Natalie Boll-Avetisyan | René Kager

OCP-PLACE in Speech Segmentation

Suggested citation referring to the original publication:
Language and Speech 57 (2014) 3 pp. 394–421
DOI http://dx.doi.org/10.1177/0023830913508074
ISSN (print) 0023-8309
ISSN (online) 1756-6053

Article

Language
and Speech

Language *and* Speech
2014, Vol. 57(3) 394–421

DOI: 10.1177/0023830913508074
las.sagepub.com
SAGE

OCP-PLACE in Speech Segmentation

Natalie Boll-Avetisyan
Utrecht University, The Netherlands; University of Potsdam, Germany

René Kager
Utrecht University, The Netherlands

Abstract
OCP-PLACE, a cross-linguistically well-attested constraint against pairs of consonants with shared [place], is psychologically real. Studies have shown that the processing of words violating OCP-PLACE is inhibited. Functionalists assume that OCP arises as a consequence of low-level perception: a consonant following another with the same [place] cannot be faithfully perceived as an independent unit. If functionalist theories were correct, then lexical access would be inhibited if two homorganic consonants conjoin at word boundaries—a problem that can only be solved with lexical feedback.

Here, we experimentally challenge the functional account by showing that OCP-PLACE can be used as a speech segmentation cue during pre-lexical processing without lexical feedback, and that the use relates to distributions in the input.

In Experiment 1, native listeners of Dutch located word boundaries between two labials when segmenting an artificial language. This indicates a use of OCP-LABIAL as a segmentation cue, implying a full perception of both labials. Experiment 2 shows that segmentation performance cannot solely be explained by well-formedness intuitions. Experiment 3 shows that knowledge of OCP-PLACE depends on language-specific input: in Dutch, co-occurrences of labials are under-represented, but co-occurrences of coronals are not. Accordingly, Dutch listeners fail to use OCP-CORONAL for segmentation.

Keywords
Artificial language learning, OCP-PLACE, phonotactics, speech segmentation, pre-lexical processing

---

Corresponding author:
Natalie Boll-Avetisyan, University of Potsdam, Cognitive Sciences, Department Linguistics, Karl-Liebknecht-Straße 24-25, Potsdam, 14476, Germany.
Email: nboll@uni-potsdam.de

I Introduction

Numerous languages restrict their lexicons such that they prohibit similarity between co-occurring segments, typically within a morpheme. This has been first described by Greenberg for Arabic (Greenberg, 1950): in Arabic, verbs are derived from consonantal roots. For example, the root /k t b/ can be derived to *katab-a* 'he wrote', *kutib-a* 'it was written', and *kuttib-a* 'he was made to write' (example taken from Frisch et al., 2004). Consonantal root morphemes, with rare exceptions, do not contain two or more labial consonants. That is, roots such as **fbm*, **bfk* or **kbm* are systematically unattested. This phonotactic restriction falls into a class of constraints that are part of the so-called *Obligatory Contour Principle* (OCP) requiring segments to be featurally non-identical (i.e. disharmony constraints). Originally, the OCP was formulated to account for similarity avoidance between tones in West-African languages (Leben, 1980). Later, McCarthy (1986) extended the OCP to the notion of OCP-PLACE to formally account for similarity avoidance between consonants with shared [place].

Currently, there are very few accounts of how OCP-PLACE might be represented in listeners' minds and how this affects processing. This topic, however, is an interesting one as—taking a computational perspective—using OCP-PLACE for processing should not be a trivial task; the effect of OCP-PLACE crosses intervening vowels, constraining the co-occurrence of pairs of consonants that are phonetically non-adjacent (although consonants are represented as adjacent on the "consonant tier", an abstract representation that omits vowels; McCarthy, 1986). In spite of this, strong OCP-PLACE effects are found across many languages; for example, they are found in Semitic languages such as Arabic (e.g., Frisch et al., 2004; Greenberg, 1950; McCarthy, 1985) and Hebrew (e.g., Berent & Shimron, 1997; McCarthy, 1985). Furthermore, many genetically and geographically unrelated languages display non-categorical, gradient effects of OCP-PLACE (e.g., English: Berkley, 1994; Muna: Coetzee & Pater, 2008; Niger-Congo languages: Pozdniakov and Segerer, 2007; Dutch: Shatzman and Kager, 2007). In the latter type of language, pairs of consonants with shared place features are attested, such as /pVm/ in English *spam* or /mVb/ in *mob*. However, such pairs occur significantly less often than expected if non-adjacent consonants co-occurred at random.

Over the past few years, a number of experimental studies have shown that listeners have unconscious knowledge of OCP-PLACE. For example, in studies with native listeners of Hebrew (Berent & Shimron, 1997), Arabic (Frisch & Zawaydeh, 2001), and English (Coetzee, 2008), nonwords that violate OCP-PLACE are judged to be less well-formed than nonwords that do not. Furthermore, OCP-PLACE affects performance in lexical decision tasks: nonwords that violate OCP-PLACE are rejected faster than nonwords composed of consonants that do not share [place] by native listeners of Hebrew (Berent et al., 2001) and Dutch (Shatzman & Kager, 2007). Also, OCP-PLACE biases phoneme identification such that forms containing sequences that violate OCP-PLACE tend to be perceived as sequences of non-harmonic consonants by native listeners of English (Coetzee, 2005). These studies suggest a role for OCP-PLACE in processing.

The typological preponderance of OCP-PLACE and the fact that listeners are influenced by the constraint in processing tasks has led to functional theories of a perceptual bias against consonants with shared [place]. We will first introduce these accounts, and then argue against the functionalist proposal by putting the role for OCP-PLACE in a new perspective, that is, speech segmentation, and provide experimental support for our claim.

The functionalist accounts are based on typological studies (e.g., Mayer et al., 2010; Pozdniakov and Segerer, 2007) that have shown that it is difficult to find a language that is not restricted by similar place avoidance. Such typological observations have lead researchers to consider

OCP-Place to be a (statistical) universal (e.g., McCarthy, 1986; Pozdniakov & Segerer, 2007). Most phonologists assume that OCP is a restriction that holds at the level of underlying forms (for a surface-based implementation of OCP-Place in an Optimality Theory framework, see Gafos, 2003). Underlyingly, phonological features are assumed to be represented on different autosegmental tiers, i.e. a vowel and a consonant tier. These tier representations are further subdivided into, for instance, a place feature tier and a laryngeal tier. The constraint OCP-Place disfavors repetitions of features on the place tier (McCarthy, 1986). If a surface form contains two adjacent consonants with a shared feature, OCP-Place makes sure that their representations in the underlying form are merged, that is, one feature specification is doubly linked to two consonant positions rather than there being two separate specifications (e.g., Goldsmith, 1979; Leben, 1980).

Autosegmental phonologists like Goldsmith or Leben do not attempt to account for a role of OCP in perception, being merely concerned with the representational merger. Boersma (1998, 2000, 2009) elaborates this idea in a speech perception context and proposes that OCP results in a perceptual merger. While traditional phonological theory assumes that OCP affects the mapping of surface and underlying form, Boersma proposes that OCP affects perception at a lower level, specifically the output of pre-lexical perception—the mapping of an auditory form onto a surface form. If a listener perceives an auditory form that has two adjacent phonemes with shared [place], then OCP-Place—disfavoring repetitions of features in surface forms—has the perceptual effect of merging these phonemes into one higher-order unit in their surface form: “A sequence of two acoustic cues *cue*1 and *cue*2 is perceived as a single value *x* on the perceptual tier *f*, despite the presence of some intervening material *m*” (Boersma, 2000: 23). In some way, Boersma’s view of OCP as a perceptual merger comes close to the assumption of the *gestalt* principle of similarity that is used to explain human preferences to group similar elements (e.g., Thorpe & Trehub, 1989; Wertheimer, 1923). In his view, it is crucial that this merger affects pre-lexical perception: the purpose of the output of pre-lexical perception is to facilitate lexical access, hence it will be favorable to avoid attempts of mapping two homorganic units onto an underlying form during lexical access, if pairs of homorganic units are not allowed in the lexicon (Boersma, 2000: 14; 2009: 24).

Frisch (2004) opens yet another functional perspective on OCP. He assumes that similarity avoidance creates a functional advantage during pre-lexical perception in speech processing. In pre-lexical perception, speech sounds are serially encoded. Once a phoneme is encoded, its co-encoded features are immediately inhibited to speed up the encoding of the following phoneme. Frisch proposes that the second of two consonants with shared features has a high chance of not being correctly encoded because of an inhibition of relevant features (Stemberger et al., 1985; Stemberger and MacWhinney, 1986). Phoneme inhibition can be viewed as the speech processing instantiation of repetition blindness (Eriksen & Schutze, 1978; Kanwisher, 1987). This is useful for speeding up the lexical recognition process as words typically do not contain pairs of consonants with shared place (although there is also evidence against pre-lexical phoneme-level inhibition, cf. McMurray et al., 2009). Frisch explains the typological recurrence of OCP through the functional bias: as words containing sequences of homorganic consonants are more difficult to process, they will—in the course of language change—be replaced by words that do not contain OCP-violating sequences.

In sum, functionalist accounts of similarity avoidance assume that listeners do not faithfully and/or individually perceive both consonants in a pair of homorganic consonants. Frisch assumes that the perception of the second of two homorganic consonants is distorted, whereas Boersma assumes that homorganic consonants are merged into one percept.

In the current study, we want to look into a different role that OCP-Place may play in speech processing. Potentially, the requirement for consonants to disagree in [place] within a word may provide a useful cue for speech segmentation. In particular, we want to investigate whether OCP-Place has the potential of giving *pre-lexical* cues to segmentation *without* a need for lexical feedback. For listeners whose native language is restricted by OCP-Place, if one encounters a sequence of two consonants with shared [place] in continuous speech, it is likely that this sequence straddles at a word boundary (e.g., /pVm/ in *happy man*), as it is unlikely to reside within words. Several findings from previous segmentation studies suggest that potentially, OCP-Place could be a pre-lexical segmentation cue: First, troughs in probability can be interpreted as word boundaries (e.g., Mattys and Jusczyk, 2001; Saffran et al., 1996). Second, in artificial language learning experiments, statistical cues to non-adjacent dependencies can be picked up from continuous speech input and used as segmentation cues (e.g., Newport and Aslin, 2004; Peña et al., 2002). Third, it has been demonstrated that infants and adults locate word boundaries between pairs of disharmonic vowels, albeit only if their language is categorically restricted by vowel harmony (Turkish, but not French: Kabak et al., 2010; Finnish: Suomi et al., 1997; Turkish: Van Kampen et al., 2008; Finnish, but neither French nor Dutch: Vroomen et al., 1998). Hence, it is likely that the phonotactic constraint OCP-Place joins the sum of pre-lexical cues for segmentation such as lexical stress (e.g., Cutler & Norris, 1988), coarticulation (e.g., Johnson & Jusczyk, 2001; Mattys, 2004), probabilistic phonotactics (e.g., Mattys & Jusczyk, 2001) and segment duration (e.g., Shatzman & McQueen, 2006).

