

The Obligatory Contour Principle in the perception of English¹

This paper reports on a phoneme identification experiment that establishes two facts: (i) That phonological grammar influences perception. (ii) That the Obligatory Contour Principle (OCP) has psychological reality for English listeners. In particular, the experiment shows that the OCP causes perceptual biases in phoneme identification tasks. Assuming that the OCP is part of phonological grammar, this shows that grammar does influence perception. It also shows that the OCP shapes the linguistic performance of English listeners, and therefore that it is a psychologically real part of their grammar.

In this paper I report on a phoneme identification experiment that establishes two facts: (i) That phonological grammar influences perception. (ii) That the OCP has psychological reality for English listeners. The experiment shows that the OCP causes perceptual biases in phoneme identification tasks. Assuming that the OCP is part of phonological grammar, this shows that grammar does influence perception. It also shows that the OCP shapes the linguistic performance of English listeners, and therefore that it is a psychologically real part of their grammar. The rest of this paper is structured as follows: In §1, I give the background necessary to understand the rationale for the experiment, presenting an overview of the literature on the non-acoustic factors that contribute to speech perception, as well as on the literature about the psychological reality of the OCP. In section §2, I discuss the experiment and its results. Finally, in §3, I show how the results of the experiment relate to the two questions stated above.

1. The background

1.1 Grammar mediated perception

Much research has been done to establish the factors other than the acoustical signal that influence perception. Two kinds of factors have been identified, namely lexical statistics and phonological grammar.

Lexical statistics. There is a large body of literature showing that statistical patterns in the lexicon influence perception. Three kinds of lexical statistics have been found relevant: (i) lexical frequency (Connine et al. 1993, Ganong 1980), (ii) lexical neighborhood density (Vitevitch and Luce 1998), and (iii) transitional probabilities (Newman et al. 1997).

There is a perceptual bias towards percepts with higher lexical frequencies. As an example: In a phoneme identification experiment, Connine et al. (1993) presented their listeners with a token ambiguous between *goat* and *coat*, and found a bias to identify the initial stop as [k], since *coat* has a higher lexical frequency than *goat*. But in a token between *gave* and *cave*, the bias was to [g], since *gave* has a higher lexical frequency than *cave*.²

Lexical neighborhood density also influences perception. Vitevitch and Luce (1998, 1999) found that, in tasks that require lexical access (e.g. lexical decision), non-words with denser neighborhoods are perceived slower and less accurately than non-words with sparser neighborhoods. They attribute this to competition between lexical neighbors. In a lexical decision task listeners search through the lexicon to determine whether the percept is in the lexicon. Listeners consider at least all the words in the immediate lexical neighborhood of the percept. For a token in a dense neighborhood, there are more possibilities to exclude before the token can be rejected as a non-word.

However, Vitevitch and Luce also found that perception is aided by a dense lexical neighborhood in tasks that do not require lexical access (e.g. same/different judgments). Newman et al. (1997) investigated this further with a phoneme identification task. They presented their subjects with a token ambiguous between two non-words, one with a denser neighborhood than the other. They found a bias toward the token with the denser neighborhood. For instance, between *gice* and *kice*, they found that the initial stop was identified more as [g] because *gice* has a denser neighborhood than *kice*. However, [k]-responses were more frequent for a token between *gipe* and *kipe*, because *kipe* has a denser neighborhood than *gipe*. Their interpretation: Non-words with dense lexical neighborhoods typically also have phoneme sequences that are frequent. The transitional probabilities of such tokens are higher than that of non-words from sparser neighborhoods. They argue that the listener can access this transitional frequency information without accessing the lexicon. In tasks that do not require lexical access, the higher transitional probabilities favor a token.

Phonological grammar. Phonological grammar, and in particular, phonotactics, has also been shown to influence perception (Dupoux et al. 1999, Massaro and Cohen 1983, Moreton 2002, Pitt 1998). The general finding is that, when listeners are presented with a token ambiguous between a phonotactically legal percept and a phonotactically illegal percept, there is a perceptual bias towards the phonotactically legal percept.

Consider the following example from Massaro and Cohen (1983). They presented listeners with tokens ambiguous between [tri] and [tli], with the task to identify the liquid as [r] or [l]. They found a bias towards [r]. This can be interpreted as a bias towards the phonotactically legal percept – #*tr*- is allowed word-initially in English but #*tl*- not. On tokens between [sri] and [sli], they found the opposite bias – more [l]-responses. This also corresponds to English phonotactics – #*sl*- is legal in English but #*sr*- is not.

However, there is a confound. Since #*tl*- and #*sr*- are illegal in English, there are no words with these initial sequences. The transitional probability between word initial [t] and following [l] is zero. Similarly, the transitional probability between word initial [s] and following [r] is zero. The results of Massaro and Cohen can therefore also be interpreted as a bias towards the token with the higher transitional probability *a la* Newman et al. (1997).

This identifies an important question: Can all perceptual biases apparently caused by phonological grammar also be explained in terms of transitional probabilities? If so, then grammar does not have to be afforded an autonomous role in perception. The influence of grammar can then be subsumed under that of lexical statistics. This is the route taken by most of the models of speech perception developed over the last few decades – TRACE (McClelland and Elman 1986), MERGE (Norris et al. 2000), and NAM (Luce and Pisoni 1998).

