The Obligatory Contour Principle

Consonant co-occurrence restrictions in Dutch

MA-thesis Taal, Mens en Maatschappij
Utrecht University
Wil van Goch
Supervisor: Prof. Dr. René Kager
Co-reader: Prof. Dr. Wim Zonneveld
June 2010

Hilversum, June 2010

Cover illustration: Gerard Janssen

The Obligatory Contour Principle

Consonant co-occurrence restrictions in Dutch

Contents

1 Introduction

Phonology studies the fascinating world of language sounds. Since the claim made by Chomsky (1964) that language is built on a universal innate language system that needs language-specific filling-in, investigations in linguistics have focused on innate language and the ways in which the system is organized. For the past decades research has focused on general factors and how these fit into a system with universal language as its basis. Generative linguistics studies phenomena that are encountered cross-linguistically, trying to find evidence for a common basis for the phenomena in the universal foundation. Generative linguists search for similarities and distinctions between languages to establish which aspects found in languages can be attributed to Universal Grammar and which cannot.

One of the topics studied over the years is a phonological restriction on the co-occurrence of consonants with the same features in each other's vicinity. The phenomenon of co-occurrence restrictions was first recognized by Leben (1973). Goldsmith (1976) introduced the term Obligatory Contour Principle, henceforth also referred to as OCP:

(1)

Obligatory Contour Principle:	At the melodic level of the grammar, any adjacent tonemes must be distinct

This was later adapted by McCarthy (1988):

(2)

Obligatory Contour Principle:	Adjacent identical elements are prohibited

The study into the phenomenon has been taken up in connection with various languages. It was established in studies into the phenomenon that restrictions are often not absolute, but that languages show gradient OCP effects claimed to be connected with diverse matters. Since 1973 several theories on influences on OCP have been developed.

Languages, as diverse as Arabic, (McCarthy 1988, Pierrehumbert 1993, Frisch et al. 2004), Javanese (Yip 1989), Japanese (Ito & Mester 1998, Kawahara 2005), Russian (Padgett 1995), Muna (Coetzee & Pater 2008), English (Berkley 1994, 2000) and Latin (Berkley 2000) have been shown to make use of OCP restrictions. There is evidence of the psychological reality of OCP in Jordanian Arabic (Frisch and Zawaydeh 2001), Hebrew (Berent and Shimron 1997, Coetzee 2006), English (Coetzee 2005, 2009) and Dutch (Kager and Shatzman 2007).

In the present thesis I check the claims made in the literature against the facts of co-occurrence restrictions in Dutch. There are two ways of approaching this linguistic phenomenon; the first is finding evidence of the phenomenon in the language based on the investigation of language data, and the second is testing the psychological reality of the phenomenon for speakers of a language. I have opted for the first variety. To my knowledge there is no investigation into OCP in Dutch. My aim is to thoroughly investigate the subject and connect the findings on Dutch with claims from the literature, thus defining common ground and determining which aspects of OCP are universal and which are not. The present thesis will contain a comprehensive inventory of literature on the subject, followed by a detailed analysis of OCP in Dutch. Aspects that will emerge from the literature on OCP in diverse languages will be examined in relation to OCP in Dutch.

This thesis is structured as follows. Chapter 2 provides a survey of OCP as found in the literature on co-occurrence restrictions, including publications on the psychological reality of the phenomenon in various languages. Chapter 3 introduces some phonotactics of Dutch relevant to the subject. In Chapter 4 I focus on OCP in Dutch and investigate the aspects which in Chapter 2 were professed to influence OCP, as well as two more aspects possibly influencing OCP. Chapter 5 supplies conclusions on OCP in Dutch and provides an OT analysis of OCP in Dutch. Finally, the overall conclusion in Chapter 5 contains information on cross-linguistic common ground on OCP and marks out universal and language-specific properties of the phenomenon.

2 The Obligatory Contour Principle

This chapter will be concerned with the history of the Obligatory Contour Principle, as discussed in the literature. Also publications on the psychological reality of the phenomenon for native speakers of different languages will be dealt with. I will first give a survey of the history of OCP based on the exploration of language data. I will start with OCP in connection with tone languages, which have suprasegmental co-occurrence restrictions on melody patterns. This will be followed by OCP in autosegmental phonology, in which not only suprasegmental restrictions but also restrictions between segments and features of segments are discussed. In connection with this the issue of gradient OCP will come up. Aspects in the literature that have been claimed to influence gradient OCP will be under consideration. I will finish the section on the history of OCP with literature on OCP in Optimality Theory. After the literature on OCP based on language data, I will discuss a different approach to the subject, viz. research into the psychological reality of OCP. Experiments on the psychological reality of OCP for native speakers of Jordanian Arabic, Hebrew, English and Dutch, will be dealt with. Chapter 2 will be divided into two parts: 2.1 will contain a survey of the history of OCP, 2.2 will deal with psychological reality. At the end of the chapter I will introduce four hypotheses for the subject of the present thesis: OCP in Dutch.

2.1 OCP: a survey

In this section the first instance of OCP and the development of the phenomenon in linguistic theory will be presented. A description of co-occurrence restrictions in melody patterns will be followed by a description of OCP in autosegmental tiers, and the introduction of features as (partial) definers of OCP. Furthermore, the discovery of gradience and attempts at accounting for this with the help of features, distance and salience will be dealt with. Analyses in Optimality Theory, with surface-based OCP, will complete the survey.

2.1.1 Co-occurrence restrictions in melody patterns

The phenomenon of the non-occurrence of identical sounds within grammatical domains was initially discovered in connection with tone languages. Leben (1973) examined the tone language Mende, a Mande language of Sierra Leone. The list below sums up the tonal patterns he observed in monomorphemic nouns of maximally three syllables; H indicates a high tone, L represents a low tone.

(3) Surface tone pattern on nouns in Mende:

H	L		
HH	LL	HL	LH
HLL	LHH	LHL	

Leben did not find any HHL or LLH nouns. His analysis of the tonal patterns was that underlyingly five tonal patterns could be recognized in Mende, viz. H, L, HL, LH, LHL[1]. These tonal patterns must be mapped to a string of segments. Since the number of segments does not always coincide with the number of tones, mapping rules are needed. The mapping rules for Mende are (Leben 1973, p 44):

(4) Mapping rules for Mende

Tone Mapping (Mende)

a If the number of level tones in the pattern is equal to or less than the number of vowels in the word possessing the pattern, put the first tone on the first vowel, the second on the second, and so on; any remaining vowels receive a copy of the last tone in the pattern.

E.g. ^{LHL}nikili becomes nikili
L H L

^{H}pɛlɛ becomes pɛlɛ and then pɛlɛ
H H H

[1] His explanation for the absence of HLH is that: "at most a single prosodeme stands out among the others", p 170. HLH would contain two high-toned prosodeme, and only one is permitted.

b If the number of level tones in the pattern is greater than the number of vowels in the word possessing the pattern, put the first tone on the first vowel, the second on the second, and so on; remaining tones are expressed as a sequence on the last vowel available.

