A ROLE FOR PHONOTACTIC CONSTRAINTS IN SPEECH PERCEPTION

Keren B. Shatzman and René Kager

Utrecht Institute of Linguistics OTS, University of Utrecht, the Netherlands
E-mail: Keren.Shatzman@let.uu.nl, Rene.Kager@let.uu.nl

ABSTRACT

This study investigated whether abstract, gradient phonotactic constraints play a role in speech processing. Dutch listeners performed an auditory lexical decision task, in which the nonword stimuli either did or did not violate a phonotactic constraint. Listeners were faster to reject nonwords that violated a phonotactic constraint. This effect remained significant even after partialling out the effects of lexical factors, such as the similarity of the nonwords to existing words in the lexicon. This finding constitutes, to our knowledge, the first demonstration of the involvement of pure abstract phonotactic constraints in on-line speech perception.

1. INTRODUCTION

Is the perception of spoken language influenced by abstract phonotactic constraints? While there is ample evidence showing that phonotactic knowledge is used in spoken-word recognition (e.g. [1]), and that categorical constraints affect speech perception (e.g., [2]), evidence for the independent role of abstract, gradient phonotactic constraints has been lacking. Claims about the psychological reality of phonotactic constraints have mostly been based on wordlikeness judgments studies, showing that participants judge nonwords which violate a constraint to be less wordlike than words that do not violate that constraint (e.g., [3], [4]). Although these studies certainly demonstrate listeners' gradient sensitivity to phonotactic well-formedness, they do not provide proof that this knowledge is used in the on-line process of speech perception. The wordlikeness judgment task is an off-line metalinguistic task, and could well be influenced by processes that do not occur in normal speech perception. Furthermore, in most of these studies the stimuli were presented orthographically. It is possible that these results may be valid only to reading. Moreover, the similarity of the nonwords to lexical entries was not controlled for in many of these studies. That is to say, words that violate a constraint tend to be less similar to existing words. It is, therefore, unclear whether participants' judgments are really due to pure phonotactic constraints or reflect the influence of the similarity of the nonwords to existing words. The few studies that used on-line measures with auditory stimuli and reported effects of abstract phonotactic constraints (e.g., [5], [6]) also did not manage to deconfound abstract phonotactic constraints from lexical statistics. Thus, clear evidence for the involvement of abstract phonotactic constraints in speech perception remained so far unattested. The current study examined this issue by measuring listeners' responses in a lexical decision task to nonwords that did or did not violate a gradient and abstract phonotactic constraint, and using regression analysis to tease apart the effects of lexical factors from those of constraints.

The constraint we focused on is one governing the distribution of labials in Dutch. Like many other languages, Dutch shows a statistic under-representation of words with consonants that share place of articulation. As many languages attest, the restriction is stronger for sequences of labials than for coronals. This was affirmed by analyzing a lexicon of Dutch underived stems (N=8,305). The Observed/Expected (O/E) ratio was computed for C_1VC_2 sequences for which C_1 and C_2 share place of articulation, by counting the number of such CVC sequences in the lexicon (Observed value) and dividing it by the Expected value, which is the value to be expected if C_1 and C_2 were combining freely. The expected value is computed as the probability that C_1 occurs in the initial position of CVC, multiplied by the probability that C_2 occurs in the final position of CVC, multiplied by the total number of CVC sequence tokens in the lexicon. This calculation was done for C_1VC_2 sequences in which C_1 and C_2 were either identical or non-identical, and where C_1 was either in word-initial or non-initial position. The results of this analysis are shown in Table 1, separately for labials (the phonemes /p, b, f, v, m/; hereafter, P) and coronals (the phonemes /t, d, s, z, n/; hereafter, T).