To decide whether grammar plays an autonomous role in perception, it is necessary to control for the possibility of interpreting phonotactics as a reflex of transitional probabilities. Moreton (2002) reports an experiment in which he did exactly this. He claims that although neither [bw] nor [dl] occurs word-initially in English, the absence of [bw] is accidental, while the absence of [dl] is due to a phonotactic constraint against it.³ The stop [g] differs from both [d] and [b] – it combines with [w] and [l] word-initially, [gw] and [gl]. Moreton constructed one array of stimuli ambiguous between [gwæ, glæ, bwæ, blæ], and a second ambiguous between [gwæ, glæ, dwæ, dlæ]. Since [gw] and [gl] are both attested, he assumed that there will be no biases in the identification of the [w] or [l] when the initial consonant is identified as [g]. The [gwæ, glæ] tokens act as a baseline for comparison. However, since [dl] is illegal word-initially, Moreton hypothesizes that identification of the initial consonant as [d] will result in a bias towards [w] for the second consonant. On the other hand, since [bw] is only accidentally absent from the lexicon, identification of the initial consonant as [b] should not bias the identification of the second towards [l]. This prediction of Moreton was borne out by the results of his experiment. Listeners were more likely to identify the second consonant in the cluster as [w] when they identified the first consonant as [d]. But when the initial consonant was identified as [b] he found no bias towards [l] or [w] for the second consonant. He interprets this as a result of [dlæ] being phonotactically illegal, and both [blæ] and [bwæ] phonotactically legal.

These results cannot be ascribed to transitional probabilities. Given a vowel [æ], the probability of the preceding sound being [l] is 0.04, while the

probability of it being [w] is 0.12.⁴ This could explain the bias to [w] in the [dwæ, dlæ]-condition. But had it been this transitional probability that caused the bias, then the same bias should have been found in the [bwæ, blæ]-condition. And Moreton found no bias in this condition. It is also unlikely that lexical neighborhood density contributed much towards the results. Since [dw] is found word-initially but [dl] not, [dwæ] will have a denser neighborhood than [dlæ]. This could contribute to the [w]-bias in this condition. However, [bw] and [bl] stand in the same relationship – [bl] is found word-initially and [bw] not, so that [blæ] will have a denser neighborhood than [bwæ]. If the [w]-bias after [d] were caused by a denser neighborhood, then there should be a bias to [l] after [b]. Moreton's results are evidence that (i) phonological grammar influences perception, and that (ii) the effects of lexical statistics can be overridden by grammar.

There is therefore evidence for the influence of both lexical statistics and phonological grammar on perception. Most research has focused on lexical statistics and has tended to interpret phonotactics as transitional probability. The first aim of this paper is therefore to add to the evidence in favor of the independent contribution of phonological grammar to perception.

1.2 The psychological reality of the OCP

It is known for several languages that words with homorganic (i.e. also identical) consonants are underrepresented in their lexicons. This has been shown of Arabic (Frisch et al. 2000), Hebrew (Greenberg 1950, McCarthy 1985), French, English and Latin (Berkley 2000). These effects are usually ascribed to the Obligatory Contour Principle (Leben 1978, McCarthy 1986). However, restrictions like these are seldom involved in active phonological alternations. It is therefore not clear what the status of the OCP is in the grammar of these languages. Is it an active part of the grammar, or is it simply a stative statistical generalization over the structure of the lexicon?

Recent research into the OCP started to address this question, and the evidence points to the OCP shaping the linguistic competence and performance of language users. Even though it does not take part in active alternations, it is therefore a true and psychologically real part of grammar.

Berent and her colleagues have studied the effect of the OCP on the linguistic performance of Hebrew speakers in a series of papers. I discuss as an example one of the results of Berent et al. (2001). The OCP in Hebrew dictates that tri-consonantal roots are not allowed to have homorganic consonants as first two consonants. Berent et al. found in a lexical decision task that subjects were faster in rejecting non-words that violate the OCP than non-words that do not violate the OCP. They found this effect even

when they controlled for the contribution of lexical statistics (particularly transitional probabilities). They interpret this as showing that the OCP is a constraint that is actively part of the Hebrew speaker's linguistic competence independently from mere lexical statistical effects.

Frisch and Zawaydeh (2001) obtained similar results for Jordanian Arabic. Arabic has the same version of the OCP as Hebrew. Frisch and Zawaydeh constructed a list of novel tri-consonantal roots. Their list had some roots that violate the OCP (i.e. with homorganic consonants in first and second position), and some that do not. The roots with an OCP-violation represent a systematic gap in the Arabic lexicon. The non-OCP-violating roots that they used were constructed such that they also represent gaps in the lexicon. Since these forms obey the OCP, these gaps are assumed to be accidental. The novel roots all had very similar lexical neighborhood densities and transitional probabilities. The only relevant difference between the two kinds of novel roots was whether they violated the OCP or not.

Their subjects had to rate the roots for their well-formedness/wordlikeness. They found that the novel roots with an OCP-violation were judged as less acceptable than roots that do not contain an OCP-violation. These results show that speakers of Jordanian Arabic have access to a phonotactic constraint (the OCP) independently from frequency patterns in the lexicon.

There is therefore evidence that the OCP is an active, psychologically real part of the grammar of Hebrew and Arabic. This paper will provide evidence that the OCP is similarly an active part of the grammar of English, even though the OCP is not involved in any phonological alternation in English.

2. The experiment

2.1 The rationale behind the experiment: an OCP effect in English

English has restrictions on words of the form [sCvC] where the two C's are both stops (Davis 1991, Fudge 1969). Words with two [k]'s or two [p]'s are not allowed (**speap* and **skeak*).⁵ However, two coronal stops [t] are allowed (*state*, *stout*, *stoat*, etc.). Also, words with two heterorganic stops are allowed (*speak*, *skip*, *stop*, *skate*, etc.). Since this is a restriction on the co-occurrence of homorganic (identical) consonants within a word, it can be attributed to some form of the OCP.⁶

Based on the results of Moreton (2002), Berent et al. (2001), and Frisch and Zawaydeh (2001), we can expect that this restriction on possible words will influence perception in English. The following is expected: If English listeners are presented with a token between [skʌk] and [skʌp], there will be a bias towards identifying the final consonant as [p]. Identifying it as [k] will

result in an OCP-violating percept [skɹk]. On the other hand, if presented with a form between [spɹp] and [spɹk], listeners will be biased towards [k]. Here [p] will result in an OCP-violating percept [spɹp]. We expect more [p] responses in a [skɹp]~[skɹk]-context than a [spɹp]~[spɹk]-context. This is due to a bias against [k] in a [skɹp]~[skɹk]-context, and against [p] in a [spɹp]~[spɹk]-context. The first hypothesis of this paper is therefore:

(1) *Hypothesis 1:* The OCP does influence perception in English.