If we take the functionalist theories of OCP described above at face value, then it should be difficult for OCP-Place to serve as a cue for speech segmentation without lexical feedback. Using OCP for segmentation requires that humans perceptually *separate* two consecutive consonants with shared [place]. Crucially, perceptual segregation can only work if both homorganic consonants are faithfully perceived as individual elements. Frisch's assumption about an inhibition of the second of two homorganic consonants is problematic for the following reason: as for words, continuous speech also needs to be encoded serially. Hence, for example, in the sequence /pVm/ straddling a word boundary in *happy man*, the word onset /m/ would not be properly encoded due to inhibition of the labial feature after activating the preceding word offset /p/. Hence, Frisch's theory predicts that word recognition of *man* after *happy* would be slowed down. This, however, would be harmful to word recognition, as it would slow down the processing of a word onset; yet word onsets are usually most important for lexical access. In any case, the prediction is not confirmed by the results of a word-spotting experiment carried out in our lab (Kager and Shatzman, forthcoming). Here, listeners were actually faster at spotting labial-initial words if they were preceded by a labial rather than when they were preceded by a coronal consonant. Hence, Frisch's theory of an inhibited encoding of the second of two similar consonants can only be correct if inhibition of the place feature stops at the end of a word or a morpheme. Yet it is not clear how the inhibitory process might stop at word boundaries when listening to continuous speech. Moreover, if pre-lexical perception of homorganic consonant pairs were generally inhibited, it leaves unexplained the results of lexical decision experiments, in which nonwords that violate OCP-Place are rejected *faster* than nonwords composed of consonants that do not share [place] (Berent et al. 2001; Kager and Shatzman, 2007). This raises the question of how the functional effect of OCP to inhibit the encoding of a second consonant in a homorganic pair should be stopped at word boundaries. It necessarily would have to involve lexical knowledge that the homorganic consonants are tautomorphemic.

Boersma's assumption of a merger that creates a single percept of two adjacent homorganic consonants is problematic, if these are part of two separate words. This perceptual merger would inhibit segmentation and, consecutively, lexical access, since the merged percept would have to be undone by reallocating the merged feature to two separate lexical units. Boersma is aware of this problem, when discussing the case of nasal spreading in Guaraní, for which he assumes a merged surface representation:

> … on the prelexical level listeners do not hear word boundaries. The question then is: …, e.g. would listeners interpret [tũpã] as having two nasals if there were a word boundary between [tũ] and [pã]? They could indeed do this if the lexicon is allowed to pass on information about word boundaries to the lower prelexical level. (Boersma, 2009: 26)

Hence, Boersma proposes that in these cases, feedback from the lexical level will be at play. Alternatively, in an earlier paper he proposes that listeners may learn a pre-lexical constraint based on lexical knowledge:

> Add a morpheme boundary between any two identical adjacent consonants, even if they differ in voicing […]. This step is reasonable because no English word-like morpheme contains a sequence of identical adjacent consonants or of adjacent consonants that differ only in voicing. (2000: 6)

However, it is not clear how such a pre-lexical constraint interacts with the pre-lexical merger, particularly in cases in which effects of OCP are just probabilistic (as, e.g., in English in *happy man*).

In sum, in both functionalist accounts, OCP takes pre-lexical effects on the phonological encoding that may be advantageous for lexical access but disadvantageous for speech segmentation. It is intrinsic to Boersma's proposal of a perceptual merger that OCP unifies two consonants with shared place rather than that it segregates them. Frisch's proposal of a slowed-down identification of the second of two similar consonants has the consequence that at word boundaries, the recognition of the word onset would be slowed down. Although both Boersma and Frisch suggest that OCP is active at a pre-lexical level, its use as a segmentation cue would have to be triggered by lexical feedback: in continuous speech, a merger/inhibitory effect of OCP can only be overcome if lexical knowledge is drawn upon, such that a listener recognizes two homorganic consonants (such as /pVm/ in *happy man*) as belonging to two separate words.

The conclusion that the merger/inhibitory effect of a phonotactic constraint OCP-PLACE would have to be "deactivated" at word boundaries through lexical feedback in order to allow for successful lexical access in speech processing is at odds with studies on segmentation that suggest that phonotactic knowledge can be used as a cue to word boundaries independently of lexical knowledge for example in pre-lexical infants (e.g., Mattys and Jusczyk, 2001). The literature on speech segmentation suggests that pre-lexical cues for speech segmentation can be induced through exposure to continuous speech streams in which co-occurrences of linguistic units vary in transitional probability (e.g., Saffran et al., 1996), and phonotactic cues, i.e., co-occurrences of phonemes with a low transitional probability, have proven to be a classical example of distributional cues to segmentation (e.g., Mattys and Jusczyk, 2001).

For this reason, we hypothesize that OCP-PLACE is a phonotactic constraint influencing speech segmentation without involving lexical organization. Knowledge of this speech segmentation cue may reflect pure knowledge of an organization of phoneme co-occurrences in continuous speech. An assumption of a perceptual bias against co-occurring phonemes with shared [place] would lapse.

Consequently, we carried out experiments testing Dutch participants on their use of OCP-PLACE in speech segmentation. Dutch, as will be described in Section 2, is gradiently restricted by OCP-PLACE, and moreover, the restriction is not balanced between natural classes: consonant distributions in Dutch give more reason for assuming an effect of OCP-LABIAL than an effect of OCP-CORONAL, as pairs of labials are strongly under-represented, whereas pairs of coronals are not. This asymmetry makes Dutch particularly well-suited for testing the hypothesis that knowledge of OCP-PLACE reflects input distributions.

For our experiments, we used the artificial language learning paradigm. For testing our hypothesis, it was essential that participants would not be able to access their lexicons, so we needed stimulus material that was well controlled for lexical statistics. Artificial languages are ideal for this purpose, as information can be reduced to a minimum such that the impact of segmentation cues can be tested in isolation (i.e., lexical or phonological cues to word boundaries can be maximally excluded). In this task, participants are first familiarized with a continuous stream of an artificial language, a highly reduced miniature language. Subsequently, segmentation performance is assessed in a test phase by means of a two-alternatives-forced-choice task. In our study, the only cue for segmenting the artificial language stream was OCP-Place. The learning and segmentation of artificial languages can be affected by the transfer of native language phonological knowledge (Finn and Hudson Kam, 2008; Onnis et al., 2005; Vroomen et al., 1998). Hence, we predict that if listeners use OCP-Place for speech segmentation in their native language, they will transfer the constraint to segmenting an artificial language.

The paper is structured as follows: In the next section, non-adjacent consonant distributions in Dutch will be analyzed. In the subsequent sections, three experiments will be presented. Experiment 1 assesses whether Dutch listeners use OCP-Labial as a segmentation cue in an artificial language. Experiment 2 serves as a control for whether the effects of Experiment 1 were simply an effect of the participants' intuitions about the well-formedness of Dutch word forms. By testing the use of OCP-Coronal on segmentation, Experiment 3 was carried out to test whether the results of Experiment 1 are due to an effect of OCP-Labial or a more general constraint OCP-Place, and simultaneously controlled for the possibility that a general cognitive preference for similarity between segments in edges or avoidance of similarity in adjacency was at play.

2 OCP-Place in Dutch

2.1 Method

In order to test our hypothesis, we used Dutch. The Dutch lexicon has been found to be gradiently restricted by OCP-Place (Kager and Shatzman, 2007). Furthermore, Dutch listeners possess unconscious knowledge about this restriction, as revealed in experiments with lexical decision tasks (Shatzman and Kager, 2007). We calculated the probabilities of C_1VC_2 sequences in which $C_1$and C_2 share [place]. We used two different databases: First, we extracted a lexicon of 8305 Dutch monomorphemic stems from the CELEX lexical database (Baayen et al., 1995). All CVC sequences, independent of whether they occurred word-initially, -medially, or -finally, were included. Each occurrence of a C in a database was counted as C_1, no matter whether it was counted as a C_2 before or not. That is, for instance, from the Dutch word *begin* /bəxɪn/ 'begin', we extracted two CVC sequences (i.e., /bəx/ and /xɪn/) and not just one. By this, 2700 CVC types (11,092 tokens) were extracted. After filtering to CC, 403 types were left, on which our calculations were based. Second, we used CGN, a corpus of phonetically transcribed spoken Dutch (Goddijn and Binnenpoorte, 2003). Here, we did not only calculate the probabilities of C_1VC_2 sequences that occurred within words, but also those that occurred across word boundaries. So, for example, for the sequence *de naam* /də na:m/ 'the name' we would extract two CVC sequences: /dən/ and /na:m/. Probability "troughs" in the speech stream are useful cues for speech segmentation (e.g., Saffran et al., 1996). Hence, even though OCP-Place is viewed as a constraint that targets morphemes, we include counts of between word probabilities, as there is the possibility that OCP-Place reflects knowledge of phoneme distributions in the input. A total of 6312 CVC types (11,817,772 tokens) were extracted from CGN. Calculations were based on 516 CC types that were left after filtering.

As a measure, we used the Observed/Expected (*O/E*) ratio (Pierrehumbert, 1993). This measure compares the observed counts of consonant pairs to the counts expected if consonants combined at random (Frisch et al., 2004). O stands for the number of C_1VC_2 sequences in the corpus (or lexicon). It is divided by E, 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, which is multiplied by the total number of CVC sequence tokens:

$$O/E = N(C_1VC_2) \,/\, p(C_1) * p(C_2) * N(CVC)$$

$N(C_1VC_2)$: the number of C_1VC_2 sequences; $N(CVC)$: the number of CVC sequences
$p(C_1)$, $p(C_2)$: the probability of C_1 and C_2 calculated as proportions of consonants
$O/E > 1$: sequence is over-represented; $O/E < 1$: sequence is under-represented

2.2 Results and discussion

The results of our calculations show that in Dutch, CVC sequences in which both Cs share the feature [place] are under-represented both in the lexicon as well as in continuous speech. This holds in particular for sequences of labials (the phonemes /p, b, f, v, w, m/; hereafter, P) and dorsals (the phonemes /k, g, x, ŋ/; hereafter, K). The distribution of coronals (the phonemes /t, d, s, z, ʃ, ʒ, r, l, n/; hereafter, T) is somewhat less restricted (see Figure 1).