Table 1: Observed/expected values for C_1VC_2 sequences (C_1 and C_2 share place of articulation)

C_1 in word initial position		
	Coronals	*Labials*
Identical consonants	0.725	0.956
Non-identical consonants	0.880	0.444
C_1 in non-initial position		
	Coronals	*Labials*
Identical consonants	0.281	0.134
Non-identical consonants	0.827	0.091

The results of this analysis showed an interesting distributional pattern. Words with two identical coronal consonants (e.g., *totaal* [total]) are slightly under-represented in the lexicon, while words with two labial consonants (e.g., *papier* [paper]) are not under-represented. In contrast, words with two non-identical labial consonants are under-represented (i.e., words like *maf* [crazy] are rare), but words with two non-identical coronals (e.g., *tas* [bag]) only slightly so. In non-initial position, both coronals and labials show an under-representation of words with identical consonants (e.g., *staat* [state]). However, with non-identical consonants, coronals are just slightly under-represented (i.e., there are many words like *steen* [stone]), while there are only a handful of words with two non-identical labials (e.g., *spam* [spam]).

Traditionally, the under-representation of words in which consonants sharing place of articulation co-occur has been attributed to the Obligatory Contour Principle ([7], [8]). In classical Optimality Theory (OT) this has led to the postulation of OCP-PLACE constraints, which prohibit adjacent identical elements. In OT terms these constraints would be ranked: OCP-LAB » OCP-COR (following from the universal markedness relation *LAB » *COR, [9]). Alternatively, a self-conjoined constraint has been suggested [10, 11], prohibiting two identical elements per word. Again, following the universal markedness relation, this leads to the ranking $*\text{LAB}^2$ » $*\text{COR}^2$. Neither of these constraints could, however, account for the effects of initial versus non-initial CVC sequences that was observed in Dutch. That is, they do not explain why there are so few words like *spam*, while words like *steen* are not uncommon. We therefore tentatively propose the place feature alignment constraints ALIGN-LAB (every labial must be word-initial) and ALIGN-COR (every coronal must be word-initial), which are ranked ALIGN-LAB » ALIGN-COR. The constraint ALIGN-LAB adds one violation for each labial that is non-initial.

Regardless of which of these gradient constraints accounts best for the distributional pattern of labials in Dutch, the main question of the current study was whether any of these constraints might be involved in speech processing. We used the lexical decision task to test if we could find any evidence of such involvement. The rationale of the experiment was that if constraints influence perception then nonwords should be rejected more quickly if they contain a phonotactically ill-formed structure. Hence, reaction times to nonwords that violate a constraint should be shorter. There is, however, a caveat to this reasoning. Nonwords that violate a constraint tend to carry less resemblance to existing words. Shorter reaction times are, therefore, also predicted just on the basis of the similarity of the nonword to words in the lexicon. Although in constructing the stimuli, we tried to minimize the differences between constraint violating and non-violating nonwords in their similarity to words in the lexicon, it was impossible to balance the sets completely. We therefore computed for each nonword several measures of its similarity to words in the lexicon, and made use of multilevel regression analysis to disentangle the effects of the lexicon from possible effects of phonotactic constraints. If gradient phonotactic constraints influence perception, reaction times should be shorter for words violating a constraint, independently of the effects attributed to lexical factors.

2. METHOD

2.1. Participants

Twenty volunteers from Utrecht University were paid for their participation. They were all native speakers of Dutch, with no known hearing problem.

2.2. Materials

The experimental items were 192 bisyllabic nonwords. Half of these had a CVCVC structure, with consonants being either a coronal ([t], [n] or [s]) or a labial ([p], [m] or [f]), resulting in 8 combination types (TTT, TTP, TPT, TPP, PTT, PTP, PPT and PPP). The first consonant was either a plosive or a nasal, and adjacent consonants never shared manner of articulation. For each combination type there were, thus, six manner patterns: plosive-nasal-plosive, plosive-nasal-fricative, plosive-fricative-nasal, nasal-fricative-nasal, nasal-plosive-fricative and nasal-plosive-nasal. Two items of each pattern were constructed, differing in the vowels, resulting in 96 (2x6x8) items. The other 96 nonwords had a sCVCVC structure and were otherwise identical to the CVCVC items. All items had word-final stress. None of them contained strictly illegal sequences.