Hypothesis 1 can be confirmed by finding evidence for the kind of perceptual biases described in the paragraph above (1).

Why are coronals exempt from this restriction? Why is *state* a word but **skake* and **spape* are not even possible words? One possibility is that the OCP does not refer to coronal place. If this were so, then we would expect no bias on a token between [stet] and [stek], because neither [t] nor [k] will result in an OCP-violating percept. However, on a token between [skək] and [sket], we will expect a bias towards [t], since [k] will result in the OCP-violating percept [skək]. We will therefore expect more [k]-responses in the [stek]~[stet]-condition than the [skək]~[sket]-condition. However, this will be due to only one bias (against [k] in the [skək]~[sket]-condition). The effect in this K~T-condition is then due to only one bias, while the effect in the P~K-condition as discussed above is due to two biases. If there were indeed no OCP-constraint against [sTvT] forms, then it is likely that the bias effect will be smaller in the K~T-condition than in the P~K-condition.

There is another way to interpret the fact that [sTvT]-forms are allowed in English. It is possible that there is a constraint against English words of the form [sTvT], but that English is willing to violate this constraint. This seems more likely. Berkley (2000) found that words with two coronals are underrepresented in the English lexicon, even if less so than words with two labials or velars. Similar patterns have been identified for Hebrew and Arabic – there are restrictions on the co-occurrence of coronals within a root, even if these restrictions are less strict than those on labials and velars (Frisch, 2000). It therefore seems to be the case that the OCP also makes reference to coronal place. This is the interpretation that I will accept here.

Under this interpretation the bias effect in a K~T-condition will also be due to two biases. On a token between [stet] and [stek], we now expect more [k]-responses. Although [stet] is a possible word in English, it still violates the OCP while [stek] does not. Under this interpretation it is more likely that there will not be a difference in the size of the bias effect between the K~T-

condition and the P~K-condition. Since this is the interpretation taken in this paper, we can state a second hypothesis for the paper as in (2).

(2) *Hypothesis 2*: The OCP also makes reference to coronal place.

If no difference is found in the bias size between the K~P-condition and K~T-condition, that could be interpreted as evidence for *Hypothesis 2*.

2.2 Experimental design

The experiment was designed to test both hypotheses. Three sets of continua were constructed: (i) **[k]~[p]**: [skap]~[skak] and [spap]~[spak]; (ii) **[k]~[t]**: [skek]~[sket] and [stek]~[stet]; (iii) **[p]~[t]**: [spap]~[spat] and [stap]~[stat]. The tokens on these continua were presented to listeners, whose task was to identify the final stop consonant. The expected perceptual biases in each condition are summarized in *Table 1*.

Table 1. Predicted perceptual biases

Continuum	Predicted bias	Comment
[skap]~[skak]	[p]	Because *[skak]
[spap]~[spak]	[k]	Because *[spap]
[spap]~[spat]	[t]	Because *[spap]
[stap]~[stat]	[p]	Uncertain, because [stat] legal
[skek]~[sket]	[t]	Because *[skek]
[stek]~[stet]	[k]	Uncertain, because [stet] legal

Perceptual bias was measured in two ways. I discuss the P~K-condition as an example, but the other conditions were treated in the same way: (i) Total response pattern: The total percent [p]-responses on [skap]~[skak]-continuum and on the [spap]~[spak]-continuum was measured. Since a bias toward [p] is expected on the [skap]~[skak]-continuum and a bias against [p] on the [spap]~[spak]-continuum, a higher total percent [p]-responses is expected on the [skap]~[skak]-continuum. (ii) Category boundary/crossover-point: This is the point on the continuum that is identified equally often as [p] and [k]. This point is expected to be closer to [p] on the [spap]~[spak]-continuum (fewer tokens identified as [p]) because of the bias against [p]. Similarly, it is expected to be closer to [k] on [skap]~[skak]-continuum because of the bias against [k].

2.2.1 Method

2.2.1.1 Selection of tokens

Since the aim of this experiment was to find evidence for the effect of phonological grammar on perception, it was attempted to control for all other factors that are known to cause perceptual biases. For this reason the

tokens were chosen such that both endpoints of all continua are non-words. No lexical bias effect (Ganong 1980) should therefore be found.

I also tried to select tokens such that the transitional probabilities and the lexical neighborhood frequencies of the four endpoint tokens in each condition were equal. It turned out to be impossible to control for both of these statistics. I decided to control for the transitional probabilities rather than the lexical neighborhood densities, for the following reasons: (i) Since phoneme identification does not require lexical access, it is expected that transitional probabilities will be more important than lexical neighborhood density in the processing of the tokens (Vitevitch and Luce 1998). (ii) All tokens are monosyllabic. It is known that lexical effects tend to decrease or even disappear in experimental setups where the stimulus lists contain only monosyllabic tokens (Cutler et al. 1987). This also implies that the lexical neighborhood frequency is less likely to contribute to the processing of the stimuli used in the experiment.

The four endpoint tokens in each condition were selected such that their transitional probabilities were all of the same order of magnitude. *Table 2* contains the lexical neighborhood densities and the transitional probabilities for the twelve endpoint tokens. These statistics were calculated using the CELEX database (Baayen et al. 1995).⁷ The method used in the calculation is in all relevant respects identical to that used by Vitevitch and Luce (1999).