E. g. HL mbu becomes mbu
HL
LHL nyaha becomes nyaha
L HL

With the help of these mapping rules and the five tonal patterns of Mende we find the following surface realizations of nouns[2]:

(5) Tonal patterns in Mende with different surface realizations:

H	kɔ H	*person*	pɛlɛ H H	*house*		
L	kpa L	*debt*	bɛlɛ L L	*trousers*		
HL	mbu HL	*owl*	kenya H L	*uncle*		
LH	mba LH	*rice*	nika L H	*cow*		
LHL	mba LHL	*companion*	nyaha L HL	*woman*	nikili L H L	*groundnut*

Leben's explanation for the surface realization of underlying H and LH as HH and LHH, respectively, and of underlying L and HL as LL and HLL, was tone spreading. In words with more vowels in the word than tones in the tonal pattern, the tones of the pattern are applied from left to right. The final tone of the pattern is copied and spreads to all remaining vowels in the word. Surface realizations with adjacent identical tones are thus a consequence of the Mende mapping rules with rightward spreading. There are no underlying (lexical) tonal patterns with adjacent identical tones. Leben was the first to describe a restriction on the co-occurrence of identical features in a linguistic domain.

[2] There are no examples in Leben 1973 of rightward spreading of HL to HLL or LH to LHH

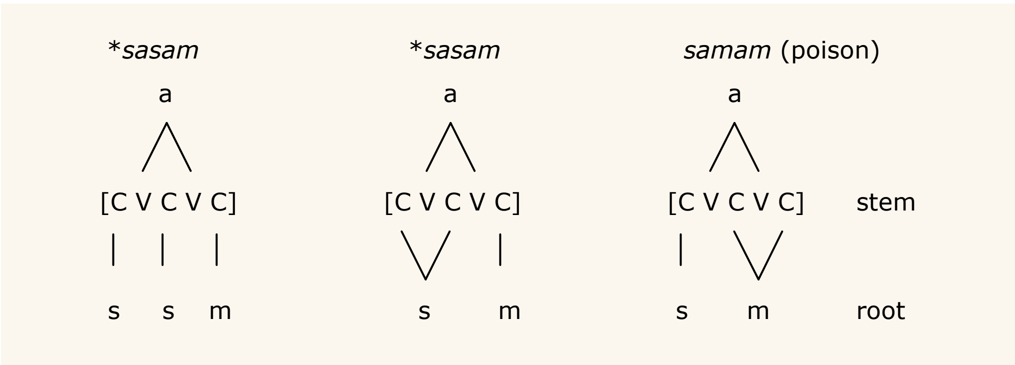
2.1.2 Autosegmental tiers

Important work on the OCP was performed in autosegmental phonology. Autosegmental phonology distinguishes levels in phonological representations – also called tiers – such as the consonantal level, vowel level, melodic level, syllabic level, etc. Autosegmental phonology opened new possibilities for dealing with a single feature that was not connected to a single segment, but to a sequence of segments, as well as possibilities for dealing with multiple features that had to be connected to a single segment. One of the suprasegmental feature levels is the melodic level. In his dissertation of 1976, Goldsmith analyses Leben's findings on tone languages in autosegmental phonology: "... when a vowel desyllabifies or is deleted by some phonological rule, the tone it was bearing does not disappear – rather, it shifts its location and shows up on some other vowel. The toneme, or more generally the tone melody, has a stability which maintains it independently of the other aspects of the signal, and thus the tone melody preserves itself despite modifications to the syllabic structure..... If the tone were a feature of a segment it would be deleted along with the segment" (Goldsmith 1976, p 30). As indicated above, Goldsmith was also the first to use the term Obligatory Contour principle for the facts Leben had found for tone languages. He stated the principle as follows (Goldsmith 1976, p 36):

(6) <u>Obligatory Contour Principle:</u> At the melodic level of the grammar, any two adjacent tonemes must be distinct

McCarthy (1986) discovered that the Obligatory Contour principle was applicable not only to suprasegmental levels or tiers, but that the principle could also be applied to identical segments. He investigated the OCP at the segmental level in Semitic roots and found a ban on *[C_1VC_1X], whereas there was no ban on [$C_1VC_2VC_2$]. There were no stems like **sasam*, but there were stems like *samam*. He explained the phenomenon as a ban on identical consonants in roots. A root in Classical Arabic is a combination of consonants which forms the basis for inflections. The root *ktb,* e.g. represents 'write'*.* Stems derived from the root are CVCVC, *katab* 'to write', *kutib* 'to be written'. In Classical Arabic, roots are analyzed as having 2, 3 or 4 consonants. McCarthy (1986) discussed roots with 2 and 3 consonants. As soon as there is a triconsonantal root, one sees a limitation. There are no triconsonantal CVCVC stems with identical consonants 1 and 2; there do appear to be triconsonantal CVCVC stems with identical consonants 2 and 3[3]. However, according to McCarthy, there is a constraint against any adjacent identical consonants in a root. He called the phenomenon anti-gemination. Triconsonantal stems with identical consonants 2 and 3 are based on biconsonantal roots. Similar to Leben's analysis of tone-spreading, McCarthy's explanation for the occurrence of final geminates in CVCVC-stems was autosegmental spreading with rightward mapping. His analysis of **sasam* and *samam* is in (7) below:

(7) analysis of *sasam and samam



Classical Arabic prohibits the occurrence of two identical adjacent consonants in a morpheme, in this case a root.

2.1.3 Features

Another important development in the history of the OCP was the introduction of features as OCP markers. This section will be concerned with the features [place], [manner] and [voice].

[3] "...it is worth noting that there are no quadriliteral verb forms with doubling of any consonant except the final one, ..." (McCarthy 1986, p 209).

2.1.3.1 The feature [place]

McCarthy (1988) was the first to discover that OCP affected not only supra-segmental features and segments as a whole, but that OCP was also active between non-identical segments sharing a feature. There were OCP-place effects across an intervening consonant in Arabic verbal roots. He investigated the role of the feature [place] in relation to OCP. McCarthy argued in favour of a division of the category of place into four articulatory classes, viz. [labial], [coronal], [dorsal] and [pharyngeal], each with their separate tiers, and found proof of OCP restrictions within these categories. In Arabic, e.g., the root **ftb* is represented as:

(8) **ftb*

```
[lab]    [lab]
  |        |
  f    t   b
       |
     [cor]
```

In this analysis *f* and *b* are adjacent on the labial tier and therefore prohibited. McCarthy thus introduced the feature [place] in the history of the OCP. It must be noted that, with this division, McCarthy could not account for the fact that the OCP-place restriction was not applicable to all place tiers. Coronal consonants appeared not to be affected. The consonantal tier – the segments in **sasam* – required absolute OCP, for the place tier he observed a "tendency" for OCP (McCarthy 1988).

2.1.3.2 The feature [manner]

With regard to manner of articulation of consonants, we distinguish the following features: [consonantal], [sonorant], [continuant], [nasal], [liquid], [approximant], [lateral]. We find evidence of OCP restrictions within the category of liquids, e.g. in liquid dissimilation in Latin (Kenstowicz 1994, in Van der Torre 2003). In Latin the adjectival suffix -*alis* dissimilates to -*aris* when the preceding stem contains /l/[4]:

(9)	a.	nav + alis	navalis	*naval*	b. sol + alis	sola**r**is	*solar*
					milit + alis	milita**r**is	*military*

To avoid identical consonants, lateral /l/ is replaced by a consonant from the same category, the category of liquids, but with a different manner of articulation, viz. rhotic /r/. The manners of articulation of the two sounds differ in the manner feature [lateral]. In the history of the English language, suffix-boundaries were also influenced by OCP; -*ity* and -*ness* are two possible suffixes for nominalization of an adverb. Latinate stems occur freely with both suffixes, but Latinate stems ending in /t/ are significantly underrepresented with -*ity* (O/E = 0.038)[5]. Similarly the adverbial suffixes -*ish* and -*al* are influenced by OCP: fish → *fishish, fishy; scale → *scalal, scalar (Berkley 1994b). Identical consonants are avoided.