The distributions of Cs with shared [place] in CVC justify the assumption that OCP-PLACE holds gradiently in the lexicon. More specifically, the Dutch lexicon is restricted by the gradient phonotactic constraints OCP-LABIAL (CELEX: $O/E = 0.45$, CGN: $O/E = 0.58$) and OCP-DORSAL (CELEX: $O/E = 0.58$, CGN: $O/E = 0.69$), and possibly also OCP-CORONAL, if phonotactics were learned from the lexicon (CELEX: $O/E = 0.77$), but not when phonotactics are acquired from continuous speech (CGN: $O/E = 1.24$). Since OCP-LABIAL shows the most significant under-representation, we tested our assumptions by means of this constraint.

For Experiment 1, we created an artificial language in which coronal-initial CV syllables (hereafter, T) were always followed by two labial-initial CV syllables (hereafter, P). The syllables were concatenated into a speech stream without pauses (…$P_1P_2TP_1P_2TP_1P_2T$…). We predict that if Dutch participants have knowledge that PP sequences are under-represented in their native language, and are able to apply this knowledge when segmenting speech, they should insert a boundary between P_1P_2. Hence, in a test phase, PTP words should be preferred over PPT and TPP words. Furthermore, there should be no preference of PPT over TPP or vice versa, as both segmentations violate OCP-LABIAL and, hence, should not be segmented from the stream.

3 Experiment 1

3.1 Method

3.1.1 Participants. Fifty-four students, (12 male, 42 female, $M_{age} = 23.67$ years, age range: 18–68 years) from Utrecht University participated in Experiment 1a, and 39 students, (three male, 36 female, $M_{age} = 23.03$ years, age range: 18–50 years) participated in Experiment 1b. All participants reported normal hearing, and Dutch to be their only native language. They were compensated for their efforts.

3.1.2 Materials. Familiarization stimuli: Two versions of an artificial language of a …$P_1P_2TP_1P_2T$-P_1P_2T… structure were created: Language A (Experiment 1a) and Language B (Experiment 1b). They consisted of nine syllables that were assigned to one of three fixed slots (see Table 1).

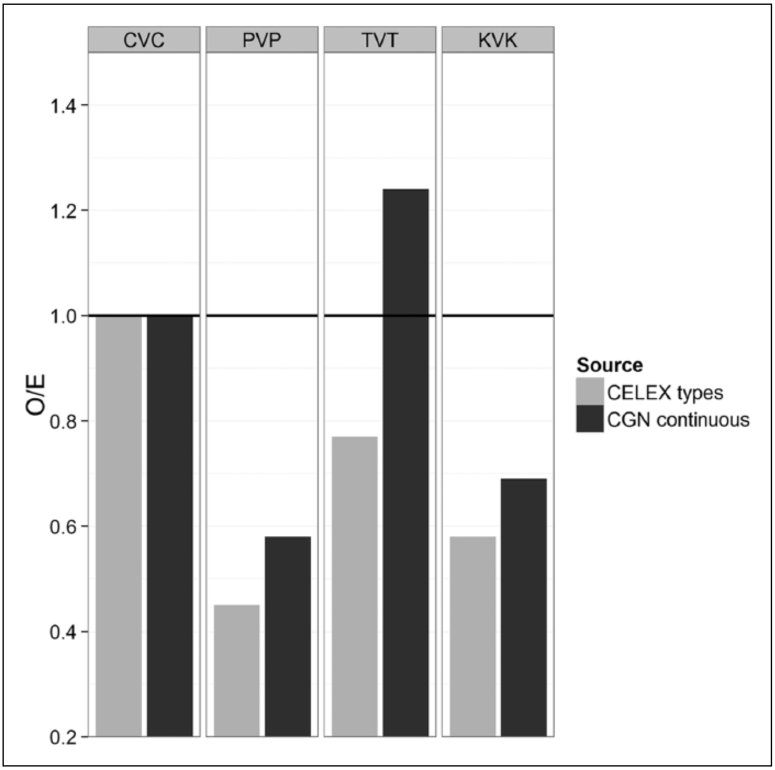

Figure 1. *O/E* values for consonant pairs sharing place of articulation (P = labial, K = dorsal, T = coronal) in CVC sequences. Calculations are based on counts in CELEX lemma types (CELEX types) and an unsegmented corpus of spoken Dutch (CGN continuous).

Table 1. Syllable inventory of the artificial languages used in Experiments 1 and 3.

Experiment	A	B	A
1a	P_1	T	P_2
	/po/	/tu/	/pa/
	/be/	/do/	/bi/
	/ma/	/ne/	/mo/
1b	P_1	T	P_2
	/pe/	/tu/	/po/
	/bo/	/de/	/be/
	/mi/	/na/	/ma/
3	T_1	P	T_2
	/ta/	/po/	/tu/
	/di/	/be/	/do/
	/no/	/ma/	/ne/

One of the three P_1 syllables was always followed by one of the three T syllables, which was followed by one of the three P_2 syllables, which was followed by one of the three P_1 syllables and so forth. Syllables were concatenated into a speech stream with no pauses (e.g., …/pamatumomatubibetu/…). Transitional probabilities between both adjacent and non-adjacent syllables were always 0.33, so that the language itself did not contain a distributional cue for segmentation. Consequently, all 3^3=27 possible words occurred in the language.

The syllable inventory for the two languages was selected on the basis of controls for lexical, low-level phonotactics, and other phonological factors to assure that segmentation would be only cued by OCP-LABIAL, independent of the intervening vowels. To do so, we controlled for potential cues for segmentation that may resemble cues from OCP-LABIAL.

The consonants /p/, /b/, and /m/ were selected to represent P, and /t/, /d/ and /n/ to represent T. By selecting three consonants of different manner (voiceless and voiced plosive, and nasal) for each consonant class (P and T), we avoided any confounding effects of manner on segmentation. We did not include any fricatives (e.g., /f/, /v/, /z/, /s/), as voiced and voiceless realizations of fricatives are phonetically highly variant (and hence, perceptually confusable) in Dutch. Sequences of perceptually similar syllables were avoided (e.g., /mu-nu/, /bi-pi/). Furthermore, sequences that are real words (e.g., /bebi/ 'baby') that might affect segmentation were avoided. The vowels were selected such that the language was controlled for cues from positional syllable frequencies which might give cues to word boundaries (e.g., a certain syllable might be infrequent in word-initial but frequent in word-medial position). Hence, based on word types in the CELEX lexical database (Baayen et al., 1995), we selected syllables for the sets (T, P_1, P_2) such that their frequency would be evenly distributed in utterance-initial, -medial, and -final position (see Appendix A). Language A was, moreover, controlled for the probability of identical vowel sequences between parsings (i.e., /podoP_2/, /domoP_1/, /pamaT/).

For Language A, this resulted in the syllable inventory P_1 = {/po, be, ma/}, T = {/tu, do, ne/} and P_2 = {/pa, bi, mo/}. Language B was a minimal variant of Language A, the crucial difference being that syllables in word-initial position P_1 were now assigned to word-final position P_2 (see Table 1). A new set of syllables was selected for T and P_2 in order to meet the same controls that we applied to Language A (P_1 = {/pe, bo, mi/}, T = {/tu, de, na/}, and P_2 = {/po, be, ma/}, see Appendix A).[1]

The language was synthesized with a male voice "nl2" based on Dutch biphones, provided by FLUENCY using the MBROLA speech synthesizer (Dutoit et al., 1996). The synthesizer produced the stream with a monotone pitch (F0 = 100 Hz) with an average syllable duration of 232 ms. In order to prevent the endpoints (utterance boundaries) from giving a cue for segmentation, intensity faded in for the first five seconds and out for the last five seconds. The natural relationships in phoneme durations (voiced plosives being shorter than unvoiced plosives etc.; Waals, 1999) were maintained.

The artificial language stream lasted for 10 minutes, and consisted of 2700 syllables. If it is parsed into tri-syllabic units, the speech stream can be parsed into 900 PTP, 900 PPT and 900 TPP tokens. The artificial speech stream was pseudo-randomly organized in such a way that no tri-syllabic string occurred twice in succession.

Test stimuli: For Language A, we synthesized three types of stimuli: PTP, PPT, and TPP items for three different conditions. Condition 1 contrasted PTP (e.g., /ponebi/) and PPT (e.g., /pamado/) items, Condition 2 PTP (e.g., /benemo/) and TPP (e.g., /domoma/) items, and Condition 3 PPT (e.g., /mobedo/) and TPP (e.g., /nebibe/) items. For Language B, we had two instead of three different test phases: Condition 1 (PTP versus PPT) and 2 (PTP versus TPP). We dropped PPT versus TPP, as this comparison was less essential for testing our predictions.

The test items were balanced for phonological and lexical factors that potentially affect decisions: First, we controlled for manner of articulation. That is, within one condition (e.g., in Condition 1) each test item of one type (here: PTP) received a counterpart (here: PPT) that contained a nasal, a voiced and an unvoiced plosive in the same position. For illustration, the PTP item /ponebi/ received the PPT item /pamado/ as a counterpart, as both start with an unvoiced plosive, followed by a nasal and end with a syllable that starts with a voiced plosive (see Appendix B and C). Second, we controlled for the lexical factors cohort density and lexical neighborhood density. Cohort density was defined as the sum of the logged frequencies of the set of words in CELEX starting with the same three phonemes, whereas lexical neighborhood density was defined as the sum of the logged frequencies of the set of words in CELEX that differed from the test item by a single segmental change, for example deletion, addition, permutation, or alteration (Luce, 1986).

The best controls for cohort density were guaranteed using 12 test items per comparison. So, Condition 1, for example, used 12 PTP and 12 PPT items. The cohort densities for the test items can be viewed in Appendix D. Lexical neighborhood density zero for all test items. The test phase consisted of 48 trials in total, of which 36 were real test pairs (e.g., PTP versus PPT), and 12 filler pairs with two items from the same class (e.g., PTP versus PTP). Each item occurred four times in a different test pair, thrice in a test trial pair, and once in a filler pair with a test item of the same type. The test pairs were combined at random with the prerequisites that two items only made a pair once, and exactly half of the test pairs started with one type of stimulus.