Table 2. Lexical Neighborhood Densities and Transitional Probabilities⁸

Endpoint	Lexical Neighborhood Density	Probability of V given final C	Probability of final C given V
[spʌp]	21.41	0.064	0.036
[spak]	35.69	0.052	0.021
[skʌp]	31.57	0.064	0.036
[skak]	18.41	0.052	0.021
[skɛk]	12.65	0.079	0.078
[skɛt]	34.37	0.048	0.018
[stɛk]	47.58	0.079	0.078
[stɛt]	38.11	0.048	0.018
[spʌp]	11.38	0.033	0.050
[spʌt]	31.13	0.064	0.027
[stʌp]	46.08	0.033	0.050
[stʌt]	42.19	0.064	0.027

2.2.1.2 Construction of stimuli

The same procedure was used in all six continua. The procedure is described here only for the [skap]~[skak]- and [spap]~[spak]-continua. Recordings were made in a sound attenuated room. A male speaker of American English was recorded reading the non-words [skap] and [spak] in the carrier sentence *John said _____ again to me*. Four repetitions of each token were recorded.

The [skap]~[skak]-continuum: A 20-step continuum was synthesized. For the initial part of each of the tokens on the continuum, the [sk] and initial 40% of the nucleus of the recorded non-word [skap] were used to set the synthesis parameters. For the last part of [skap]-endpoint token, the last 60% of the nucleus plus the final stop [p] of the recorded non-word [skap] supplied the synthesis parameters. For the last part of [skak]-endpoint token, the last 60% of the nucleus plus the final stop [k] of the recorded non-word [spak] supplied the parameters. The synthesis parameters for the last half of the 18 intermediate tokens were formed by mixing the last 60% of the nuclei plus final stops ([p] or [k]) of the two recorded non-words in different proportions. This resulted in a 20-step continuum from [skap] to [skak]. The tokens were identical in their onset clusters [sk] and the first 40% of their nuclei. They differed only in the last 60% of their nuclei and final stops.

The [spap]~[spak]-continuum: This continuum was constructed in the same way, except that the first part of the tokens was synthesized using the [sp] and initial 40% of nucleus of the non-word [spak]. The tokens in the [skap]~[skak]-continuum and the [spap]~[spak]-continuum were therefore identical, except for the initial cluster and the first 40% of the nucleus.

The [skək]~[sket]-, [stek]~[stet]-, [spʌp]~[spʌt]-, and [stʌp]~[stʌt]-continua were constructed in a similar way. This resulted in tokens of high quality that sounded like natural speech.⁹

2.2.1.3 The subjects

All subjects were undergraduate students from the University of Massachusetts, who took part in the experiment for extra course-credit. Subjects were native speakers of American English, and reported no hearing disabilities. A total of 37 subjects took part. In each condition only those subjects who could identify the endpoint stimuli correct at least 75% of the time were included in the analysis. This resulted in 15 subjects in the K~P-condition, 26 in the K~T-condition, and 26 in the P~T-condition.

2.2.1.4 The procedure

Three experiment files were created, one each for the P~K-, P~T- and K~T-conditions. The structure of the P~K experiment file is discussed here. The other two files were constructed in an analogous way.

The P~K experiment file contained three repetitions of the four endpoint tokens [skak], [skap], [spak] and [spap]. Of the 18 intermediate tokens from each continuum, the six that surrounded the most ambiguous region on the continuum were selected.¹⁰ Six repetitions of each of these six ambiguous tokens were included in the experiment file. The six ambiguous tokens from the [skap]~[skak]-continuum and the [spap]~[spak]-continuum were identical in the last 60% of their nuclei and final stops. They differed only in their initial consonant clusters and the first 40% of their nuclei.¹¹ Filler items were also included. In the [p]~[k]-condition the filler items came from a [trap]~[trak]-continuum. This continuum was constructed in the same way as described above for the other [p]~[k]-continua. Three repetitions of the two endpoint tokens from the filler continuum, as well as six repetitions of the six intermediate tokens were included. The filler items were included to ensure that there was enough variation in token initial position so that listeners will pay attention to this.

Experiment files for the other conditions were created in the same way. In the P~T-file, the fillers came from a [krʌp]~[krʌt]-continuum, and in the K~T-file from a [prɛk]~[prɛt]-continuum.

The three conditions were tested separately. Stimuli were played to subjects over headphones, who had to press a button corresponding to the sound on which the token ended. The experiment was conducted in two blocks. In the first block, each of the experiment files was presented separately to the subjects. The tokens in each file were presented in random order. The second block was a repetition of the first. Each time an experiment file was presented, it was preceded by 18 practice trials without feedback, to familiarize subjects with the procedure and the tokens. The practice trials consisted of three repetitions of each of the six endpoint tokens in an experiment file.

2.2.2 The results

Each of the three conditions was treated separately. Response times were recorded, starting at the onset of the presentation of an experimental token. Responses that were recorded within 200 ms were not included in the analysis. When a subject responds as quickly as this, it is likely that he/she has not yet heard enough of the token to have received the auditory information necessary to identify the final consonant.

K~P. The percent [p]-responses of each subject on the [skap]~[skak]- and [spap]~[spak]-continua was calculated. According to *Hypothesis 1* there should be more [p]-responses on the [skap]~[skak]-continuum, because of a bias against [k]. Conversely, fewer [p]-responses are expected on the [spap]~[spak]-continuum because of a bias against [p]. The average percent [p]-responses on the [spap]~[spak]-continuum was 39%, and on the [skap]~[skak]-continuum it was 48% (one-tailed $t(14) = 1.712, p < .05$). *Figure 1* displays the mean identification functions for these two continua.