There are also examples of an obstruent/sonorant distinction in OCP as a secondary constraint on place harmony in languages of the world. As we saw above, McCarthy had recognized that there were no or very weak OCP-place restrictions among coronals when all coronals were taken into account. However, division of the category of coronals into obstruents and sonorants showed OCP-place restrictions within the separate categories. In addition to the four place categories he therefore introduced the features [consonant] and [sonorant] as relevant features for OCP coronal in Arabic. Similar restrictions are found in English and Russian. In English, coronal obstruents are underrepresented with coronal obstruents, not with coronal sonorants and vice versa (Berkley 1994a). Padgett (1995) recognizes "place-stricture interaction" in Russian. He distinguishes 3 coronal classes:

[4] The process is blocked when the preceding stem has /l/ and /r/ (in that order); litor + alis remains litoralis, 'of the shore' (Van der Torre 2003)

[5] The O/E rate is the number of observed combinations divided by the number that would be expected if consonants co-occurred freely. For a more thorough explanation of O/E I refer to Chapter 4, section 1.

(10) coronal classes in Russian

[+son]	l, l’, r, r’, n, n’
[-son, -cont]	t, t’, d, d’
[-son, +cont]	s, s’, z, z’

Within these classes there are co-occurrence restrictions, among these classes the limitations are less strict. In Russian, coronal subsidiary manner features are both [sonorant] and [continuant].

On the basis of these examples there is reason to include manner features in the OCP analysis.

2.1.3.3 The feature [voice]

There are reasons to differentiate consonants not only by [place] or [manner], but by the feature [voice] as well. Yip (1988) investigated how OCP can block rules that would otherwise have been applied. An example of this is Rendaku voicing in Japanese. Rendaku voicing requires voicing in the second element of a compound. This process is blocked when the second element already contains a voiced obstruent. This OCP-like restriction is known as Lyman’s Law. An example is found in (11) below (Yip 1988, p 68):

(11)

a.	maki + sushi	→	maki**z**ushi	*rolled sushi*
b.	kami + kaze	→	kamika**z**e	*divine wind*

Voiceless /s/ becomes voiced in *maki**z**ushi;* voiceless /k/ in *kamika**z**e* remains voiceless, because the second element already contains a voiced segment /z/[6].

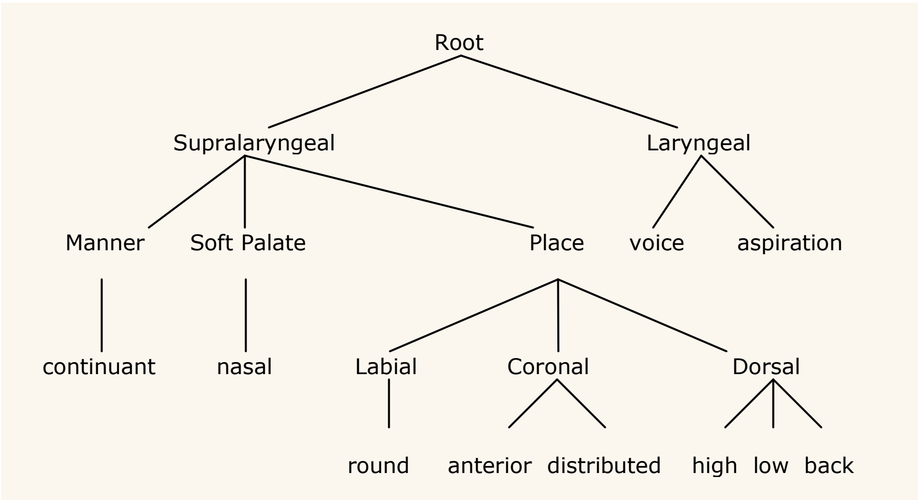
[6] It must be noted that Rendaku voicing in Japanese is also influenced by an obstruent/sonorant distinction. Alderete (1997) noticed that Rendaku voicing is blocked in stems with a post-initial voiced obstruent, but not with a post-initial sonorant:

a.	kami + kaze	kamika**z**e, *kamigaze	*divine wind*
b.	ori+kami	ori**g**a**m**i	*folding paper*

The stem *kaze* has a voiced obstruent and Rendaku voicing is blocked; the stem *kami* has a sonorant as second element and Rendaku voicing takes place. The fact that voicing effects do not affect sonorants indicates that in Japanese OCP voicing is connected with the manner feature [sonorant] as well.

Yip analyzes long-distance OCP as a restriction on adjacent nodes in a tree.: "...that distinctive features make up a tree structure whose terminal nodes are single features that in turn are dominated by higher-order articulator nodes and class nodes, until eventually the whole tree is dominated by the root node ...". The tree structure is in (12) below:

(12) Structure of the feature system



For OCP to be effective segments must be adjacent on some tier in the tree structure, e.g. the root tier, the place node or the coronal node. Intervening segments that are unspecified for the relevant feature are skipped when determining adjacency. "If a node/feature is missing from a given segment, the comparable nodes/features of the next-door segments will be adjacent, since nothing will intervene" (Yip 1988, p 71). Thus, in kami + kaze the stem *kaze* is analyzed on the voice node as: /k/ - [-VOI], /a/ - unspecified, /z/ - [+VOI], /ə/ - unspecified. On the voice node /k/ and /z/ are adjacent, because both are specified for the feature [voice]. Even though on the surface the segments are not adjacent, on a deeper level they are. Lyman's Law in Japanese is clearly an OCP effect connected with the feature [voice].

In this section I discussed three features in which OCP effects have been proven to exist in languages of the world, viz. [place], [manner] and [voice]. It is clear that OCP not only affects identical segments, but that it is also active between segments with identical features. It is also the case that OCP is present across segments. With the introduction of features as secondary OCP markers in e.g. coronal combinations, it appeared that the concept of absolute OCP had to be abandoned.

2.1.4 Gradience in OCP

McCarthy's observation that the OCP was not always categorical, i.e. active in all circumstances with adjacent consonants with identical features, was further investigated. It appeared that the non-categorical nature of the OCP – the "tendency" McCarthy had observed in Arabic – was also present in other languages. The OCP phenomenon turned out to be gradient. A grammar with gradient constraints recognizes degree of violation of a constraint. Forms that violate a particular constraint to a lesser degree are more frequent than forms that violate the same constraint to a greater degree. The word *constraint* is here used in the sense of a limitation of the language. Constraints define the well-formedness of an utterance.

In experiments in several languages, (English, Hebrew, Jordanian Arabic[7]), it was shown that speakers of a language with gradient OCP restrictions can make judgements not only about acceptable and unacceptable forms in their language, they can also differentiate degrees of acceptability in possible as well as in impossible forms. When not forced to choose, speakers of a language will make categorical distinctions, when forced to choose, gradient distinctions are made (Alderete & Frisch 2006, Coetzee 2006). Autosegmental phonology could not cope with gradience. What was needed was a model of gradient linguistic constraints and constraint combinations. Several attempts at explaining gradience were made. The next section will discuss the influence of different properties of words and segments on gradience in OCP.