3.1.3 Apparatus and procedure. The procedure was borrowed from Peña et al. (2002). Each participant was tested individually in a sound-attenuated booth. Stimuli were presented over headphones. To acquaint participants with the two-alternatives-forced-choice task of the test phase, they first entered a pre-familiarization phase, in which they were asked to find the syllable /so/ in a pair of CV syllables. Subsequently, participants entered the familiarization phase. They were informed that they would listen to an artificial language for the next 10 minutes and were instructed to listen carefully, because they would later be tested on their knowledge about the words of this language. After the familiarization phase, participants were tested in a two-alternative-forced-choice task, in which they had to indicate via a mouse-click on a button on the screen whether the first or the second of two items was more likely to be a word from the language they just heard. Test pairs were presented with an inter-stimulus interval of 500 ms. Using a between-subjects-design, participants were randomly distributed to one out of three conditions (Language A: Condition 1: N = 18; Condition 2: N = 18; Condition 3: N = 18; Language B: Condition 1: N = 20; Condition 2: N = 19). Within a condition, all participants were tested on the same list of test pairs. The order of the presentation of the items within a test pair (e.g., 1: /ponebi/, 2: /tupabe/ or 1: /tupabe/, 2: /ponebi/) was counterbalanced between participants.

3.1.4 Data processing and analysis. The filler pairs were excluded from the analysis. One outlier per condition was excluded from each of Experiment 1a and 1b. The dependent variable was the categorical response (“well-formed” versus “ill-formed”).[2] We analyzed the data in “R” (R Core Team, 2012) using the package “lme4” (Bates et al., 2012). As the data is binomially distributed, a maximal generalized linear mixed model (Jäger, 2008) with fixed and random effects was calculated.

The fixed factors were “condition” (three levels: Conditions 1, 2 and 3), “trial number” (continuous) and “identity” (two levels: strict-identity versus near-identity). The latter was introduced to control for the possibility that response preferences for one item type over the other in a given test pair might be influenced by the presence of an item in which two adjacent Ps were identical (e.g., /papotu/).

Two random factors—one for items and another for participants—were added to account for the variability between individual participants and individual items. As some of the items could occur across conditions, a random slope for “condition” was added to the random factor “items”. Given

Table 2. Tests of response preferences in Experiment 1 for each condition against chance. Coefficients (β) and their standard errors (*SE*) are logit transformations.

Estimates of fixed effects

Parameters	β	*SE*	*z*	*p*
Condition 1 (PTP vs. PPT)	0.36	0.14	2.48	< 0.05, two-tailed
Condition 2 (PTP vs. TPP)	0.28	0.14	2.01	< 0.05, two-tailed
Condition 3 (PPT vs. TPP)	0.27	0.17	1.62	> 0.1, n.s.
Trial number	–0.004	0.004	–1.04	> 0.1, n.s.

Table 3. Results of the comparisons between conditions in Experiment 1. Coefficients (β) and their standard errors (*SE*) are logit transformations.

Estimates of fixed effects

Parameters	β	*SE*	*z*	*p*
Intercept	0.27	0.11	2.41	< 0.05, two-tailed
Condition 1 vs. Condition 2	0.13	0.13	1.01	> 0.1, n.s.
Condition 1 + 2 vs. Condition 3	–0.14	0.08	–1.72	< 0.05, one-tailed
Trial number	–0.002	0.003	–0.67	> 0.5, n.s.
Condition 1 vs. Condition 2 * Trial number	–0.007	0.004	–1.69	> 0.1, n.s.
Condition 1 + 2 vs. Condition 3 * Trial number	0.005	0.003	2.01	< 0.05, two-tailed

the between-subjects design, no random slope for “condition” was entered for the random factor “participants.” Furthermore, random slopes for “trial number” were entered for both random factors, as the order of trial pairs was the same for all participants.

Two tests were carried out: First, preferences in each condition were compared to chance by setting the intercept at zero. Second, preferences between conditions were compared in a helmert contrast, which is an orthogonal contrast that in a first step compares the mean of Condition 1 with the mean of Condition 2, and in a second step, the mean of Conditions 1 and 2 with the mean of Condition 3.

For the subsequent analysis, the data from Experiments 1a and 1b were pooled, as there were no significant differences in their outcomes (see Appendix E). As “identity” neither had an effect on response preferences nor contributed to the model by significantly accounting for more variance in the data, the factor was not included in the consecutive analysis. In the model, response preferences were logit transformed, but for the ease of interpretability, we give back-transformations to the mean percentages of response preferences in the description of the results.

3.2 Results and discussion

3.2.1 Results. Participants significantly preferred PTP over PPT items in Condition 1 (58% for PTP) and over TPP items in Condition 2 (56% for PTP). In Condition 3, participants gave slightly more answers for PPT (55%) than for TPP, but this was not a significant preference. The results of the tests against chance are given in Table 2.

Comparisons between conditions showed that there was no difference in responses between Conditions 1 and 2. Responses to Condition 3, however, differed significantly (one-tailed) from

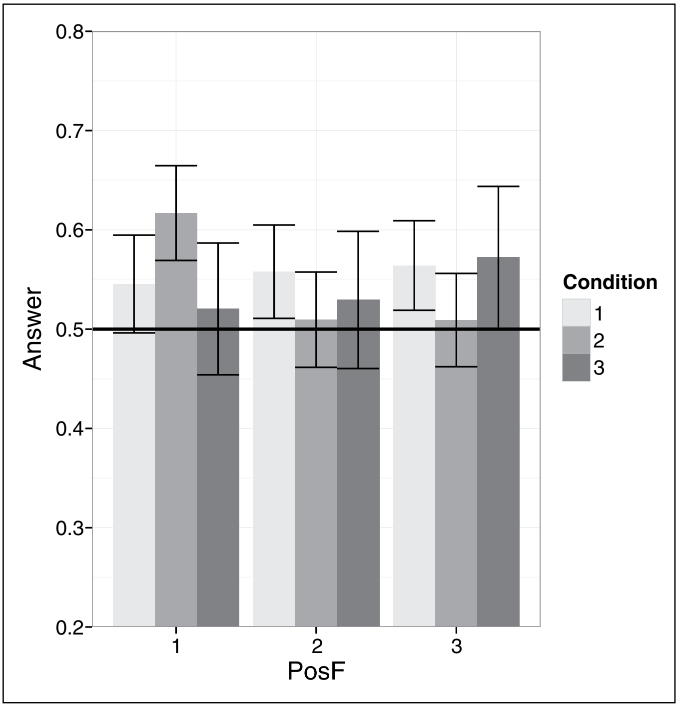

Figure 2. Average segmentation preferences and their standard errors in Experiment 1. Horizontal axis: 1 = test pairs 1–12; 2 = test pairs 13–24; 3 = test pairs 25–36. Bars are clustered by Condition: 1 = PTP > PPT, 2 = PTP > TPP, 3 = PPT > TPP.

Conditions 1 and 2 (see Table 3 for the fixed effects and Appendix F for the random effects). Furthermore, as can be seen in Figure 2, trial number interacted significantly with condition: for Condition 3, the proportion of response preferences for the PPT over TPP items increased throughout the test phase. In Conditions 1 and 2, this effect was reversed: the response preferences for PTP over PPT and TPP respectively was most pronounced in the beginning of the test phase and then gradually declined.

3.2.2 Discussion. Our prediction was that Dutch listeners—when listening to an artificial language in which two Ps are followed by a T (i.e., …PPTPTPPT…)—would transfer knowledge of an under-representation of non-adjacent labials in their native language, and hence would locate word boundaries within PP sequences. As predicted, PTP items were preferred over OCP-violating PPT and TPP items. The results suggest that preferences cannot be a result of linguistic knowledge other than a phonotactic constraint OCP-Labial, as the stimuli were rigidly controlled for low-level phonotactics and lexical statistics.

All in all, the result is in line with our hypothesis that OCP-Labial is used as a cue to detect word boundaries in the speech signal. However, as OCP-Labial affects consonant distributions in the Dutch lexicon, it might be argued that the preferences in Experiment 1 were not due to segmentation of the artificial language during the familiarization phase, but were merely a result of differences in well-formedness between test items that were applied as late as the test phase, but not during familiarization with the artificial language. Previous studies have shown that listeners use

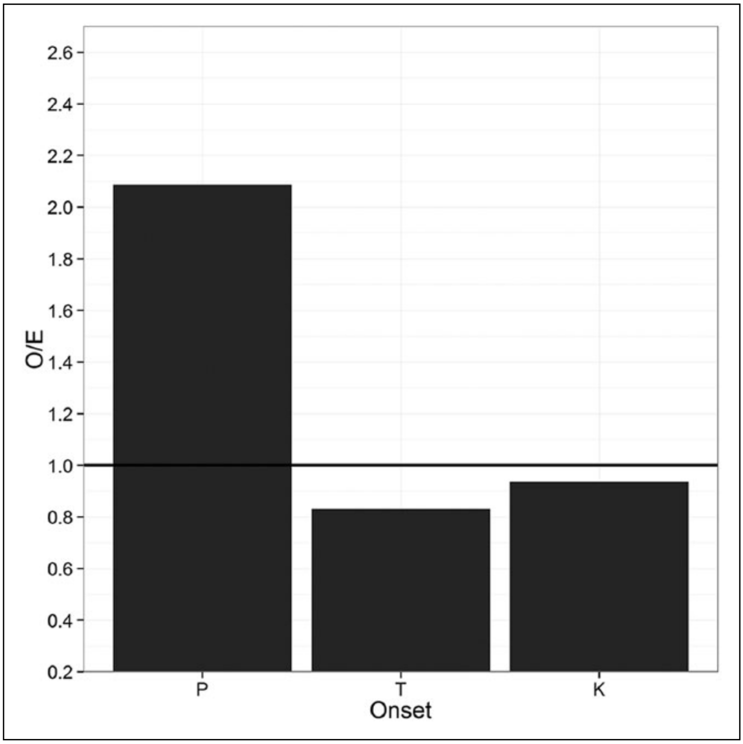

Figure 3. *O/E* probability for labials (P), coronals (T) and dorsals (K) to occur in word-initial position. Calculations are based on counts in CELEX lemma types.

knowledge of OCP-Place for phonotactic well-formedness judgments (e.g., Hebrew: Berent & Shimron, 1997; Arabic: Frisch & Zawaydeh, 2001; English: Coetzee, 2008).

There is some evidence for the interpretation that the preferences in Conditions 1 and 2 are not merely a reflection of the participants' phonotactic well-formedness intuition that—due to OCP-Labial—an item of a shape PTP might make up for a better word of Dutch than an item of a shape TPP or PPT: First, the null-result in Condition 3 and the difference between Condition 3 versus Conditions 1 and 2 is reasonable under the assumption that OCP-Labial has been used to segment the artificial language into …PTP-PTP-PTP…: if participants have segmented sequences of PTP, then it should be difficult for them to display a preference during a test phase that only offered them a forced choice between non-segmented PPT and TPP items, both OCP violators. If the forced choice task calls upon verbatim recall of items that were actually perceived during listening to the artificial language, then participants are expected to have problems in the case of a PPT versus TPP comparison.