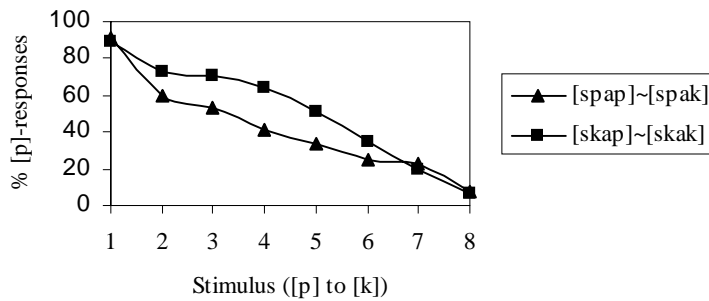


Figure 1. Group identification functions for the P~K-condition

A second measure of the bias is the category boundary. This is a less accurate measure, as it measures the influence of the bias at only a single place on the continuum. Comparing the total percent [p]-responses across a continuum measures the bias across the whole continuum.¹² In the calculation of category boundary only those subjects for whom the crossover point of both continua was in the middle six tokens were used. Crossover points were calculated by linear interpolation. Of the 15 subjects in this condition, 9 had crossover points in the ambiguous region. The mean crossover point on the [skap]~[skak]-continuum was 5.09, and on the [spap]~[spak]-continuum it was 3.69. This difference is in the predicted direction – the lower crossover point on the [spap]~[spak]-continuum corresponds to the lower percent of [p]-responses, and is predicted by the bias against [p]-percepts. A one-tailed t -test on the differences in category boundaries for each subject yielded the following: $t(8) = 1.852, p < .05$.

K~T. The percent [k]-responses of each subject on the [skək]~[sket]- and [stek]~[stet]-continua were calculated. Fewer [k]-responses are expected on the [skək]~[sket]-continuum, because of a bias against [k] on this continuum. The average percent [k]-responses on the [skək]~[sket]

continuum was 47%, and on the [stek]~[stet]-continuum 55% (one-tailed $t(25) = 2.822, p < .005$). *Figure 2* displays the mean identification functions for these two continua.

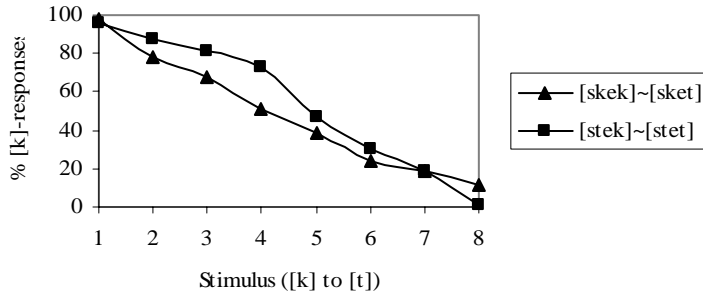


Figure 2. Group identification functions for the K~T-condition

Of the 26 subjects in this condition, 21 had crossover points in the ambiguous region. For these subjects, the mean crossover point on the [skek]~[sket]-continuum was 4.66, and on the [stek]~[stet]-continuum 4.95. The lower crossover point on the [skek]~[sket]-continuum corresponds to the lower percent of [k]-responses on this continuum, and is predicted by the bias against [k]. However, the difference is small and does not turn out to be significant (one-tailed $t(20) = 1.118, p < .14$).

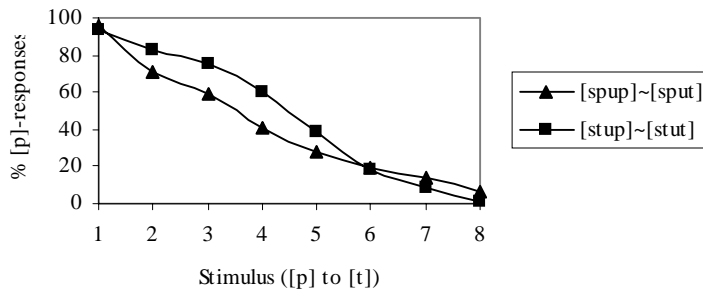


Figure 3. Group identification function for the P~T-condition

P~T. The total percent [p]-responses of each subject on the [spup]~[sput]- and [stup]~[stut]-continua were calculated. We expect fewer [p]-responses on the [spup]~[sput]-continuum, because of a bias against [p] on this continuum. The average percent [p]-responses on the [spup]~[sput]-

continuum was 40%, and on the [stɒp]~[stɒt]-continuum 47% (one-tailed $t(25) = 1.897, p < .04$). *Figure 3* displays the mean identification functions for these two continua across listeners.

Of the 26 subjects in this condition, 20 had crossover points in the ambiguous region. For these subjects, the mean crossover point on the [spɒp]~[spɒt]-continuum was 3.74, and on the [stɒp]~[stɒt]-continuum 4.49. The lower crossover point on the [spɒp]~[spɒt]-continuum corresponds to the lower percent [p]-responses, and is predicted by the bias against [p]-percepts on this continuum. A one-tailed t -test also reveals that this difference is significant ($t(19) = 2.460, p < .01$).

Table 3 and *Table 4* summarize the results of the experiment.

