the participants to process the sub-lexical chunks and not focus on the lexical neighbourhood. The lexical effect was still present for words, though. In a lexical decision task, this effect wins for both words and non-words: altogether, good phonotactic properties slowed down processing.

The V&L experiments show the existence of independent, sub-lexical phonotactic effects, and were not intended to look into details of this effect. The measure of phonotactic well-formedness used is phoneme and biphone frequency, taken over the whole word. The present study focusses on phonotactic effects, and is therefore controlled for lexical effects. This allows to address the details of the type of phonotactic well-formedness that facilitates processing on a sub-lexical level.

In order to do so, two types of phonotactic combinations are contrasted. One type is dubbed ‘good’, and is hypothesised to have a special status in the phonotactic grammar. We do not wish to claim that this grammar is of a special cognitive type on the basis of behavioural data only. However, processing good clusters as units might be beneficial in speech recognition, as they occur frequently; the experiments show that human listeners indeed process frequent combinations differently.

1.2. Word initial clusters of /s/ and consonant
This study investigated Dutch words starting with s and one other consonant after the s, /sC/-onsets. The phonotactic well-formedness is operationalised as the O/E ratio of the s and the consonant as found in continuous speech, in this case the CGN Corpus of Spoken Dutch [3]. O/E ratios are calculated by dividing the observed frequency O by the estimated frequency E. E can be calculated by multiplying the probability of both components and the total number of tokens. This number reflects how often two phonemes would be observed next to each other if there was no preference for combinations at all.

The actual observed frequencies of the /sC/-clusters can deviate from the expected value. In the case the O/E ratio is below 1, there seems to be a preference against combining the phonemes. If it is larger, the combination can be considered preferred (for brevity’s sake, the term ‘good’ will be used in the rest of the paper). Values around 1 seem to indicate that a language has no preference for or against the combination.

O/E ratios are given in table 1 for all non-marginal Dutch /sC/-clusters that occur word-initially. The clusters come in two groups: O/E above 2, or around 1. Note that statistical measures such as transitional probability are unidirectional.

The good clusters’ overrepresentation might be explained in part by their phonetic properties, but the point here is that they have a special status, which is a phonotactic property of Dutch. It is language specific: many languages do not allow consonant clusters at all and have zero O (and therefore O/E) values for all combinations. The neutral clusters, however, have observed values that can be explained by the frequency of their parts (phoneme frequencies), i.e. the expected value.

This study proposes to look into the extra knowledge that the good clusters are indeed good. This knowledge does not follow from the parts and is therefore connected to the cluster as a unit. Below the possible effects on processing are discussed, after which they are tested.
signal to a phonetic perception, but to the mapping of the pho-

To assim to what extent a manipulated version of a word can

2. Experiment

Table 1: sC-clusters, ranked on O/E as found in the phonetically

annotated part of the corpus of spoken Dutch.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Observed</th>
<th>Expected</th>
<th>O/E</th>
</tr>
</thead>
<tbody>
<tr>
<td>[sp]</td>
<td>5573</td>
<td>1783</td>
<td>3.12</td>
</tr>
<tr>
<td>[sx]</td>
<td>10762</td>
<td>3596</td>
<td>2.99</td>
</tr>
<tr>
<td>[st]</td>
<td>32173</td>
<td>15840</td>
<td>2.03</td>
</tr>
<tr>
<td>good</td>
<td>48508</td>
<td>21219</td>
<td>2.28</td>
</tr>
<tr>
<td>[sn]</td>
<td>3548</td>
<td>3541</td>
<td>1.00</td>
</tr>
<tr>
<td>[sl]</td>
<td>3667</td>
<td>3304</td>
<td>1.11</td>
</tr>
<tr>
<td>[sn]</td>
<td>3742</td>
<td>3136</td>
<td>1.19</td>
</tr>
<tr>
<td>[sk]</td>
<td>2703</td>
<td>3241</td>
<td>0.83</td>
</tr>
<tr>
<td>neutral</td>
<td>13660</td>
<td>13222</td>
<td>1.03</td>
</tr>
</tbody>
</table>

13. Segmentation and epenthesis effects

The good/neutral dichotomy found for /sC/-clusters is possibly relevant to two speech processing issues. First, it could affect segmentation of words. We choose to calculate the O/E ratios over continuous speech to make sure the property is in principle available to infant learners before they construct a lexicon or have access to segmentation, as infants seem to have knowledge of phonotactics early, before they acquire a lexicon [4, 5]. In that sense sub-lexical phonotactics could be the origin of lexical phonotactic effects.

Illegal phonotactics impede certain segmentations and fa-

cilitate others, as McQueen showed [6]. To test for that effect of good clusters, clusters were epenthesised in two ways: to [r+sC] and to [suC]. The effect of the manipulations could differ; if good clusters are recognised faster because the segmentation keeps them together, they should suffer less from the [r+sC] ma-

nipulation than neutral ones.

However, McQueen’s findings cannot be directly extrapo-

lated to the present materials, as the benefit of keeping the good clusters together is in conflict with the bad phonotactic well-

formedness of the residue [r], a vowel that is not a possible syllable in Dutch on its own. Under a null hypothesis that both clusters are nothing more than the sum of their parts, the [suC] manipulation cannot have a different effect, as both phonemes are still there. However, if clusters are processed as units on some level, the insertion of [u] would cause more problems.

The second important aspect of phonotactic status of clusters on processing is perceptual epenthesis; illegal phonotactic structures can be corrected to legal ones in perception [7, 8]. In Dutch the clusters are legal and therefore not necessarily broken by epenthesis (sC → sVC). However, they might be preceded by vowels in continuous speech, which matches the epenthesis that best fits fricatives, as Fleischhacker found [9]. That context might therefore be less damaging to the word recognition process.