[7] The experiments will be dealt with below in 1.2, the psychological reality of OCP.

2.1.4.1 The influence of features on gradience

In order to be able to account for gradience, other than place features were included in the analysis. McCarthy had already added the features [consonant] and [sonorant] as elements in the OCP coronal restriction. Pierrehumbert (1993) introduced degree of similarity based on features to explain a systematic range of intermediate OCP effects. It was also observed by Frisch in the Arabic roots he had studied, based on the dictionary, that identical pairs showed stronger OCP-restrictions than homorganic non-identical pairs (Frisch 1997). This led to the development of the similarity matrix.

The similarity matrix

The concept of feature similarity was developed by Frisch, Broe and Pierrehumbert, FBP, in their study of Arabic (1997). They claimed that the aim of OCP is similarity avoidance. Pierrehumbert and Frisch developed a similarity matrix based on the features [consonant], [sonorant], [continuant], [nasal], [lateral], [labial], [dorsal], [coronal], [anterior], [back], [voice]. Not all features are relevant for all segments. The feature [nasal], for instance, is not relevant for obstruents, because there are no obstruents which can have the feature [+NAS]. When a feature is not relevant for a particular segment, this segment is said to be unspecified for the feature, hence obstruents are said to be unspecified for the feature [nasal]. This means that segments can be marked + or - for a particular feature, or they can be unspecified for that feature. The similarity matrix is based on marked features.

The formula for the similarity matrix was presented as below:

(13)

$$\text{Similarity} = \frac{\text{Shared features}}{\text{Shared features + non shared features}}$$

An aspect that proved relevant in connection with the similarity matrix, was the markedness of features withìn the feature classes. For instance, with regard to the feature class of sonorants, co-articulation in nasal-obstruent clusters is very common cross-linguistically, whereas, liquid assimilation is rare, e.g. Latin luna + lis → luna***r***is (Alderete & Frisch 2006). This indicates a difference in markedness within the category of sonorants. In connection with OCP it appeared that there was not only a difference in [place] and [manner], but also within the place and manner categories. Hence the similarity matrix was adapted using natural classes instead of single features: a natural class is a set of segments with a feature shared by all members of that class.

The formula became:

(14) $$\text{Similarity} = \frac{\text{Shared natural classes}}{\text{Shared natural classes + non shared natural classes}}$$

"Identical consonants have similarity 1, as they participate in exactly the same natural classes. Highly dissimilar consonants share very few natural classes (because they share very few features) and have very low similarity" (Frisch, Broe, Pierrehumbert 2004, p 198). An example from FBP will illustrate how the matrix is applied. "The natural classes given are only those whose definition includes the labial place feature. The pair /f, m/ share 2 such natural classes, namely {b, f, m, w} (the labials) and {b, f, m} (labial consonants). They have 7 non-shared classes: {b, f} (obstruents), {f, w} (continuants), {f} (voiceless continuants), {b, m, w} (voiced), {b, m} (voiced stops), {m, w} (voiced sonorants) and {m} (nasals). The similarity of /f, m/ is 2/9 = 0.22 by [the equation above, ed]. The pair /b, f/ share 3 classes, {b, f, m, w}, {b, f, m} and {b, f}. They have 5 non-shared classes: {f, w}, {b, m, w}, {b, m}, {b, w}, {b}. The similarity of /b, f/ is 3/8 = 0.38" (FBP 2004, p 199). Frisch, Broe and Pierrehumbert investigated the effect of the similarity matrix for Arabic and found that it had predictive value for Arabic verbal roots.

The similarity matrix was applied to English by Berkley (2000). In her dissertation she analyzed English and Latin OCP restrictions. Both English and Latin turned out to show labial and dorsal OCP effects in adjacent segments as well as across segments. She also observed that Latin had coronal OCP, whereas English only has coronal OCP within the categories of obstruent and sonorant. She connects this with the number of classes that could theoretically be shared – for Latin coronals the number of classes is 24, for English coronals the number of classes that could theoretically be shared is 52. She calculated similarity values on the basis of the number of natural classes occurring between two consonants. An example of the similarity values for English labials used by Berkley is in the table below (Berkley 2000, p 55). The numbers indicate degree of similarity, 1 represents 100% similarity.

(15) Similarity values for English labials

	p	b	f	v	m	w
p	1					
b	0.45	1				
f	0.33	0.17	1			
v	0.18	0.36	0.38	1		
m	0.27	0.45	0.09	0.18	1	
w	0.20	0.27	0.11	0.22	0.50	1

Berkley divided the similarity values of all homorganic consonant pairs in monomorphemic monosyllables into 10 categories and calculated O/E scores. O/E scores express the difference between observed scores and expected scores in a database. When observed score and expected score coincide the O/E score is 1; scores below 1 indicate underrepresentation, scores above 1 indicate overrepresentation. For an extensive explanation of O/E I refer to Chapter 4, section 4.1. The table below gives Berkley’s results per similarity category (Berkley 2000, p 57).

(16) O/e ratios for homorganic consonant pairs from English monomorphemic monosyllables, grouped by similarity.

Table 3.5. O/e ratios for homorganic consonant pairs from English monomorphemic monosyllables, grouped by similarity			
Similarity Values	Observed	Expected	O/e ratios
0 - 0.10	1103	956.0	1.15
0.11 - 0.20	157	170.2	0.92
0.21 - 0.30	100	109.0	0.92
0.31 - 0.40	23	37.9	0.61
0.41 - 0.50	88	126.9	0.69
0.51 - 0.60	0	0.0	…
0.61 - 0.70	0	0.0	…
0.71 - 0.80	18	40.8	0.44
0.81 - 0.90	0	0.0	...
0.91 - 1.00	42	90.3	0.47
total	1531	1531.0	

It can be seen from the table that "as similarity between consonants increases, the o/e ratios decrease" (Berkley 2000, p 56). She found that the greater the similarity between two consonants in English monomorphemic monosyllables, the stricter the OCP constraint was applied.

For both Arabic and English the similarity matrix proved a useful prediction device in accounting for gradience in OCP.

Total identity versus partial identity

For both Classical Arabic and English gradience in OCP could be accounted for with the help of the similarity matrix. However, the similarity matrix in the form discussed above was not applicable to all languages investigated. There are differences between languages in the strictness of OCP with regard to completely identical segments. If we apply the similarity matrix, total identity must be considered the worst possible violation of the OCP constraint. This is indeed what we find in several languages. The literature on Arabic shows strictest OCP for geminates. Berkley (2000) stated that for English the similarity matrix is applicable; there is no significant difference between identical pairs of consonants and almost similar pairs of consonants in monosyllables, O/E rates of 0.47 and 0.44 respectively. We can therefore assume for English that total identity is included in the OCP restrictions. Van de Weijer (2003) illustrated for Swedish that co-occurrence of two identical consonants in a monomorphemic word is rare. However, a language like Yamato Japanese does not have a problem with total identity. "It seems that total identity provides an escape hatch from the OCP" (Kawahara 2005). Also in Javanese, (Yip 1989) and Muna, (Coetzee and Pater 2008) total identity is exempt from OCP restrictions. The role of total identity vs. degree of similarity cannot be analysed uniformly for the languages discussed in the literature. In some languages it must be considered a phenomenon separate from partial or featural OCP. These languages will make use of a constraint allowing total identity, which outranks the OCP constraint(s).