Table 3. Total percent responses

Continuum	Bias predicted toward	Responses	One-tailed t -test
[skɒp]~[skɒk]	[p]	% [p] = 48	$t(14) = 1.71, p < .05$
[spɒp]~[spɒk]	[k]	% [p] = 39	
[stɛk]~[stɛt]	[k]	% [k] = 55	$t(25) = 2.82, p < .005$
[skɛk]~[skɛt]	[t]	% [k] = 47	
[stɒp]~[stɒt]	[p]	% [p] = 47	$t(25) = 1.90, p < .04$
[spɒp]~[spɒt]	[t]	% [p] = 40	

Table 4. Category boundaries

[P]~[K]	[p]	Crossover	[k]
[spɒp]~[spɒk]	1	3.69	8
[skɒp]~[skɒk]	1	5.09	8
$t(8) = 1.852, p < .05$			
[K]~[T]	[k]	Crossover	[t]
[skɛk]~[skɛt]	1	4.66	8
[stɛk]~[stɛt]	1	4.95	8
$t(20) = 1.118, p < .14$			
[P]~[T]	[p]	Crossover	[t]
[spɒp]~[spɒt]	1	3.74	8
[stɒp]~[stɒt]	1	4.49	8
$t(19) = 2.460, p < .01$			

Comparison between conditions. According to *Hypothesis 2* there should be a bias against [t] on the [stɹp]~[stɹt]-continuum, and on the [stek]~[stet]-continuum even though English tolerates words of the form [sTɹT]. If this is correct, the difference in the two identification functions in the P~T- and K~T-conditions, is the result of a double bias. The size of the effect in these conditions should then be comparable to the size of the difference in the K~P-condition – where the double bias follows because English tolerates neither [sKvK] nor [sPvP] words. To test this, a difference score was computed for each subject in each experimental condition. For instance, in the K~P-condition this score was computed by subtracting the percent [p]-responses of the subject in the [spɹp]~[spɹk]-continuum from the percent [p]-responses on the [skɹp]~[skɹk]-continuum. The difference scores in the other conditions were computed in an analogous manner. Two-tailed *t*-tests on each of the three pairs of difference scores were performed to test whether the size of the effect was different between the three conditions. No significant difference was found. The values obtained are: (i) P~K- and P~T-conditions: $t(13) = 0.403$, $p < 0.70$; (ii) P~K- and K~T-conditions: $t(10) = 0.748$, $p < 0.48$; (iii) P~T- and K~T-conditions: $t(21) = 0.829$, $p < .42$.

The same analyses were done in terms of the difference scores for category boundaries. Two-tailed *t*-tests show that there is no difference between the difference scores for the three conditions. (i) P~K- and P~T-conditions: $t(8) = 0.548$, $p < 0.60$; (ii) P~K- and K~T-conditions: $t(4) = 0.495$, $p < 0.65$; (iii) P~T- and K~T-conditions: $t(12) = 0.662$, $p < .53$.

2.2.3 Discussion

The results confirm both *Hypotheses*. *Hypothesis 1* (that the OCP influences perception) was confirmed by the fact that in all three conditions a bias was found against identifying the final consonant in such a way that the resulting percept would violate the OCP. In all three conditions, this was found when the total percent responses for one of the two percepts was used as a measure of the effect. In the P~T- and P~K-condition, this was also found for category boundaries. In the K~T-condition, the category boundary shift was in the predicted direction, but did not reach significance.

This gives more evidence that phonological grammar influences perception. The tokens were selected with very similar transitional probabilities, and experimental conditions made it unlikely that lexical neighborhood density would have a significant influence.¹³ The only relevant difference between the percepts in each condition was whether a percept violated the OCP.

Hypothesis 2 was that the size of the effect should be approximately equal in all three of the conditions. This would imply that there is a bias against [t] on

the [stʌp]~[stʌt]- and the [stɛk]~[stɛt]-continua, although English tolerates [sTʌT] words. This was confirmed by the fact that no significant difference was found in the size of the bias between the experimental conditions. This implies that even possible (or real) words can violate phonotactic constraints and that listeners have access to this knowledge. Language users can distinguish between degrees of acceptability, even above the absolute cut-off point for grammaticality. [stɛt] and [stɛk] are possible words of English, but [stɛt] violates a phonotactic constraint that [stɛk] does not violate.

3. General discussion

This section is structured as follows: Section §3.1 discusses the possible contribution of lexical neighborhood density to the effects observed in the experiment. Section §3.2 then summarizes the relevance of the results for the theories of speech perception and for our understanding of the OCP.

3.1 The possible contribution of lexical neighborhood density

As explained in the discussion of the experimental design (§2.2.1.1), it was not possible to equalize both transitional probabilities and lexical neighborhood densities across the tokens used in experiment. I used tokens with nearly equal transitional probabilities, since the experimental design was such that lexical neighborhood densities were unlikely to cause perceptual biases. Even so, comparing the lexical neighborhood counts of the tokens (*Table 2*) with the biases found, it is clear that the biases are in the direction of the token with the higher lexical neighborhood in every instance. For example, [spak] inhabits a denser neighborhood than [spap], and on the [spap]~[spak]-continuum, there was a bias towards [k]. Although it is not expected that lexical neighborhood would have much of an influence, it is still necessary to consider this possibility.

Consider the lexical neighborhood counts for [spak] (35.69) and [spap] (21.41). The density of the [k]-endpoint is 14.28 higher than that of the [p]-endpoint. This means that there is negative bias of 14.28 against the [p]-percept on the [spap]~[spak]-continuum. On the [skap]~[skak]-continuum, the neighborhood bias is in favor of [p]. The neighborhood density of [skap] is 31.57, and of [skak] 18.41. The difference between these two translates into a 13.16 positive bias towards [p]. Similar calculations were done for the other conditions. The results are shown in *Table 5*.

A regression analysis can be done using the values from *Table 5*. The values from the table can be used as the independent variable. The percent [p]-responses in the K~P-condition, the percent [k]-responses in the K~T-condition, and the percent [p]-responses in the P~T-condition for each

subject are used as dependent variable. The r^2 value attained for such an analysis is 0.065. The lexical neighborhood density differences account for only a small fraction of the variation observed in the data.