Second, there was one factor that the test items could not be controlled for: sometimes, the two labials differed in manner or voice (e.g., /pVm, bVp/), and sometimes, they were identical (e.g., /pVp/, /mVm/). Our analysis did not reveal that the preferences for PTP items were either reduced or enhanced by the presence of PPT and TPP in which the two Ps were identical. This may rule out two possibilities as to how this difference may have affected the results: first, it has been argued in the literature that the effects of OCP-Place are strongest in cases of strict identity (Frisch et al., 2004). However, there are other studies showing that strict identity is exceptional to OCP

(MacEachern, 1999; Gallagher and Coon, 2009). Similarly, in Dutch, co-occurrences of identical labials are in fact less restricted than co-occurrences of near-identical labials, in particular, if they occur word-initially (e.g., *O/E* ratios in CELEX lemma types: #/pVp/ = 2.34, #/bVb/ = 0.97, #/mVm/ = 0.43). If, on the one hand, the overall preference for PTP had only been due to a dispreference for PPT or TPP with two strictly identical labials (e.g., /papotu/), this would have weakened our interpretation that a use of the constraint OCP-Place reflects knowledge of lexical distributions. If, on the other hand, the overall preference for PTP had been weakened in presence of *well*-formed PPT or TPP items with two identical labials, this might have suggested that response preferences were influenced by specific characteristics of the items rather than that a more general knowledge of OCP-Place was responsible for the effect.

A third indication is the interaction of trial number with condition. It is not unexpected that memory for segmented items is freshest immediately after the familiarization phase. Hence, it is reasonable that participants' preferences for PTP words are most pronounced during the initial portion of the test phase, if they have segmented the stream into PTP-PTP-PTP. The emerging response preferences for PPT items in Condition 3 can also be accounted for: In Dutch, Ps are likely to occur at word beginnings (*O/E* = 2.87 in CELEX lemma types, calculated as $E_{\#C1} = p(\#C_1) * N_{\#X}$, see Figure 3). Previous studies (Kager and Shatzman, 2007) have found that Dutch listeners are influenced by this distributional knowledge in a lexical decision task, giving rise to the assumption of a phonotactic constraint Align-Labial requiring labials to align with the left edge of a word. Hence, it is not unexpected that Dutch listeners would be affected by this knowledge in well-formedness judgments. If Dutch listeners know that PPT is phonotactically more well-formed than TPP, it is reasonable that this causes a response bias only in the latter portion of the test phase because at the beginning they had no opportunity to express their preference for the segmented PTP.

If our line of reasoning is correct, then we should be able to show that well-formedness judgments are influenced by Align-Labial. Thus, to further exclude the possibility that the effects found in Experiment 1 were due to preferences that only came into play in the test phase, but did not affect segmentation during the familiarization with the artificial language, we carried out Experiment 2.

Experiment 2 was the same as Experiment 1, except that the familiarization phase was omitted. The test phase stimuli were given in a well-formedness judgment task: participants were required to determine which of two items would be a better word of Dutch. Here, we expect the following results: if the judgments in Experiment 1 were independent of familiarization with the artificial language, then participants in Experiment 2 should display the same preferences as in Experiment 1. However, if the locus of the effect found in Experiment 1 was in segmentation, then the results in the following experiment should be as follows: in Conditions 1 and 2, listeners should have a well-formedness preference for PTP over PPT and TPP nonwords due to phonotactic knowledge of OCP-Place. This outcome would extend findings by Berent and Shimron (1997), Frisch and Zawaydeh (2001) and Coetzee (2008) to material that is more heavily controlled for lexical statistics. In Condition 3 using PPT and TPP nonwords, we expect Dutch listeners to judge PPT to be better-formed than TPP due to phonotactic knowledge of an over-representation of P in word-initial position, i.e., Align-Labial.

4 Experiment 2

4.1 Method

4.1.1 Participants. Thirty-nine native speakers of Dutch (22 women and 17 men, M_{age}: 27.38 years, age range: 18–52 years) with normal hearing, all of whom had not been tested in Experiment 1, participated in the experiment. One additional participant who had made more than two mistakes in the pre-test was excluded.

Table 4. Tests of response preferences in Experiment 2 for each condition against chance. Coefficients (*β*) and their standard errors (*SE*) are logit transformations.

Estimates of fixed effects				
Parameters	*β*	*SE*	*z*	*p*
Condition 1 (PTP vs. PPT)	0.37	0.17	2.18	< 0.05, two-tailed
Condition 2 (PTP vs. TPP)	0.23	0.17	1.36	> 0.1, n.s.
Condition 3 (PPT vs. TPP)	0.36	0.17	2.08	< 0.05, two-tailed

Table 5. Results of the comparisons between conditions in Experiment 2. Coefficients (*β*) and their standard errors (*SE*) are logit transformations.

Estimate of fixed effects				
Parameters	*β*	*SE*	*z*	*p*
Intercept	0.32	0.10	3.24	< 0.01, two-tailed
Condition 1 vs. Condition 2	–0.07	0.12	–0.55	> 0.5, n.s.
Condition 1 + 2 vs. Condition 3	0.02	0.07	0.28	> 0.5, n.s.

4.1.2 Materials, apparatus and procedure. The stimuli were the same as used in the test phase in Experiment 1. Thirteen participants were tested in a sound-attenuated booth. Twenty-six participants were tested in a quiet room with a laptop. The set-up of the experiment was identical to that of Experiment 1 with the only difference being that there was no familiarization phase. The instructions were to indicate via a mouse-click on a button on the screen whether the first or the second of two items would be a better word in Dutch. Thirteen participants were tested on Condition 1, 14 on Condition 2, and 12 on Condition 3.

4.1.3 Data processing and analysis. The filler pairs were excluded from the analysis. The analogous design of Experiments 1 and 2 allowed us to analyze the data in an analogous model. We used the same dependent variable and the fixed and random effects as in Experiment 1. This time, the fixed factors "trial number" and "identity" did not reach significance and were hence not included into the model. For this reason, no random slope for "trial number" was included. Again, we performed a test of condition against chance, and used a helmert contrast for comparing preferences between conditions.

4.2 Results and discussion

4.2.1 Results. The tests of response preferences against chance (see Table 4 for the fixed effects and Appendix G for the random effects) showed that participants had a significant preference for PTP (58%) over PPT in Condition 1. Moreover, in Condition 3, participants displayed a significant preference for PPT (58%) over TPP. In Condition 2, there was a slight preference for PTP (55%) over TPP, which did not show up as significant. None of the comparisons between conditions showed up as significant (see Table 5).

4.2.2 Discussion. First, the results suggest that Dutch listeners' well-formedness judgments were influenced by OCP-Place: their preferences for PTP responses were above chance in Condition 1, and the fact that there was no difference in response preferences between Condition 1 and Condition 2 might speak for a conclusion that, although PTP items were not preferred above chance, participants were doing the same in both tests. This replicates findings by Berent and Shimron (1997), Frisch and Zawaydeh (2001) and Coetzee (2008) while extending these results to material that is more heavily controlled for lexical statistics.

Second, the results of Condition 3 suggest that Dutch listeners have knowledge of the initial predominance of labials, i.e., a constraint ALIGN-LABIAL, leading them to like nonwords starting in Ps to nonwords starting in Ts. With respect to the current study, the results of Condition 3 are crucial as they indicate that the dispreferences for OCP-violating nonwords in Experiment 1 cannot be simply accounted for by well-formedness intuitions that only emerged during the test phase. Rather, the preferences found in the test phase in Experiment 1 must be a result of exposure to the artificial language, and perception must have been driven by OCP-LABIAL. The results support our line of reasoning that participants have segmented the artificial language into …PTP-PTP-PTP…, for which it was impossible for them to display a preference for PPT or TPP in Condition 3.

An issue that merits further explanation is why ALIGN-LABIAL affected well-formedness judgments in Experiment 2, but should not have affected segmentation in Experiment 1. In fact, as a …PTP-PTP-PTP… segmentation satisfies both, ALIGN-LABIAL and OCP-LABIAL, it is very likely that both constraints in conjunction influenced segmentation in Experiment 1. If only ALIGN-LABIAL, but not OCP-LABIAL were a segmentation cue, then both PTP and PPT segmentations were possible outcomes. However, as suggested by the results of Condition 3, participants were not inclined to segment PPT strings from the artificial language.

In fact, a conjoined effect of ALIGN-LABIAL and OCP-PLACE is suggested by the data: Figure 2 shows how decisions developed during the time-course of the test phase in Experiment 1. We have divided responses into three blocks: block 1 shows responses to test pairs 1–12, block 2 shows responses to test pairs 13–24, and block 3 shows responses to test pairs 25–36. As already discussed above, decisions on the first test pairs that immediately followed the familiarization phase are probably the best reflection of what happened during familiarization, as memory for the artificial language should be highest then. It can be seen that during the initial portion of the test phase (test pairs 1–12), preferences for PTP compared to TPP (Condition 2) were highest, which offers an indication of a conjoined effect of ALIGN-LABIAL and OCP-LABIAL.

This result furthermore rules out an alternative explanation evoked by a reviewer, who suggested that the preferences in Experiments 1 and 2 might both be based on well-formedness intuitions with the difference being that in Experiment 2, ALIGN-LABIAL was active, while in Experiment 1, the relative strength of ALIGN-LABIAL was suppressed. The suppression of ALIGN-LABIAL might have been caused by exposure to a language with a high percentage of labials. Although we agree with the reviewer that familiarization with the artificial language must have had some effect on the judgments in the test phase, these effects need not necessarily indicate that these effects were due to segmentation. However, the presence of an effect of ALIGN-LABIAL during the first block of the test phase in conjunction with OCP-PLACE speaks against a deactivation of ALIGN-LABIAL as a result of, and immediately contingent upon, exposure to a language with a high percentage of labials. A second objection that can be raised against the idea that suppression of ALIGN-LABIAL was caused by exposure to a language with a high percentage of labials is that, if this was the case, a similar suppression should have affected OCP-PLACE due to exposure to a language with a high percentage of labial sequences.

Lastly, we will address another issue, namely the question of whether segmentation performance in Experiment 1 was driven by the participants' distributional knowledge about an under-representation of PVP in Dutch. There are two alternative possibilities: First, functional theories might be on the right track to the extent that the use of OCP-PLACE as a speech segmentation cue might be based on a perceptual universal. Second, listeners might display a more general cognitive preference for similarity at edges (ABA > AAB, ABB, regardless of values of A, B). For example, Onnis and colleagues (2005) showed that English speakers could only learn an X_1-C-X_2 pattern if X_1 and X_2 shared manner of articulation, i.e., both start with plosives or both with continuants. To rule out these alternative accounts, we conducted Experiment 3.