2.1.4.2 The influence of distance on gradience

The relevance of distance between the segments as a defining aspect of the strength of OCP-effects was taken up by Berkley (1994a). She observed for English that, next to similarity, distance between segments was a determining factor. Frisch (1997) noted that both in non-concatenative languages, such as Arabic, Hebrew, Jordanian, as well in concatenative languages, e.g. English, not only strictly surface adjacent consonants were affected by the OCP, but that the OCP was also active in consonant pairs separated by one or more intervening segments. Berkley (2000) established for English that distance between two segments in a linguistic domain influenced gradient OCP-place effects. In her view this was connected with the salience of two consonants ***as a pair*** : "the salience of two consonants as a pair contributes to their vulnerability to the OCP" (Berkley 2000, p 178). She claims that "the closer two consonants are, the more salient they are as a pair" (Berkley 2000, p 197). It stands to reason that consonants with one segment between them are more salient as a pair than consonants with more intervening segments, hence more prone to OCP. It must be noted that this is a language-specific aspect of OCP. There are also languages in which OCP is only active in strictly surface adjacent consonants, e.g. Cambodian (Yip 1989).

2.1.4.3 The influence of salience on gradience

This section will deal with salience as a defining factor in the gradience of OCP. The most prominent consonants are the most salient consonants. Salience has been claimed to be connected with distance between two segments, as indicated above, as well as with prosody, i.e. position in the word, stress, and sonority. Berkley posits that the initial syllable is the most salient syllable in the word, with onset position being more salient than coda position. The number of activated consonants is smaller when only word onset is concerned; following consonants are embedded in more intricate constructions, hence the consonant at word onset has the most prominent position and is therefore the most salient. In syllables further down in the word the onset position of the syllable is again more salient than coda position. The salience of a consonant pair is also connected with stress. Stressed syllables are more salient than unstressed syllables.

Position in the word

With regard to position in the word, Berkley found that in disyllables, word onset pairs, i.e. pairs with both consonants in the initial syllable of the word, /**k**æ**k**-təs/, were affected more than word coda pairs, i.e. pairs with both consonants in the final syllable of the word, /æl-**b**ʌ**m**/, whereas onset-onset pairs, i.e. pairs with one consonant as onset of the initial syllable and one consonant as onset of the second syllable, /**b**e-**b**i/*,* were in between. As indicated above, the consonant in word onset position is the most salient consonant; across syllables the salience decreases by the intervention of a syllable boundary, indicating more distance between the two consonants than between consonants within the first syllable. For English the most salient consonant pairs are most likely to be affected by OCP. In accordance with this, Berkley found a weaker OCP effect across syllable boundary than within the initial syllable. The second onset is less salient than the first; (second) syllable onset and coda were third in line in strictness.

Word stress

The salience of a consonant pair is also connected with stress. Stressed syllables are more salient than unstressed syllables, hence more prone to OCP. The OCP-place restriction was indeed found to be stronger in English in stressed syllables than in unstressed syllables (Berkley 2000).

Sonority

Elmedlaoui (1995) discusses an additional sub-regularity in consonant co-occurrence patterns in Arabic. This sub-regularity is connected with sonority. He discovered that, when the first of two consonants was more salient than the second, i.e. less sonorous, OCP was stricter; falling sonority is preferred over flat or rising sonority in homorganic consonants.

We have seen in this section that OCP is often a gradient phenomenon. Gradience has been illustrated to be connected with features, distance, position in the word, stress and sonority. Especially the recognition of the influence of features as secondary OCP markers was important.

2.1.5 Optimality Theory

This section will complete the survey of the history of OCP with analyses applied in recent phonological theory, Optimality Theory. Optimality Theory (Prince and Smolensky 1993), introduced constraints determining phonotactic well-formedness. Optimality Theory, henceforth also referred to as OT, makes use of universal constraints which are ranked in a language-specific dominance hierarchy, denoted as $C_1 >> C_2$. The constraints have an evaluative task with respect to phonological representations. There are two types of constraints, markedness constraints and faithfulness constraints. Markedness constraints require output forms to agree with a specific structure in order to be well-formed; they bear on the surface structure only. Faithfulness constraints require faithfulness between input, i.e. the lexicon, and output. Languages will give priority to markedness in some aspects and to faithfulness in other. An example in Dutch of markedness outranking faithfulness is found in final devoicing, denoted by the constraint NoVoicedCoda – coda consonants must be voiceless. In accordance with this the /d/ in /hɑn**d**ən/ 'hands' is pronounced in agreement with the input /hɑnd/ + plural /ən/; singular input /hɑnd/ is realized as /hɑn**t**/ with final /t/ on account of the final devoicing constraint outranking the faithfulness constraint. The interaction between markedness and faithfulness, specified in the hierarchical ranking between the two types of constraints, will determine what are grammatical forms in a particular language and what are not.

One of the central notions of OT is that constraints are violable. The candidate with the least offensive violations in a hierarchically ranked constraint set is the optimal output in that language for a particular input. A constraint outranking other constraints has priority over the other constraints, no matter how many or how few violations of lower ranked constraints follow. Phonotactic well-formedness is about the "best possible" form, not about the perfect form. A grammatical form is thus a form with minimal constraint violation and not a form with no constraint violation.

The differences in rankings across languages account for language-specific properties. All universal constraints are present in the rankings of the languages of the world. Constraints that are inactive in a particular language are simply ranked so low that their effect does not emerge in the language.

2.1.5.1 Categorical OCP in OT

A fixed ranking of individual constraints will create an all or nothing situation. "...the grammar bans pairs of consonants that contradict them, and permits ones that do not" (Coetzee and Pater 2008). Judgements of well-formedness in this system are categorical. Coetzee and Pater (2008) introduce OCP constraints that correspond to McCarthy's subdivision for Arabic (McCarthy 1988), including featural markedness of coronals:

(17)

OCP-LAB	Adjacent labials are prohibited
OCP-DOR	Adjacent dorsals are prohibited
OCP-PHAR	Adjacent pharyngeals are prohibited
OCP-COR	Adjacent coronals are prohibited
OCP-COR$_{[\alpha SON]}$	Adjacent coronals agreeing in [+/-sonorant] are prohibited

I here add the corresponding faithfulness constraints:

(18)

IDENT- LAB	Input labials must be realized in the output
IDENT- DOR	Input dorsals must be realized in the output
IDENT- PHAR	Input pharyngeals must be realized in the output
IDENT- COR	Input coronals must be realized in the output

A constraint ranking that will capture the phenomenon of Arabic, viz. categorical OCP restrictions on labials, dorsals and pharyngeals, as well as on coronals provided they share the feature [sonorant] is in (19).