Table 5. Lexical neighborhood biases in the three experimental conditions

K~P	Bias to [p]
[spap]~[spak]	-14.28
[skap]~[skak]	13.16
K~T	Bias to [k]
[skæk]~[skæt]	-21.72
[stæk]~[stæt]	9.47
P~T	Bias to [p]
[spʌp]~[spʌt]	-19.75
[stʌp]~[stʌt]	3.89

It can therefore be concluded that the response bias observed was not due to the lexical neighborhood differences between the tokens.

These regression analyses also show another possible interpretation of the results to be insufficient. It is possible that the results are caused by non-contiguous co-occurrence dependencies. Newport and Aslin (2004) have recently shown in an artificial language learning experiment that listeners can learn co-occurrence probabilities between non-adjacent segments. This begs the question of whether the results cannot be explained by such non-contiguous dependencies. Consider the K~P-condition as an example: In this condition a bias was found against [p] on the [spap]~[spak]-continuum. Could this not be attributed to the fact that [sp_k] has a higher probability than [sp_p]? I calculated transitional probabilities only between contiguous segments. The question is therefore whether transitional probabilities should not also be calculated for non-contiguous segments – i.e. abstracting away from the intervening vowels.¹⁴

Non-contiguous dependencies such as these are accounted for in calculation of the lexical neighborhood densities. All words of the form [sp_k] count as lexical neighbors to the [spak]-endpoint. The fact that words of the form [sp_k] exist while words of the form [sp_p] does not exist therefore results in the [spak]-endpoint having a denser lexical neighborhood than the [spap]-endpoint. By showing that lexical neighborhood density contributes insignificantly toward explaining the response patterns in the data, we have therefore also shown that non-contiguous co-occurrence dependencies do not explain the data.

3.2 The relevance of the results

This paper had two primary goals: (i) to provide additional support for the claim that phonological grammar influences perception; and (ii) to provide additional support for the psychological reality of the OCP.

It was found in the experiment that, when presented with a token ambiguous between a percept that does and a percept that does not violate the OCP, listeners are more likely to perceive this token as the OCP-obeying percept. This shows that listeners use phonological grammar in speech processing. Since the tokens were equalized for transitional probability, the perceptual biases cannot be attributed to transitional probabilities. In §3.1 above it was shown that the differences between tokens in lexical neighborhood density do not contribute significantly to the biases observed. The biases can therefore not be explained in models of speech perception that allows only for lexical statistics as non-acoustic sources of information in perception – for instance NAM (Luce 1986, Luce and Pisoni 1998, Vitevitch and Luce 1998, Vitevitch and Luce 1999), TRACE (McClelland and Elman 1986), and MERGE (Norris et al. 2000). These models either allow no place for phonological grammar, or allow it only as a byproduct of the frequency statistics (a phonotactic constraint is simply equal to zero frequency).

As for the second goal of the paper, it was found that English listeners do rely on a constraint of similarity avoidance in perception. This provides additional support for the psychological reality of the OCP in English.¹⁵

Even though English allows words of the form [sTvT], a perceptual bias was found against [sTvT]-forms. This sheds light on how the OCP should be understood. It is not a constraint that is obeyed by all words of a language. Some actual words do violate this constraint. The perceptual bias against [sTvT]-forms can only be explained under this interpretation of the OCP.

This also corresponds to patterns observed in the lexicon of several languages. It is known for Arabic (Frisch et al. 2000), English, French and Latin (Berkley 2000), that words that contain identical or highly similar sounds are either absent or under-represented in the lexicon. These patterns are interpreted as the result of the OCP – even though words with identical or very similar sounds are allowed in these languages, there is a bias against incorporating words like these into the lexicon.

The perceptual bias against [sTvT]-forms in English also supports this interpretation of the OCP. Even though [sTvT]-strings do occur in the English lexicon, these forms violate the OCP and there is a bias against them. This implies that the OCP must also make reference to coronal place.

Notes

- ¹ Thanks to John Kingston, John McCarthy, the Phonology Group of the University of Massachusetts, the PaPI 2003 audience, the editors of this volume and two anonymous reviewers for helpful discussion of and feedback on this paper.
- ² I am collapsing the lexical bias effect found by Ganong (1980) with the lexical frequency effect found by Connine et al. (1993). Ganong found that when presented with a token ambiguous between a word and a non-word (for instance *kiss*~*giss*), listeners are biased to identify the percept as the word (i.e. as *kiss*). However, *giss* being a non-word has a lexical frequency of zero. It is therefore the case that *kiss* has a higher lexical frequency than *giss*.
- ³ Moreton motivates this as follows: Cross-linguistically consonant clusters with a larger sonority difference between the consonants in the cluster are preferred. The voiced stop [d] is closer in sonority to the liquid [l], than the voiced stop [b] is to the glide [w]. Also, English tolerates [voiced stop] [glide] sequences in the coronal [tw] and dorsal place [gw]. It can therefore be expected that this should also be possible at the labial place [bw].
- ⁴ For how these probabilities were computed, see §2.2.1.1 below.
- ⁵ This is part of a more general constraint. For instance, neither identical liquids nor stops differing only in voice are allowed (**sLvL* and **sPvB*) (Fudge 1969).
- ⁶ The form that the OCP takes here will have to take into account that the initial [s] is crucial – [PvP] and [KvK] are well-formed in English (cf. *pipe* and *cake*). However, how to formulate this OCP-constraint is not relevant for this paper. All that is relevant is that there is a restriction that rules out [sKvK] and [sPvP].
- ⁷ This database was “Americanized” before the calculations were done. The “Americanization” consisted primarily in changes in vowel quality and in the presence of post-vocalic /r/.
- ⁸ See §3.1 below for motivation that lexical neighborhood density and transitional probabilities between non-contiguous segments did not significantly influence the results of the experiment.
- ⁹ A more detailed description of the synthesis process is available from the author.
- ¹⁰ These six tokens were identified as follows: In a pre-test, a 20-step [hap]~[hak]-continuum was synthesized. The tokens on this continuum were identical to the tokens on the [skap]~[skak]-continuum and the [spap]~[spak]-continuum in the last 60% of the nucleus and the final stop. They differed only in that they started with the place neutral consonant [h-]. This continuum was presented to 20 subjects in a pre-test. Subjects had to identify the final consonants as either [p] or [k]. Those six tokens that surrounded the 50%-mark on this continuum were assumed to be the six most ambiguous tokens. The six tokens on [skap]~[skak]-continuum and the [spap]~[spak]-continuum that correspond to these six most

ambiguous tokens on the [hap]~[hak]-continuum were used in the experiment. The six most ambiguous tokens in the other two conditions were selected in a similar way.