Experiment 3 tested whether Dutch listeners would make use of OCP-Coronal as a speech segmentation cue. In Dutch, distributional properties give less reason to use OCP-Coronal (TVT: CELEX: *O/E* = 0.77; CGN: *O/E* = 1.24) for segmentation than OCP-Labial (PVP: CELEX: *O/E* = 0.45; CGN: *O/E* = 0.58). We created an artificial language similar to the language in Experiment 1, the only difference being that one P was replaced by T (…TTPTPTTPTTP…). If the result in Experiment 1 is due to a) a general ABA preference or b) a perceptual bias against homorganic consonants, then participants should segment TPT words from the speech stream in Experiment 3. Our hypothesis, however, is that listeners do not insert boundaries between similar consonants unless this is driven by knowledge of the native language's distributions in the lexicon. So, Dutch listeners should not insert boundaries between two Ts. In light of their knowledge of Align-Labial, it is instead to be expected that Dutch listeners should put P in initial position.

5 Experiment 3

5.1 Method

5.1.1 Participants. Participants were 43 students (six male, 37 female, M_{age} = 20.05 years, age range: 18–31 years) without hearing difficulties and Dutch as a native language, none of whom participated in any other experiment of this study.

5.1.2 Materials. The artificial language was similar to that of Experiment 1, with the only difference being that P_1={pa, bi, mo} was replaced by T (T_2={ta, di, no}). Syllables and their order met the same controls as the languages created for Experiment 1 (see Table 1 and Appendix A). Each 12 TPT, TTP and PTT test items were selected that matched in cohort density and manner of articulation, just as the test items in the previous experiments (see Appendices D and H). No test items had any lexical neighbors. As in Experiment 1, 48 test pairs were constructed with each test item being repeated four times: three times as a target item and once as a filler item.

5.1.3 Apparatus and procedure. Apparatus and procedure were identical to Experiment 1. After familiarization, participants were randomly distributed to one out of three different test phases. Fifteen participants were tested in Condition 1 on pairs of TPT versus TTP, 14 participants were tested in Condition 2 on pairs of TPT versus PTT, and 14 participants were tested in Condition 3 on pairs of PTT versus TTP.

5.1.4 Data processing and analysis. The filler pairs were excluded from the analysis. For the analysis, we used the same fixed and random factors as in Experiment 1. Again, two tests were carried out: a test of each condition against chance, and a test comparing preferences between conditions. In Experiment 3, different comparisons between conditions were of theoretical interest than in Experiments 1 and 2, for which we specified a sliding contrast. This is an orthogonal contrast with which first, Condition 1 is compared to Condition 2, and then Condition 2 is compared to Condition 3. In both Conditions 1 and 2, TPT items were present, which should be the preferred, if Dutch listeners used OCP-Coronal or edge-identity as a segmentation cue. In both Conditions 2 and 3, PTT items were present, which should be the preferred, if Dutch listeners used Align-Labial as a

Table 6. Tests of response preferences in Experiment 3 for each condition against chance. Coefficients (β) and their standard errors (*SE*) are logit transformations.

Estimates of fixed effects				
Parameters	β	*SE*	*z*	*p*
Condition 1 (TPT vs. TTP)	–0.12	0.22	–0.53	> 0.5, n.s.
Condition 2 (TPT vs. PTT)	–0.78	0.25	–3.18	< 0.001, two-tailed
Condition 3 (PTT vs. TTP)	–0.30	0.22	–1.35	> 0.1, n.s.

Table 7. Results of the comparisons between conditions in Experiment 3. Coefficients (β) and their standard errors (*SE*) are logit transformations.

Estimates of fixed effects				
Parameters	β	*SE*	*z*	*p*
Intercept = grand mean	−0.40	0.13	−3.01	< 0.01, two-tailed
Condition 1 vs. Condition 2	−0.67	0.33	−2.02	< 0.05, two-tailed
Condition 2 vs. Condition 3	0.48	0.33	1.44	> 0.1, n.s.

segmentation cue. The intercept in a sliding contrast is the grand mean. Again, "trial number" and "identity" did neither show up to be a significant predictor nor did an inclusion of these fixed factors improve the model. Hence, they were neither included as fixed factors nor random slopes in the consecutive model.

5.2 Results and discussion

5.2.1 Results. As can be seen in Table 6, tests against chance revealed a significant preference for PTT (64%) over TPT (36%) in Condition 2. However, participants' response preferences were neither significantly different from chance in Condition 1 (TPT: 48% versus TTP: 52%) nor in Condition 3 (PTT: 57% versus TTP: 43%).

Comparisons between conditions (see Table 7 for fixed effects, and Appendix I for random effects) showed that there was a difference between the two conditions involving TPT items: in Condition 2 (in the presence of PTT items), participants gave fewer responses for TPT items than in Condition 1 (in the presence of TTP items). However, there was no difference between Condition 2 and 3 (both in which PTT items where present).

5.2.2 Discussion. No preference for TPT items was found. If TPT had been the preferred segmentation, the result in Experiment 1 might have been attributed to a use of OCP-Coronal based on a perceptual bias against homorganic consonant sequences or to a general cognitive preference for identical elements at edges. This explanation can now be rejected, strengthening the interpretation that the effects found in Experiment 1 are due to an influence of language-specific knowledge of OCP-Labial on segmentation.

Instead, results suggest that listeners preferred a PTT segmentation that assigns P to an initial position. The preference for PTT may be due to participants' use of the constraint Align-Labial, reflecting knowledge of the over-representation of Ps at word beginnings in the Dutch lexicon. As there is no interaction, there is no evidence that preferences for PTT items were stronger in

Condition 2 when in context with TPT items than in Condition 3 when in context with TTP items. Still, the preference for PTT responses only showed up as significantly different from chance in Condition 2 when contrasted with TPT, but not in Condition 3, when contrasted with TTP. As already mentioned in the introduction, humans in general like to group similar sounds (e.g., Thorpe and Trehub, 1989; Wertheimer, 1923). As the distribution of TVT is not specified by distributional knowledge of the native language's lexicon, a general bias for grouping similar sounds together may be responsible for the lack of a significant preference for PTT over TTP. In any case, the result of Experiment 3 supports the conclusion that the results in Experiment 1 are due to knowledge of OCP-LABIAL.

6 General discussion

The experiments presented in this study show that Dutch listeners prefer inserting word boundaries between two consecutive labials in an artificial language, suggesting their use of OCP-LABIAL for speech segmentation. Importantly, in our study, participants were able to use OCP-LABIAL pre-lexically as a speech segmentation cue without top-down feedback from the lexicon.

This result poses a problem for current functionalist accounts of OCP that assume that the perception of two homorganic adjacent consonants is inhibited such that it is not possible to faithfully perceive both of them. Frisch (2004) assumes that features that are activated when processing a phoneme are consecutively inhibited, leading to difficulties in encoding the second of two homorganic consonants. Boersma (2000) assumes that two homorganic consonants are pre-lexically merged into a single percept. Both functionalist accounts can explain effects of OCP on word processing (i.e., at the lexical level). However, they fail to account for effects of OCP on (pre-lexical) speech segmentation, as in these accounts, OCP-violating sequences can only be identified as belonging to separate words if the pre-lexical merger/inhibitory effects of OCP are stopped (through lexical feedback) whenever two homorganic entities meet at word boundaries.

It could be argued that our participants tried to access their lexicon whilst listening to the artificial languages in our experiments. However, our stimulus material was rigidly controlled for cues for segmentation from lexical statistics (such as lexical neighborhood, cohort density). Hence, it should have been almost impossible for participants to map word candidates onto lexically similar items.

This suggests that adjacent homorganic consonants must be both faithfully, accurately and individually perceived during pre-lexical processing. Otherwise, OCP-PLACE could not be a speech segmentation cue, as, when crossing word boundaries, homorganic consonants must be identified as belonging to two separate words. Hence, they need to be segregated in perception.

The comparison between Experiment 1 and Experiment 3 suggests that Dutch listeners assign a role to OCP-LABIAL but not to OCP-CORONAL in speech processing, which suggests that they represent the two constraints differently. In the case presented here, the use of the constraints matches distributional information contained in the input language, as in Dutch, there is an evident under-representation of PVP sequences, while this is less so the case with TVT sequences. This difference makes it most likely that the use of the constraint is language-specific and acquired from or triggered by the input language. We take the viewpoint that if the use of OCP-PLACE were attributable to a general cognitive bias independent of distributions in the lexicon then Dutch listeners should have used both OCP-LABIAL and OCP-CORONAL as a cue for speech segmentation.

Still, alternative accounts that might demand further investigation in future studies are the following: First, there might be functional reasons for assuming that listeners have greater difficulty in perceiving or representing consecutive labials than consecutive coronals, although it would need to be explained how a functional bias would exclusively target co-occurrences of labials but not co-occurrences of coronals.

Second, alternatively, the results can be interpreted as evidence for the under-specification theory advocated by Lahiri and colleagues (e.g., Lahiri and Reetz, 2010) suggesting that [coronal] is the default place of articulation, and hence need not be specified. Under this theory, the processing system only needs to take action when processing labials and dorsals. In the framework of this theory it could be assumed that the processing of pairs of coronals should be unproblematic, relative to pairs of labials. It will be important that future studies address this issue further, for example by testing the effects of OCP-Coronal and OCP-Labial on speech segmentation in listeners, in whose native language pairs of coronals and pairs of labials are equally under-represented. Such languages seem to exist, although they are few (Graff, 2012).

In sum, the results of the current study can be interpreted as evidence against current functionalist theories of OCP. Not only do they indicate that the effects of OCP-Place are triggered by or acquired from the input. There can neither be a perceptual merger nor an inhibition of the features shared by homorganic consonants. It is necessary to re-conceive of OCP as a constraint that *exclusively* takes action at the morpheme level as our results indicate that OCP is active during pre-lexical auditory speech processing, that is, before they have access to morpheme-sized elements, without lexical feedback.

Several points remain for discussion. First, it needs to be noted that segmentation preferences found in the current study may seem rather weak when compared with segmentation preferences reported in other artificial language learning studies and studies on the psychological reality of OCP-Place. An explanation for the relative weakness of the effect might be that the material used in the current study was more rigidly controlled for any potential confounding effect on speech segmentation than the material used in previous studies.

For the purpose of this study, this was necessary as our aim was to show that OCP-Place by itself affects pre-lexical processing without additional feedback from lexical statistics. In natural language processing, however, it is most likely that native listeners of Dutch will rely on a sum of speech segmentation cues that is available. Regarding the processing of PVP sequences in natural Dutch, it is to be expected that they will not only rely on OCP-Place, but also on the sparse number of lexical cohorts (i.e. words that start in PVP), the biphone probabilities of PV and VP and so forth. So, logically speaking, the less rigidly stimulus material is controlled for, the more participants will rely on such additive cues for processing.