For each nonword item, the following lexical factors were computed using CELEX [12]: (1) Lexical neighborhood density (the sum of logged frequencies of all the words that arise by inserting, deleting or substituting one phoneme of the nonword; hereafter LND); (2) Cohort density (the sum of logged frequencies of all words that share the initial three phonemes of the nonword); (3) Transitional probability (the probability that a segment occurs given the preceding segment; computed per nonword as the logged product of the biphone transitional probabilities), and (4) Isolation point (the number of segments of the nonword after which there are no real words that match it). Additionally, each item was coded for the number of violations it caused, for each gradient constraint. For example, items of the type PTP violate the constraints ALIGN-LAB (the last labial is misaligned) and *LAB^2 (there are two labials in the word), but not OCP-LAB (no two adjacent labials), while TPP items violate *LAB^2 (there are two labials in the word) and OCP-LAB (two adjacent labials), but ALIGN-LAB is violated twice (there are two misaligned labials).

In addition to the experimental items, 192 bisyllabic words were selected for the "YES" trials. A further 40 nonwords were constructed as fillers, as well as 10 practice trials. All stimuli were read by a female speaker of Dutch in a sound-attenuated cabin and recorded with a sample frequency of 44.1 kHz.

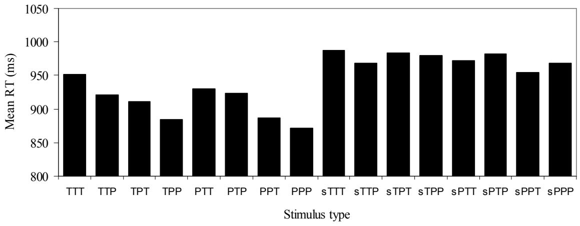
2.3. Procedure

Participants were tested individually in a sound-attenuated cabin. The stimuli were presented over headphones at a comfortable volume. Participants were asked to listen to the words, and to decide as quickly and as accurately as possible whether what they heard was a real Dutch word. They had to respond by pressing one of two buttons, labelled "YES" and "NO". The "NO" button was assigned to participants' preferred hand. All participants received the practice trials, followed by a random order of stimuli presentation, different for each participant.

3. RESULTS

Reaction times (RTs) were measured from the onset of the nonwords stimuli. Erroneous responses were excluded from the analysis (3% of the data). Mean RTs are shown in Figure 1, separately for each stimulus type.

Figure 1: Mean lexical decision latencies, per stimulus type



To analyze the results we fitted a multilevel regression model to the RTs, treating participant and item as random factors. The advantage of such an analysis is that it allows us to estimate the effects of lexical variables on RTs, and examine whether phonotactic constraints have an independent effect on response latencies. To normalize the distribution of variables, RTs and cohort density were logarithmically transformed.

In addition to the lexical factors, two other factors were included in the analysis, to partial out the influence of stimulus length on RTs. As Figure 1 shows, the RTs for sCVCVC items were longer than for CVCVC items. This could simply be the result of the fact that sCVCVC items are longer. Therefore, stimuli duration (in milliseconds) and the stimuli's number of phonemes were included as additional predictors.

In the first step, we fitted a regression model with log RT as dependent variable and the lexical factors (LND, logarithmic cohort density, transitional probability and isolation point), and stimulus length factors (stimuli duration and stimulus length) as predictors, but without the phonotactic constraints (hereafter, model A). The regression analysis revealed a significant inhibitory effect of logarithmic cohort density (F (1, 3706) = 26.54, $p < 0.0001$), and of stimuli duration (F (1, 3706) = 33.03, $p < 0.0001$). Thus, RTs were longer when there were many existing words overlapping with the initial three phonemes of the nonword. Likewise, longer stimuli duration caused longer RTs. After partialling out the effects of cohort density and stimulus length there were no significant effects of the other variables (Fs < 1).

In the next step, we tested whether the constraints had a significant effect on response latencies. Each constraint was added separately as an independent variable to model A. In each of these regression analyses, as with model A, a significant effect of cohort density and stimuli duration emerged, with no other significant effects of lexical factors or of stimulus length. The results for the constraints are shown in Table 2. As these results indicate, there was a small but significant effect of each constraint on its own on response latencies. The coefficient estimates for the constraints had a negative sign, denoting a facilitatory effect. That is to say, nonwords that violated a constraint had shorter RTs (i.e., were rejected faster).