(19) OCP-LAB, OCP-DOR, OCP-PHAR, OCP-COR$_{[\alpha SON]}$ >> IDENT-LAB, IDENT-DOR, IDENT-PHAR, IDENT-COR >> OCP-COR

In this constraint ranking all adjacent labials, dorsals, pharyngeals and same-sonority-coronals will be avoided. (A comma between constraints indicates that they are not ranked relative to each other). With OCP-COR$_{[\alpha SON]}$ outranking faithfulness, all adjacent same-sonority-coronals are out. The faithfulness constraint IDENT-COR guarantees that all other coronals will be realized faithfully. The general constraint OCP-COR, prohibiting any adjacent coronals, is outranked by the IDENT-COR constraint and OCP-COR is thus invisible in the language. This explains the non-existence of restrictions on adjacent coronals in Arabic other than those sharing the feature [sonorant]. The example clarifies that with OT categorical distinctions can be made.

2.1.5.2 Gradient OCP in OT

Optimality Theory

What we came across above, is that OCP is a gradient phenomenon in various languages. When constraints apply in the categorical fashion, as above, it is impossible to create a grammar compatible with gradient OCP. Coetzee (2007) investigates a processing difference between ungrammatical forms that differ in terms of ungrammaticality[8]. He investigates native speakers' judgements of nonce words of the type stVt - /stʌt/, skVk - /skɛk/, spVp - /spɑp/. The forms violate two constraints, viz. *tVt – do not allow two /t/s separated by only a vowel, *kVk – do not allow two /k/s separated by only a vowel, *pVp – do not allow two /p/s separated by only a vowel, and a separate constraint $*[s+stop]_{\sigma}$ – do not allow the sequence [s+stop] in a syllable. Single constraints can be linked to other single constraints in a particular domain (word, syllable, morpheme) by Local Conjunction. "Under *Local Conjunction,* two constraints are conjoined as a single composite constraint which is violated if and only if both of its components are violated within some domain" (Kager 1999). Coetzee (2007) makes use of a suggestion by Ito and Mester (1998), who propose a limit on a ranking of locally conjoined constraints, the 'ranking preservation'. This principle preserves the established ranking between constraints in a ranking with conjoined constraints. That is, a constraint order of e.g. *pVp >> *kVk >> *tVt, as established for English by Berkley (1994ab, 2000), will be preserved when the constraints are conjoined with e.g. $*[s+stop]_{\sigma}$. This will lead to the ranking *spVp >> *skVk >> *stVt. Since words of the type stVt are freely allowed in English, the constraint *stVt is assumed to be ranked below the faithfulness constraint IDENT-PLACE. The tableau in (20) gives a mini-grammar of English based on the above conjoined constraints.

[8] Coetzee (2007) will be dealt with more extensively in 2.2.3 on the psychological reality of OCP in English.

(20) Mini-grammar of English [sCVC]-forms

	*spVp	*skVk	IDENT-[place]	*stVt
/stVt/ ☞ stVt				*
stVk			*!	
/skVk/ skVk		*!		
☞ skVt			*	
/spVp/ spVp	*!			
☞ spVt			*	

This tableau predicts that a /stVt/-input will map faithfully onto itself and is a possible word in English, and that /skVk/ and /spVp/ do not map faithfully onto themselves and are not possible words of English. The tableau will generate categorical distinctions. However, it can also be used for gradient well-formedness between two ill-formed words. For the above ranking a comparative tableau, comparing different input~output mappings, rather than different output candidates for one input, can be composed. The comparison is shown in the *comparative tableau* in (21) (Coetzee 2006, p 23).

(21) Comparing non-words in English

	* spVp	*skVk	IDENT-[place]	*stVt
1 /stVt/ → [stVt]				*
2 /skVk/ → [skVk]		*		
3 /spVp/ → [spVp]	*			

This tableau indicates how these three candidates are related to each other. "Since /stVt/ → [stVt] violates the lowest ranking constraint, it is rated best, ... Since /spVp/ → [spVp] violates the highest ranking constraint, it is rated worst of all...". In this manner constraints that are ranked below constraints ruling out impossible forms can be ranked relative to each other, so that an OT grammar can establish the ill-formedness relationship between two ungrammatical forms and impose an ill-formedness hierarchy. In this manner gradient differences between two ill-formed words can be made.

Harmonic Grammar

A different approach to gradience is found in Harmonic Grammar (Smolensky and Legendre 2003). Harmonic Grammar, henceforth also referred to as HG, is a theory of grammar related to OT. In HG the optimal candidate is the candidate with the highest Harmony score. Harmony is the sum of weighted constraint scores. An example from Coetzee and Pater (2008) will illustrate how the evaluation works. Coetzee and Pater compare a restriction against the co-occurrence of homorganic consonants in the root morphemes of Muna, a Western-Austronesian language and Arabic. In both Muna and Arabic the restriction is gradient, depending on place of articulation, as well as on similarity of homorganic consonants in terms of other features. Coetzee and Pater divide OCP up into a family of more specific OCP constraints in OT. They make use of the constraints mentioned in (16) above on categorical OCP, viz. OCP-LAB, OCP-DOR, OCP-PHAR, OCP-COR[9]. As indicated above, these constraints lead to categorical distinctions; a combination is either ìn or out. In order to account for gradience Coetzee and Pater include the features [sonorant], [continuant], [voice] in their comparison of Muna and Arabic. In Muna, homorganic pairs that share the features [sonorant], [continuant] or [voice] occur less freely than pairs that do not agree in these features. All three features turn out to play a role. None of the three is more dominant than another in Muna. In Arabic the three features also play a role, but there is a clear hierarchy. Homorganic pairs of consonants that agree in the feature [sonorant] are by far more restricted than those that share [continuant], whereas [voice] is a very weak determining constraint in Arabic place-OCP.

In order to account for gradient OCP-effects Coetzee and Pater make use of weighted constraints. In a standard OT constraint tableau the most harmonic candidate wins. In a constraint tableau with weighted constraints, harmony is calculated by multiplying each

[9] The division is based on McCarthy 1988.

score by the weight of the constraint and summing the weighted scores as in tableau (18) below. The example is from Coetzee and Pater, 2008. Constraint weights are displayed in the top row, and the candidates' Harmony scores appear in the rightmost column. Each candidate mapping violates one of the constraints. Because the first constraint has a higher weight the candidate violating the second constraint has a higher Harmony score. The constraint violations are expressed by '*'.

(22)

Weight	2	1	*Harmony*
Input-1	Constraint 1	Constraint 2	
☞ Output 1-1		*	-1
Output 1-2	*		-2

A difference between constraint weighting and constraint ranking is that constraint weighting allows for cumulative constraint interactions, as illustrated in tableau (23).

(23)

Weight	2	1	*Harmony*
Input-2	Constraint 1	Constraint 2	
Output 2-1		***	-3
☞ Output 2-2	*		-2

From this table it is clear that three violations of the lighter Constraint 2 are worse than one violation only of Constraint 1. Whereas in a categorical system a violation of Constraint 1 would have eliminated the second output, the result here is, that the effect of Constraint 2 is stronger than that of Constraint 1, hence the candidate violating the higher ranked, but lower weighted score, is the optimal output. It is assumed by Coetzee and Pater that the constraint set for OCP is universal, and that phonological learning requires finding the appropriate weighting of the constraints. In this way language-specific grammars result from the universal constraint set. The inventory of the features influencing OCP is universal, the relative strength of the separate features is language-specific. For Muna and Arabic the constraints used are all of the type as indicated in (24) on dorsals (Coetzee and Pater 2008, p 310):

(24) Constraints against homorganic sequences

*K-K	Assign a violation mark to a sequence of consonants in which both consonants are dorsal
*K-K-Son	Assign a violation mark to a sequence of consonants in which both consonants are dorsal and both consonants have the same specification for sonorancy
*K-K-Stric	Assign a violation mark to a sequence of consonants in which both consonants are dorsal and both consonants have the same stricture
*K-K-Vce	Assign a violation mark to a sequence of consonants in which both consonants are dorsal and both consonants have the same specification for voice
*K-K-Emph	Assign a violation mark to a sequence of consonants in which both consonants are dorsal and both consonants have the same specification for emphatic
*K-K-Pre	Assign a violation mark to a sequence of consonants in which both consonants are dorsal and both consonants have the same specification for prenasalization
Ident-[place]	Assign a violation mark to a consonant whose place specification is non-identical in Input and Output

Table (25) below illustrates the differences in weight between the same features in Arabic and Muna, thus accounting for cross-linguistic variety.