In natural language processing, listeners can rely on many different pre-lexical segmentation cues from phonetic, prosodic and phonotactic distributions. In addition to pre-lexical cues, listeners are known to use lexical segmentation cues: knowledge of real words helps in predicting word boundaries of neighboring words (e.g., Mattys, White and Melhorn, 2005). Many of these cues arguably have a much higher rate of return and, hence, are more accessible and more reliable than OCP-Place. Still, under the assumption that listeners use the sum of all available cues for speech segmentation, and in light of the results of this study, OCP-Place will probably be one of them. In any case, the relative importance of OCP-Place in processing is irrelevant with respect to the theoretical implications of the results of our study, which were of major interest in this paper.

The finding of the current study that native listeners possess knowledge of non-adjacent consonant probabilities that abstracts over intervening vowels (e.g., /pVm/, /bVp/, etc.) may have important implications for language acquisition. It has been suggested that non-adjacent dependencies between phonological units are difficult to acquire due to their weak statistical support in the input (e.g., Pierrehumbert, 2003). Previous artificial language learning studies have suggested that non-adjacent dependencies can only be learned from continuous streams of speech if they affect units that are similar to some degree. Non-adjacent dependencies between phonemes (i.e., between consonants or vowels) have been found to be learnable (e.g., Newport and Aslin, 2004), while non-adjacent dependencies between syllables have been found to be learned only if sequences begin with consonants that share manner of articulation features (e.g., Newport and Aslin, 2004; Onnis et al., 2005; Perruchet and Pacton, 2006). However, in the case of OCP-Place, the phonotactic requirement seems to be rather the opposite, i.e., that words do *not* contain sequences of similar phonemes. If similarity is needed for non-adjacent dependency learning, then how can phonotactic knowledge of OCP-Place be acquired?

It is possible that not only when acquiring knowledge of an over-representation, but also when acquiring knowledge about an under-representation of phoneme co-occurrences, learners rely on a universal bias for similarity grouping (e.g., Thorpe and Trehub, 1989; Wertheimer, 1923). That is, under-representations that involve featurally identical segment pairs will naturally stand out, and hence would be easier to learn than under-representations involving featurally-non-identical pairs. It will be interesting for future studies to further investigate the role of a bias for attending to similarity in the acquisition of OCP-Place. For example, it will be interesting to test whether non-adjacent dependencies involving pairs of non-similar consonants can be used for segmentation as well, and also whether such dependencies are indeed less learnable than those involving pairs of similar consonants. In sum, the current study provides evidence against a functionalist theory of OCP that assumes natural inhibitions against perceiving consecutive homorganic consonants. However, the study leaves open the possibility that an innate cognitive bias for grouping similar elements may nevertheless play a role in the acquisition of OCP.

7 Conclusion

The results of this study offer evidence that OCP-Place can give cues to word boundaries in speech segmentation, if pairs of homorganic consonants are under-represented in the lexicon of the native language. This result challenges functional accounts of OCP for three reasons: First, effects of OCP-Place are language-specific, and not, as assumed by functionalist theory, an effect of a perceptual bias. Second, functionalists argue that the perceptual bias amounts to difficulty in faithfully and independently perceiving two homorganic elements. However, the ability to use OCP as a segmentation cue implies the ability to perceive two homorganic elements faithfully and independently, in order to correctly assign them to two separate words. Third, if functionalist theories were correct, then lexical access would be inhibited (rather than facilitated) if two homorganic consonants meet at word boundaries—a problem that can only be overcome with lexical feedback. In the present study, however, the use of OCP-Place for segmentation was exclusively pre-lexical.

Acknowledgments

We thank Theo Veenker for programming the experiment, Frits van Brenk, Jorinde Timmer, and Bart Penning de Vries for help with testing participants for Experiment 3, Frans Adriaans for supplying us with the software PhonotacTools for calculating lexical statistics, and Arthur Dirksen from *Fluency* for a script to generate the artificial language streams. This work has benefitted from comments by two anonymous reviewers, Frans Adriaans, Paul Boersma, Luca Bonatti, Elise de Bree, Tom Fritzsche, Annemarie Kerkhoff, Tom Lentz Claartje Levelt, Jacques Mehler, and the audiences at the *Linguistics in the Netherlands-Dag 2008*, the *Manchester Phonology Meeting 2008* and *Laboratory Phonology Conference 2008*.

Funding

This research was carried out as part of the project “Phonotactic constraints for speech segmentation: The case of second language acquisition” funded by an NWO grant (277-70-001) awarded to René Kager.

Notes

1. Artificial language learning studies on statistical learning classically include a control experiment in which statistical cues are redistributed in such a way that part-words become words, and words become part-words (e.g., Peña et al., 2002). This is to prevent from effects being due to idiosyncrasies of the stimulus material, in particular the order of specific syllables in the artificial language. In our case, by controlling for low-level phonotactics and lexical statistics, we already severely reduced the possibility that segmentation preferences were caused by unknown properties of the syllables and their order. Still, to ultimately exclude this confounding factor, we created two languages.
2. In Condition 3, PPT was "well-formed" for reasons that will be discussed in the Discussion of Experiment 1.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995) The CELEX Lexical Database (CD-ROM). Philedelphia: Linguistic Data Consortium, University of Pennsylvania.

Bates, D., Maechler, M., & Bolker, B. (2012) lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.

Berent, I., Everett, D. L., & Shimron, J. (2001) Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology 42*: 1–60.

Berent, I., & Shimron, J. (1997) The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition 64*: 39–72.

Berkley, D. (1994) The OCP and gradient data. *Studies in the Linguistic Sciences 24*: 59–72.

Boersma, P. (1998) The OCP in functional phonology. *Rutgers Optimality Archive 283*: 27 pages.

Boersma, P. (2000) The OCP in the perception grammar. *Rutgers Optimality Archive 435*: 1–52.

Boersma, P. (2009) Cue constraints and their interactions in phonological perception and production. In: P. Boersma, & S. Hamann (eds) *Phonology in perception*. Berlin: Mouton de Gruyter, 55–110.

Coetzee, A. (2005) The OCP in the perception of English. In:S. Frota, M. Vigario, & JM Freitas (eds) *Prosodies*. New York: Mouton de Gruyter, 223–245.

Coetzee, A. (2008) Grammar is both categorical and gradient. In: S. Parker (ed) *Phonological Argumentation*. London: Equinox.

Coetzee, A., & Pater, J. (2008) Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory 26*: 289–337.

Cutler, A., & Norris, D. (1988) The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance 14*: 113–121.

Dutoit, T., Pagel, V., & Pierret, N. (1996) The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes. *Proceedings of The Fourth International Conference on Spoken Language Processing (ICSLP'96) 3*. 1393–1396.

Eriksen, C., & Schutze, D. (1978) Temporal factors in visual information processing: A tutorial review. In: J. Requin (ed) *Attention and performance VII*. Hillsdale, NJ: Lawrence Erlbaum.

Finn, A. S., & Hudson Kam, C. L. (2008) The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition 108*: 477–499.

Frisch, S. A. (2004) Language processing and segmental OCP effects. In: B. Hayes, R. M. Kirchner, & D. Steriade (eds) *Phonetically based Phonology*. Cambridge: Cambridge University Press, 346–371.

Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004) Similarity avoidance and the OCP. *Natural Language & Linguistic Theory 22*: 179–228.

Frisch, S. A., & Zawaydeh, B. (2001) The psychological reality of OCP-Place in Arabic. *Language Acquisition 77*: 91–106.

Gafos, A. I. (2003) Greenberg's Asymmetry in Arabic: A Consequence of Stems in Paradigms. *Language 79*: 317–355.

Gallagher, G., & Coon, D. (2009) Distinguishing total and partial identity: Evidence from Chol. *Natural Language and Linguistic Theory* 27: 545–582.

Goddijn, S., & Binnenpoorte, D. (2003) Assessing manually corrected broad phonetic transcriptions in the Spoken Dutch Corpus. *15th ICPhS*. Barcelona, 1361–1364.

Goldsmith, J. (1979) *Autosegmental Phonology*, New York: Garland.

Graff, P. (2012) *Communicative Efficiency in the Lexicon*, PhD Thesis, MIT.

Greenberg, J. (1950) The patterning of root morphemes in Semitic. *Word 5*: 162–181.

Jäger, T. F. (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language 59*: 434–446.

Johnson, E. K., & Jusczyk, P. W. (2001) Word Segmentation by 8-Month-Olds: When Speech Cues Count More Than Statistics. *Journal of Memory and Language 44*: 548–567.

Kabak, B., Maniwa, K., & Kazanina, N. (2010) Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Journal of Laboratory Phonology 1*: 207–224.

Kager, R., & Shatzman, K. (2007) Phonological constraints in speech processing. In: B. Los, & Mv. Koppen (eds) *Linguistics in the Netherlands 2007*. Amsterdam: John Benjamins, 100–111.

Kager, R., & Shatzman, K. (forthcoming) Similarity Avoidance in Speech Processing.

Kanwisher, N. (1987) Repetition blindness: Type recognition without token individuation. *Cognition 27*: 117–143.

Lahiri, A., & Reetz, H. (2010) Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics 38*: 44–59.

Leben, W. (1980) *Suprasegmental phonology*, New York: Garland Press.

Luce, P. A. (1986) *Neighborhoods of words in the mental lexicon*, Doctoral dissertation, Indian University, Bloomington, IN.

MacEachern, M. (1999) *Laryngeal co-occurrence restrictions*, New York: Garland Press.

Mattys, S. L. (2004) Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance 30*: 397–408.

Mattys, S. L., & Jusczyk, P. W. (2001) Phonotactic cues for segmentation of fluent speech by infants. *Cognition 78*: 91–121.

Mattys, S. L., White, L., & Melhorn, J. F. (2005) Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology: General, 134*(4): 477–500.

Mayer, T., Rohrdantz, C., & Plank, F. (2010) Consonant co-occurrence in stems across languages: Automatic analysis and visualization of a phonotactic constraint. *Proceedings of the ACL 2010 Workshop on NLP and Linguistics: Finding the Common Ground*. Uppsala, Sweden, 67–75.

McCarthy, J. J. (1985) *Formal Problems in Semitic Phonology and Morphology*, New York: Garland.

McCarthy, J. J. (1986) OCP effects: gemination and antigemination. *Linguistic Inquiry 17*: 207–263.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009) Within-category VOT affects recovery from "lexical" garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language 60*: 65–91.

Newport, E. L., & Aslin, R. N. (2004) Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology 48*: 127–162.