A second pre-test with 20 subjects was run, this time replacing the initial [h-] with the relevant [sC]-clusters. This second pre-test confirmed that the tokens used on P~T-condition did indeed contain the cross-over point for this continuum. However, the tokens used on both the K~T- and K~P-conditions were a little too close to the [k]-endpoints. In the final experiment, the ambiguous tokens used for these two conditions were therefore selected a few steps closer to the [t]-endpoint and the [p]-endpoint respectively.

- ¹¹ Put differently, the tokens for the two continua were selected from the same contiguous section of the [ap]~[ak]-continuum. For instance, if the tokens from [spap]~[spak]-continuum happened to be tokens 8 through 13, then tokens 8 through 13 from the [skap]~[skak]-continuum were also used.
- ¹² See also Pitt and Samuel (1993) and Newman *et al* (1997) on this.
- ¹³ See §3.1 below for further motivation that the lexical neighborhood densities were not responsible for the biases found.
- ¹⁴ Thank you to an anonymous reviewer for pointing out this possible interpretation.
- ¹⁵ I have also shown in more recent work that these same co-occurrence restrictions influence well-formedness judgments and reaction times in lexical decision tasks (Coetzee, forthcoming). This provides additional evidence for the psychological reality of the OCP in English.

References

- Baayen, R., R. Piepenbrock, and L. Gulikers 1995 *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium.
- Berent, Iris, Daniel L. Everett, and Joseph Shimron 2001 Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology* 42:1-60.
- Berkley, Deborah Milam. 2000. Gradient Obligatory Contour Principle Effects. Ph.D. diss, North Western University.
- Coetzee, Andries W. Forthcoming. *What it Means to be Loser: Non-Optimal Candidates in Optimality Theory*. Ph.D. dissertation, University of Massachusetts.

- Connine, C.M., D. Titone, and J. Wang 1993 Auditory word recognition: extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:81-94.
- Cutler, Anne, Jacques Mehler, Dennis Norris, and J. Segui 1987 Phoneme identification and the lexicon. *Cognitive Psychology* 19:141-177.
- Davis, Stuart. 1991. Coronals and the phonotactics of nonadjacent consonants in English. In *The Special Status of Coronals: Internal and External Evidence*, Carole Paradis and Jean-Francois Prunet (eds.), 49-60. San Diego: Academic Press.
- Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler 1999 Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568-1578.
- Frisch, Stefan. 2000. Temporally organized lexical representations as phonological units. In *Language Acquisition and the Lexicon: Papers in Laboratory Phonology V*, Michael Broe and Janet Pierrehumbert (eds.), 283-298. Cambridge: Cambridge University Press.
- Frisch, Stefan A., Michael B. Broe, and Janet B. Pierrehumbert 2000 Similarity Avoidance and the OCP. Ms. University of Michigan, Ann Arbor.
- Frisch, Stefan A., and Bushra Adnan Zawaydeh 2001 The psychological reality of OCP-Place in Arabic. *Language* 77:91-106.
- Fudge, Erik. 1969 Syllables. *Journal of Linguistics* 5:253-287.
- Ganong, W.F. 1980 Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6:110-125.
- Greenberg, Joseph 1950. The patterning of root morphemes in Semitic. *Word* 6:162-181.
- Leben, Will. 1978 The representation of tone. In *Tone: A Linguistic Survey*, Victoria Fromkin (ed.), 177-219. New York: Academic Press.
- Luce, Paul A. 1986 Neighborhoods of words in the mental lexicon, MS Department of Psychology, Indiana University.
- Luce, Paul A., and David B. Pisoni 1998 Recognizing spoken words: the neighborhood activation model. *Ear and Hearing* 19:1-35.
- Massaro, Dominic W., and Michael M. Cohen 1983 Phonological context in speech perception. *Perception and Psychophysics* 34:338-348.

- McCarthy, John J. 1985 Features and tiers: Semitic root structure constraints revisited.
- 1986 OCP Effects: Gemination and antigemination. *Linguistic Inquiry* 17:207-263.
- McClelland, J.L., and Jeffrey L. Elman 1986 Interactive processes in speech perception: The TRACE model. In *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland (eds.), 58-121. Cambridge: MIT Press.
- Moreton, Elliot. 2002 Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:55-71.
- Newman, Rochelle S., James R. Sawusch, and Paul A. Luce 1997 Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology* 23:873-889.
- Newport, Elissa L. and Richard N. Aslin. 2004. Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48:127-162.
- Norris, D., J.M. McQueen, and A. Cutler 2000 Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23:299-370.
- Pitt, M.A., and A.G. Samuel 1993 An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance* 19:699-725.
- Pitt, Mark. A. 1998 Phonological processes and the perception of phonologically illegal consonant clusters. *Perception and Psychophysics* 60:941-951.
- Vitevitch, M.S., and Paul A. Luce 1998 When words compete: levels of processing in spoken word recognition. *Psychological Science* 9:325-329.
- 1999 Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40:374-408.