(25) OCP constraint weights for Arabic and Muna coronals

	Arabic	Muna
*K-K	67.14	53.80
*K-K-Son	73.39	59.84
*K-K-Stric	76.32	72.02
*K-K-Vce	93.75	96.83
*K-K-Emph		
*K-K-Pre		62.29
Ident-[place]	227.25	261.55

The table not only illustrates how cross-linguistic variety can be accounted for. Weighting constraints is also a means to distinguish gradient acceptability of forms within languages. With the language-specific weighting, harmony scores can be calculated for each possible output and gradient acceptability will result. Four HG tableaus for unattested dorsal sequences in Muna, viz. g-ʁ, ŋ-^{ŋ}g, g-^{ŋ}k and ŋ-^{ŋ}k, indicate how the hierarchy between these sequences is determined by the harmony scores. The examples are from Coetzee and Pater 2008, p 323 and 324.

(26) Unattested dorsal sequences in Muna

Weight	261.6	59.8	72.0	96.8	62.3	53.8	*H*
Input: g-ʁ	ID-[PLACE]	*K-K-SON	*K-K-STRIC	*K-K-VCE	*K-K-PRE	*K-K	
g-ʁ		-1		-1	-1	-1	-272.8
☞ d-ʁ	-1						-261.6

Weight	261.6	59.8	72.0	96.8	62.3	53.8	*H*
Input: ŋ-^{ŋ}g	ID-[PLACE]	*K-K-SON	*K-K-STRIC	*K-K-VCE	*K-K-PRE	*K-K	
☞ ŋ-^{ŋ}g			-1	-1		-1	-222.6
n-^{ŋ}g	-1						-261.6

Weight	261.6	59.8	72.0	96.8	62.3	53.8	*H*
Input: g-^{ŋ}k	ID-[PLACE]	*K-K-SON	*K-K-STRIC	*K-K-VCE	*K-K-PRE	*K-K	
☞ g-^{ŋ}k		-1	-1			-1	-185.6
d-^{ŋ}k	-1						-261.6

Weight	261.6	59.8	72.0	96.8	62.3	53.8	*H*
Input: ŋ-^{ŋ}k	ID-[PLACE]	*K-K-SON	*K-K-STRIC	*K-K-VCE	*K-K-PRE	*K-K	
☞ ŋ-^{ŋ}k			-1			-1	-125.8
n-^{ŋ}k	-1						-261.6

These tableaux predict that native speakers of Muna would accept the above forms in the order in which harmony scores are arranged, i.e. -125.8 >> -185.6 >> -222.6 >> -

272.8, which means that the above forms are accepted in the order ŋ-^{ŋ}k >> g-^{ŋ}k >> ŋ-^{ŋ}g >> g-ʁ.

Thus far we have seen the existence of OCP restrictions across languages. The OCP-constraints are universal, the ways in which they are applied are language-specific. There is both absolute OCP and gradient OCP. The recognition of the influence of features was a major contribution to understanding the gradient nature of the phenomenon. The similarity metric was an important attempt at accounting for gradience. It turned out that it could not be applied uniformly for all languages discussed; especially identical segments sometimes follow rules of their own. Another aspect that came up in the course of the discussion on the strength of OCP is the matter of linear distance. OCP not only affects strictly (surface) adjacent consonants, but also consonants separated by one or more segments. Thus the language-specific nature of OCP is also connected with distance between the segments. Also salience was found to be a defining aspect of gradient OCP. The more salient a consonant is – or, according to Berkley, the more salient two consonants are as a pair – the stricter OCP is applied. Salience is connected with prosodic properties of the word, such as the position of the two consonants in the word, word stress and sonority.

2.2 The psychological reality of OCP

The literature discussed so far investigated OCP as a characteristic of several languages. The data were based on counts from dictionaries and comparable word lists. The recognition of OCP restrictions is thus based on the lexicon only. The underrepresentation in the lexicon of words which violate the OCP-place constraint is taken as evidence for the existence of the constraint. It may, however, well be the case that a dictionary does not reflect synchronic grammatical knowledge: it may contain borrowings, obscure forms, obsolete forms. These are not part of the productive grammatical knowledge of native speakers. Another means, therefore, of addressing the question is to find out whether OCP-place has psychological reality for the speakers of a language. This section will consider studies on the psychological significance of OCP in Jordanian Arabic, Hebrew, English and Dutch.

2.2.1 Jordanian Arabic

Frisch and Zawaydeh (2001) took up McCarthy's findings for Arabic and investigated whether OCP-root was indeed a psychological reality for speakers of Jordanian Arabic. This would mean that the root would define whether a particular combination of consonants was acceptable or not, irrespective of the nature and number of intervening segments. Frisch and Zawaydeh performed an experiment to test the grammar of native speakers in the mental lexicon. In the experiment they attempted to simultaneously examine three independent possible sources of influence on co-occurrence restrictions, viz. similarity to existing words, phonotactic probability (based on onset probabilities and rhyme probabilities and combination probabilities) and phonotactic constraints. Co-occurrence restrictions were calculated by O/E scores.

(27) O/E scores for roots in Arabic beginning with /t/: (Frisch and Zawaydeh 2001, p 95):

Root type	O/E
ttC	0
tsC	0.22
tʃC	0.84
tnC	1.24
tkC	1.94

For the experiment they presented subjects with novel verbs for judgement of well-formedness. They created 254 novel verbal roots; the verbs were presented orthographically. There were equal numbers of verbs with and verbs without OCP-place violation, with high or low phonotactic probability and dense or sparse lexical neighbourhood densities (similarity to existing roots in the lexicon).

It became apparent that Expected Probability and Neighbourhood Density did not show any influence. On the other hand, there was considerable influence of OCP-violation on the ratings of the participants. Well-formedness for OCP-place was gradient. It was not the case that all OCP-place violations were rated as less acceptable than all non-violation stimuli; a striking observation was that OCP-place violations involving identical consonants were not judged as bad as those with highly similar but non-identical consonants. This is contrary to the lexical pattern and contrary to predictions of the similarity metric by FBP, and is not explained by Frisch and Zawaydeh other than that it could be a side-effect of the test – too few stimuli in the test for this result to be reliable – or it could be connected with a difference between complete identity and partial identity comparable to the difference found in other languages. The experiment did show that OCP-place is a psychologically real constraint for speakers of Jordanian Arabic.