Onnis, L., Monaghan, P., & Richmond, K. (2005) Phonology impacts segmentation in speech processing. *Journal of Memory and Language 53*: 225–237.

Peña, M., Bonatti, L., & Nespor, M. (2002) Signal-driven computations in speech processing. *Science 298*: 604–607.

Perruchet, P., & Pacton, S. (2006) Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences 10*: 233–238.

Pierrehumbert, J. B. (1993) Dissimilarity in the Arabic verbal roots. *Proceedings of the North East Linguistics Society*. 367–381.

Pierrehumbert, J. B. (2003) Probabilistic phonology: Discrimination and robustness. In: R. Bod, J. Hay, & S. Jannedy (eds) *Probabilistic linguistics*. Cambridge MA: MIT Press, 177–228.

Pozdniakov, K., & Segerer, G. (2007) Similar Place Avoidance: A statistical universal. *Linguistic Typology 11*: 307–348.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996) Statistical learning by 8-month-old infants. *Science 274*: 1926–1928.

Shatzman, K., & Kager, R. (2007) A role for phonotactic constraints in speech perception. In: J. Trouvain, & W. J. Barry (eds) *Proceedings of the 16th International Congress of Phonetic Sciences*. Dudweiler: Pirrot, 1409–1412.

Shatzman, K., & McQueen, J. (2006) Segment duration as a cue to word boundaries in spoken-word recognition. *Attention, Perception and Psychophysics 68*: 1–16.

Stemberger, J. P., Elman, J. L., & Haden, P. (1985) Interference between phonemes during monitoring: Evidence for an interactive activation model of speech perception. *Journal of Experimental Psychology: Human Perception and Performance 11*: 475–489.

Stemberger, J. P., & MacWhinney, B. (1986) Form-oriented inflectional errors in language processing. *Cognitive Psychology 18*: 329–354.

Suomi, K., McQueen, J., & Cutler, A. (1997) Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language 36*: 422–444.

R Core Team (2012) *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing.

Thorpe, L. A., & Trehub, S. E. (1989) Duration illusion and auditory grouping in infancy. *Developmental Psychology 25*: 122–127.

Van Kampen, A., Parmaksiz, G., & van de Vijver, R. (2008) Metrical and statistical cues for word segmentation: Vowel harmony and word stress as cues to word boundaries by 6- and 9-month old Turkish learners. In: A. Gavarró, & M. J. Freitas (eds) *Language acquisition and development*. Newcastle: Cambridge Scholars Publishing, 313–324.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998) The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language 38*: 133–149.

Waals, J. (1999) *An experimental view of the Dutch Syllable. Doctoral dissertation*, Utrecht: LOT.

Wertheimer, M. (1923) Untersuchungen zur Lehre von der Gestalt II. *Psycologische Forschung 4*: 301–350.

Appendix A. Syllable Frequency in Experiment 1 & 3 in % per within word position. Counts are based on word types in the Dutch CELEX lexical database.

Syllable frequencies	Syllable	% ini	% med	% fin
Experiment 1	P_1	0.50	0.43	0.07
Language A	T	0.60	0.37	0.03
	P_2	0.45	0.52	0.03
Experiment 1	P_1	0.48	0.48	0.04
Language B	T	0.61	0.37	0.02
	P_2	0.50	0.43	0.07
Experiment 3	T_1	0.48	0.48	0.04
	P	0.61	0.37	0.02
	T_2	0.50	0.43	0.07

Appendix B. Test items in Experiment 1a, Language A.

F = voiceless bilabial, V = voiced bilabial, N = nasal bilabial.

Condition 1

PPT	Manner	PTP	Manner
papotu	FFF	potupa	FFF
papone	FFN	potumo	FFN
papodo	FFV	potubi	FFV
pamatu	FNF	ponepa	FNF
pamane	FNN	ponemo	FNN
pamado	FNV	ponebi	FNV
pabetu	FVF	podopa	FVF
pabene	FVN	podomo	FVN
pabedo	FVV	podobi	FVV
mobetu	NVF	madopa	NVF
mobene	NVN	madomo	NVN
mobedo	NVV	madobi	NVV

Condition 2

PTP	Manner	TPP	Manner
potupa	FFF	tupapo	FFF
potumo	FFN	tupama	FFN
potubi	FFV	tupabe	FFV
ponepa	FNF	tumopo	FNF
ponemo	FNN	tumoma	FNN
ponebi	FNV	tumobe	FNV
betupa	VFF	dopapo	VFF
betumo	VFN	dopama	VFN
betubi	VFV	dopabe	VFV
benepa	VNF	domopo	VNF
benemo	VNN	domoma	VNN
benebi	VNV	domobe	VNV

Condition 3

PPT	Manner	TPP	Manner
papotu	FFF	tupapo	FFF
papone	FFN	tupama	FFN
papodo	FFV	tupabe	FFV
pabetu	FVF	tubipo	FVF
pabene	FVN	tubima	FVN
pabedo	FVV	tubibe	FVV
mopotu	NFF	nepapo	NFF
mopone	NFN	nepama	NFN
mopodo	NFV	nepabe	NFV
mobetu	NVF	nebipo	NVF
mobene	NVN	nebima	NVN
mobedo	NVV	nebibe	NVV

Appendix C. Test items in Experiment 1b, Language B.

F = CV syllable starting in voiceless bilabial, V = CV syllable starting in voiced bilabial, N = CV syllable starting in nasal bilabial.

Condition 1				Condition 2			
PTP		TPP		PTP		PPT	
petube	FFV	tupubi	FFV	petupo	FFF	popitu	FFV
petuma	FFN	tupumo	FFN	petube	FFV	popide	FFV
pedepo	FVF	tubope	FVF	petuma	FFN	popina	FFN
pedema	FVN	tubomo	FVN	bodepo	VVF	bebatu	VVF
penabe	FNV	tumibi	FNV	bodebe	VVV	bebade	VVV
botuma	VFN	depumo	VFN	bodema	VVN	bebana	VVN
bodepo	VVF	debope	VVF	bonapo	VNF	bemotu	VNF
bodema	VVN	debomo	VVN	bonabe	VNV	bemode	VNV
bonapo	VNF	demipe	VNF	bonama	VNN	bemona	VNN
mitube	NFV	napubi	NFV	midepo	NVF	mabatu	NVF
midepo	NVF	nabope	NVF	midebe	NVV	mabade	NVV
minapo	NNF	namipe	NNF	midema	NVN	mabana	NVN

Appendix D. Mean cohort densities with a cut-off-point of 3 phonemes of the items used in the different test phases of Experiments 1-3.

Experiment	Condition	Cohort density 1	Cohort density 2
1a, 3	Condition 1	PTP 16.89 (12.57)	PPT 16.28 (20.30)
	Condition 2	PTP 29.09 (15.57)	TPP 28.24 (17.02)
	Condition 3	PPT 16.28 (20.30)	TPP 16.13 (15.75)
1b	Condition 1	PTP 22.00 (13.35)	PPT 14.66 (14.13)
	Condition 2	PTP 37.89 (30.21)	TPP 36.36 (25.29)
3	Condition 1	TPT 39.36 (37.57)	TTP 56.33 (55.87)
	Condition 2	TPT 39.36 (37.57)	PTT 49.28 (41.00)
	Condition 3	TTP 56.33 (55.87)	PTT 49.28 (41.00)

Appendix E. Random and fixed effects when comparing Experiments 1a and 1b.

Estimates of random effects			*SD*	Correlation	
Groups	Name	Variance			
Items	Intercept	0.62092	0.78798		
	Condition 1 vs. 2	0.90579	0.95173	−0.86	
	Trial number	0.00007	0.00860	−1.00	0.86
Participants	Intercept	0.25108	0.50107		
	Trial number	0.00006	0.00754	−0.41	

(Continued)

Appendix E. (Continued)

Estimates of fixed effects	β	*SE*	*Z*	*p*
Parameter				
Intercept	0.38	0.27	1.42	> 0.1, n.s.
Experiment	−0.20	0.37	−0.54	> 0.5, n.s.
Condition	0.09	0.36	0.25	> 0.5, n.s.
Trial number	−0.004	0.01	−0.55	> 0.5, n.s.
Experiment * Condition	0.32	0.50	0.64	> 0.5, n.s.
Experiment * Trial number	0.01	0.01	0.64	> 0.5, n.s.
Condition * Trial number	−0.007	0.01	−0.43	> 0.5, n.s.
Experiment * Condition * Trial number	−0.02	0.02	−1.02	> 0.1, n.s.

Appendix F. Random effects for random factors and random slopes in Experiment 1.

Estimates of random effects		Variance	*SD*	Correlations		
Groups	Name					
Items	Intercept	0.12496	0.35349			
	Condition 1 vs. 2	0.21823	0.46715	−0.446		
	Condition 1+2 vs. 3	0.02199	0.14829	−0.995	0.472	
	Trial number	0.00010	0.01014	−0.859	0.841	0.872
Participants	Intercept	0.28054	0.52966			
	Trial number	0.00001	0.00381	−0.841		

Appendix G. Random effects for random factors and random slopes in Experiment 2.

F = CV syllable starting in voiceless bilabial, V = CV syllable starting in voiced bilabial, N = CV syllable starting in nasal bilabial.

TTP		TPT		PTT	
tudipo	FVF	tabetu	FVF	podota	FVF
tudibe	FVV	tabedo	FVV	pododi	FVV
tudima	FVN	tabene	FVN	podono	FVN
tunopo	FNF	tamatu	FNF	poneta	FNF
tunobe	FNV	tamado	FNV	ponedi	FNV
tunoma	FNN	tamane	FNN	poneno	FNN
dotapo	VFF	dipotu	VFF	betuta	VFF
dotabe	VFV	dipodo	VFV	betudi	VFV
dotama	VFN	dipone	VFN	betuno	VFN
netapo	NFF	nopotu	NFF	matuta	NFF
netabe	NFV	nopodo	NFV	matudi	NFV
netama	NFN	nopone	NFN	matuno	NFN

Appendix H. Test items in Experiment 3.

Estimates of random effects		Variance	*SD*	Correlations	
Groups	Name				
Items	Intercept	0.000000001	0.000034		
	Condition 1 vs. 2	0.20857	0.45670	0.00	
	Condition 1+2 vs. 3	0.01411	0.11879	0.00	−1.00
Participants	Intercept	0.22669	0.47612		

Appendix I. Random effects for random factors and random slopes in Experiment 3.

Estimates of random effects		Variance	*SD*	Correlations	
Groups	Name				
Items	Intercept	0.00858	0.09264		
	Condition 1 vs. 2	2.19074	1.48011	−0.008	
	Condition 2 vs. 3	2.00975	1.41765	−0.009	−0.915
Participants	Intercept	0.52036	0.72136		