Table 2: Coefficient estimates and ANOVA results for the effects of the phonotactic constraints

Model	Predictor	Coefficient estimate	*ANOVA*
Model A +	OCP-LAB	-0.01376	$F = 6.26, p < 0.05$
Model A +	*LAB^2	-0.01198	$F = 4.68, p < 0.05$
Model A +	ALIGN-LAB	-0.01469	$F = 9.08, p < 0.005$

Note. Degrees of freedom are 1, 3705

In these analyses, phonotactic constraints emerge as predictors of response latencies, after partialling out the effects of lexical factors and of stimulus length. This result suggests that abstract constraints play a role in speech processing which is not reducible to individual lexical entries.

Finally, we examined which of the constraints was the best predictor, by adding all three constraints to model A. In this analysis, therefore, the effect of each constraint was evaluated after partialling out the effects of the other variables, including the two other constraints. As before, there was a robust and significant inhibitory effect of cohort density (F (1, 3703) = 25.20, $p < 0.0001$), and stimuli duration (F (1, 3703) = 31.69, $p < 0.0001$), with no other significant effects of lexical factors or of stimulus length. With respect to the constraints, the results showed that after entering *LAB^2 and ALIGN-LAB to model A, OCP-LAB did not have a significant effect ($F < 1$). Similarly, when ALIGN-LAB and OCP-LAB were included in model A, *LAB^2 had no additional significant effect (F (1, 3703) = 1.14, $p > 0.1$). However, when entering ALIGN-LAB to the model with OCP-LAB and *LAB^2 already included, the effect of ALIGN-LAB still emerged significant (F (1, 3703) = 3.91, $p < 0.05$). This result indicates that after partialling out the effects of all the other factors, there is still a significant effect of ALIGN-LAB on response latencies. In contrast, when the effect of ALIGN-LAB is first partialled out, the effects both OCP-LAB and *LAB^2 are no longer significant. This suggests that the variance in the data that is explained by the phonotactic factors is best captured by the constraint ALIGN-LAB.

4. CONCLUSIONS

This study was designed to assess the contribution of gradient phonotactic constraints to on-line speech processing. The results of a lexical decision task showed that listeners rejected nonwords that violated abstract, gradient constraints faster than nonwords that did not violate a constraint. This effect was independent of the effects of lexical factors. This finding strongly suggests that abstract phonotactic constraints have an effect on speech perception. Furthermore, of the three constraints that could account for the distribution of labials, the constraint ALIGN-LAB appeared to be the best predictor of response latencies.

5. REFERENCES

[1] McQueen, J. M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39, 21-46.

[2] Moreton, E. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84, 55-71.

[3] Berent, I., Shimron, J., Vaknin, V. 2001. Phonological constraints on reading: Evidence from the obligatory contour principle. *Journal of Memory and Language* 44, 644-665.

[4] Frisch, S. A., Zawaydeh, B. A. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77, 91-106.

[5] Coetzee, A. W. 2005. The OCP in the perception of English. In: S. Frota, M. Vigario M. Freitas (eds.), *Prosodies*. New York: Mouton de Gruyter, 223-245.

[6] van de Weijer, J. 2005. Listeners' sensitivity to consonant variation within words. *Lund University Dept. of Linguistics Working Papers*, 225-238.

[7] McCarthy, J. 1986. OCP effects: gemination and anti-gemination. *Linguistic Inquiry* 17, 207-263.

[8] Frisch, S. A., Pierrehumbert, J. B. Broe, M. B. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22**,** 179-228.

[9] Prince, A., Smolensky, P. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report 2, New Brunswick.

[10] Alderete, J. 1997. Dissimilation as local conjunction. *Proceedings of the North East Linguistic Society 27* Amherst, 17-32.

[11] Itô, J., Mester, R. 2003. *Japanese Morphophonemics: Markedness and Word Structure*. Boston: MIT Press.

[12] Baayen, H., Piepenbrock, R., Gulikers, L. 1995. The CELEX lexical database. Philadelphia, PA: Linguistics Data Consortium, University of Pennsylvania.