Both [r+sC] and [suC] manipulations are phonotactic im-

provements (clusters are marked in general); both are attested cross-linguistically: they occur in Spanish and Japanese for loanwords, respectively. but they were expected not to match the representations of the words and therefore harm speech pro-

cessing / word recognition. This is important, as it means the phenomenon does not tap into easier mapping of the acoustic signal to a phonetic perception, but to the mapping of the pho-

netic signal to a phonological form (that is under the influence of phonotactics).
The model incorporated a faster response for good cluster targets, but not significantly so (\( p = 0.147, \text{n.s.} \)). Interaction was present between the type of manipulation and the prime type, because /su/C/ manipulation was significantly worse (\( p < 0.001^{**} \)). Looking at the data, we see that the effect of cluster type is attenuated by an interaction with prime type and group. That interaction is not significant in itself. The /rsC/-manipulated primes for good cluster words could (\( p > 0.17, \text{n.s.} \)) show some cancellation of the priming effect (reaction times going up), as can be seen by visual inspection of the data (figure 1 and ??).

As an effect of good phonotactics was expected (following V&L), a model with simple main effects was constructed including the interaction between cluster type and type of manipulation. Now \( p < 0.05^{*} \), showing that a simple main effect for condition is lost in a crossed interaction, even though the interaction is not significant.

However, we feel this model to be wrong; although this model is better than an empty model (\( p < 0.0360 \)), the full model shows that the phonotactic effects are explained better by the other factors (the full model is better than the condition model, \( p < 2.2^{16} \)).

Bonferroni correction for the number of contrasts (11) indicates that \( t \)-values over 3.1 are acceptable. An insightful model includes the interaction with prime type, excluding its main effect. This allows us to look at the interaction of condition at the different prime types. All four interactions (with as a baseline the faithful primes corrected for the possible condition main effect) now have a \( t \)-value well above the threshold: (good cluster + manipulated prime is 0.12 \( \log \text{ms} \) slower, \( t = 8.3 \), neutral cluster + manipulated prime 0.10 \( \log \text{ms} \) slower, \( t = 7.1 \), good cluster and unrelated prime are 0.21 \( \log \text{ms} \) slower, \( t = 14.37 \), and neutral clusters and unrelated primes are 0.23 \( \log \text{ms} \) slower, \( t = 16.3 \)). Together, there is a disordinal interaction: the manipulations make the priming for good clusters less effective, while it has better priming in words with unmanipulated primes.

However, this effect can only be attributed to the prime type part, given the full model. On the one hand, as usual, a \( p \) value above 0.1 cannot be used to suggest a trend, but on the other hand, it cannot be used to discard the idea that the cluster well-formedness has an effect. The solution for the disappearing phonotactic effect can be found when looking at the prime type / condition interaction for each manipulation separately. A condition main effect in the /rsC/ group \( (t = 3.2, \text{unrounded value above unrounded threshold}) \) was found. We also find a trend of an interaction: good clusters suffer more from manipulation (i.e., loose their benefit, see fig. 1). However, in the /suC/ group, no such effects can be found (fig. refchart2). As in that group the effect of manipulated primes is much stronger, we believe this underlies the "loss" of the phonotactic effect in a floor effect (hardly any priming is left anyway).

The model contains a very strong interaction for manipulated primes of the /suC/ type \( (t = 3.6) \); these primed much less than expected, close to cancelling the whole priming effect.

### 3. Discussion

The results indicate that the data follow existing predictions: phonotactics help speech processing in general, if the clusters are presented as clusters. As we corrected for word recognition times with a baseline, we isolated them from lexical

\(^{1}\)However, the crossed random effect issues that Quené and van den Bergh address did not turn out to be too severe in these data.

\(^{2}\)Not corrected for multiple comparisons.
The experimental design and the model analysis correct for the manipulation, but the V&L findings that phonotactic well-formedness play a role cannot be supported by these data for the present phonotactic contrast.

The interesting finding is that splitting a cluster with a vowel is significantly worse than a prothesis type of epenthesis [rs]. Visual inspection suggest that the good clusters suffer and even more so for being broken, but the data only suggests tendencies in these directions. The simple main effects found for the interaction were included for purposes of comparison with earlier findings, mainly V&L. More research needs to be done to independently show the effect.

4. Conclusions

The phonotactic well-formedness of the stimuli seem to have an effect on their recognition, as it shows up in priming. Destroying clusters reduces priming effects, showing that the clusters are represented as such. The different effects for the two manipulations show that just phoneme similarity is not a sufficient predictor for priming and by inference, for word recognition: both manipulations added just one vowel.

The effect on both good and neutral clusters seems to reside in the fact that clusters are easier to recognise if they are presented as clusters. This is corroborated by the fact that the [suC] manipulation was much worse than the [rsC] manipulation.

The clusters cannot be reduced to their parts, indicating that we have to assume at least some phonological property of the combination that is not derived from its parts. Still, the phonological property itself has not been identified.

5. Acknowledgements

This research is funded by the Netherlands Organisation for Scientific Research NWO, under project grant 277-70-001. The author would like to thank René Kager for helping me formulate my hypothesis. Together with Keren Shatzman he pushed me to execute the experiment. I am indebted to Keren for her help with the design of the experiment. Hugo Quéné deserves many thanks for putting me on track with the statistical analysis. I thank the three anonymous reviewers for pointing out flaws. Finally, I would like to thank the other members of our research group at the time of designing the experiment, Frans Adriaans and Natalie Boll-Avetisyan, for discussing my work, which was of great value to force me to organise my thoughts and mold them into proper hypotheses. Last but not least I thank Frits van Brenk for his work during the stimulus creation.

6. References