2.2.2 Hebrew

Research on the psychological reality of OCP was also performed on Hebrew, another non-concatenative language. In this paragraph two experiments will be discussed, i.e. one by Berent and Shimron (1997) and one by Coetzee (2006).

Berent and Shimron set out to prove that the root morpheme must be considered a distinct constituent in Hebrew. For this purpose they used the Obligatory Contour Principle. Hebrew verbs consist of a root pattern, a sequence of usually three consonants. These are combined with a word pattern containing vowels and sometimes additional consonants; the root /k t b/ e.g. represents 'write'; it can yield about 120 morphologically related words by combining it with word patterns and adding inflectional systems. In Hebrew a root morpheme is not recognizable orthographically or phonologically – different (numbers of) segments can separate the phonological basis of the phoneme. Berent and Shimron asked participants in their experiment to determine the extent to which a nonce word sounded like a possible Hebrew word. For this purpose they used two rating tasks, one where participants had to rate three words in accordance with acceptability, and one in which participants were asked to rate single words for acceptability on a scale of 1 to 5, indicating the extent to which a word sounded like a possible Hebrew word.

They demonstrated that notwithstanding a total lack of orthographic or phonological evidence, Hebrew speakers showed influence of a phonological constraint that avoids similarity in the root morpheme. With no relation to phonological or orthographic evidence in the language, the source of the constraint, according to Berent and Shimron, must be "tacit linguistic competence" (Berent and Shimron 1997, p 56).

Coetzee (2006) performed a word-likeness rating experiment to check the psychological reality of OCP and possible gradient effects of OCP in Hebrew. He used two types of experiments, one that asked subjects to rate a single nonce word on a particular scale, the other offering more tokens at a time and asking subjects to rate them according to acceptability as a native word. The first experiment, "word-likeness rating", led to categorical acceptance or rejection, although intermediate ratings were available. This proved that language users can make categorical distinctions. In the second experiment, "comparative word-likeness", subjects are required to compare words and select the one that is most word-like. The second experiment thus enforces a gradient scale. Participants in Experiment 2 also categorically rejected roots with two identical consonants, but showed a clear preference for a root with three different consonants. Native speakers of Hebrew not only show evidence of the psychological reality of OCP, they also reveal gradient OCP-effects, they make finer distinctions within the set of grammatical forms.

These two experiments show that also in Hebrew, OCP is a psychological reality for native speakers of the language.

2.2.3 English

There is not only evidence of the psychological reality of OCP in non-concatenative languages. Also concatenative languages show evidence of the psychological reality of OCP restrictions. I will here take up two concatenative languages, viz. English and Dutch. Coetzee (2005) investigated for English if OCP is "an active part of the grammar, or is it simply a stative statistical generalization over the structure of the lexicon?" (Coetzee 2005, p 4). He tested whether OCP constraints would influence perception of consonants. He hypothesized that words like [skɑk] and [skɑp], with the final consonant acoustically realized between /k/ and /p/, would show a bias in subjects' perception towards final /p/ (no OCP violation); similarly [spʌp] and [spʌk] were expected to show a bias towards /k/. Although words like 'state' freely occur in English, he also checked for restrictions on coronals with nonce words like [stɛt] and [stɛk], where he hypothesized a bias towards /k/. Three sets of continua were constructed. The tokens on these continua were presented to listeners who were asked to identify the final consonant. The expected perceptual biases on the three sets of continua are in table (28).

(28) Predicted perceptual biases based on OCP restrictions

Continuum	Predicted bias	Comment
[skɑp]~[skɑk]	[p]	Because *[skɑk]
[spɑp]~[spɑk]	[k]	Because *[spɑp]
[spʌp]~[spʌt]	[t]	Because *[spʌp]
[stʌp]~[stʌt]	[p]	Uncertain, because [stʌt] legal
[skɛk]~[skɛt]	[t]	Because *[skɛk]
[stɛk]~[stɛt]	[k]	Uncertain, because [stɛt] legal

All tokens were monosyllabic nonce words, offered in a sentence *John said .... again to me.* All items were controlled for Neighbourhood Density and Transitional Probabilities. The only relevant difference between the words in each condition was whether the word violated OCP or not. He found evidence in favour of OCP in all three place categories. The results are in table (29).

(29) Total percentage of responses per segment

Continuum	Bias predicted toward	Responses	One-tailed t-test
[skɑp]~[skɑk]	[p]	% [p] = 48	*t*(14) = 1.71, *p* <.05
[spɑp]~[spɑk]	[k]	% [p] = 39	
[stɛk]~[stɛt]	[k]	% [k] = 55	*t*(25) = 2.82, *p* <.005
[skɛk]~[skɛt]	[t]	% [k] = 47	
[stʌp]~[stʌt]	[p]	% [p] = 47	*t*(25) = 1.90, *p* <.04
[spʌp]~[spʌt]	[t]	% [p] = 40	

In all three cases the predicted bias towards segments avoiding OCP was present. Native speakers of English show evidence of the psychological reality of OCP, not only for labials and dorsals, but also for coronals.

Coetzee (2006) performed the same word-likeness rating experiment for English as he had done for Hebrew above. He used two types of experiments, one that asked subjects to rate a single nonce word on a particular scale, the other offering more tokens at a time and asking subjects to order them according to acceptability as a native word. The first experiment, "word-likeness rating", led to categorical acceptance or rejection, although intermediate ratings were available. This demonstrated that language users can and will make categorical distinctions, when offered a choice. The second experiment, "comparative word-likeness", forces a gradient scale. Participants are offered two grammatical nonce words or two ungrammatical nonce words and they are asked to select the one that is most word-like. In the word-likeness rating experiment with one word at a time they would probably accept or reject both forms. Now they are forced to choose. It appears that they prefer the most well-formed or least ill-formed token of the two, i.e. well-formed and ill-formed in agreement with the grammar of their language. Language users not only make distinctions in grammatical forms, they can and do make finer distinctions within the set of ùngrammatical forms as well. Such gradient effects show that OCP is a psychological reality for speakers of English.

2.2.4 Dutch

Kager and Shatzman (2007) investigated the psychological reality of OCP for native speakers of Dutch. Their analysis of Dutch shows that OCP-LAB exists in Dutch, that it is less strict word-initially than non-word-initially, and that it is stricter when two labial consonants are separated by a single segment than when more segments intervene. They tried to establish whether gradient well-formedness in Dutch could be connected to a specific constraint. They made use of three constraints:

(30)

OCP-LAB	no adjacent labials on the consonantal tier (McCarthy 1988)
*LAB^2	no two labials per word (Ito & Mester 1998)
AlignLab	every labial must be word-initial (new in OCP, Kager & Shatzman 2007)

The first two constraints are anti-co-occurrence constraints, of which the first constraint includes distance between the segments as a factor, the second constraint does not. The alignment constraint captures the preference for harmony in initial, rather than non-initial homorganic consonant pairs. The three constraints partly overlap, but none of the three constraints comprises all three OCP properties.
Kager and Shatzman presented their subjects with nonce words with and without OCP restrictions and measured reaction times. They expected reaction times to nonce word stimuli to be influenced by degree of phonotactic well-formedness and investigated whether phonotactic well-formedness played a role independent of lexical factors, such as Lexical Neighbourhood Density, Cohort Density and Transition Probability. It appeared that reaction times to nonce word stimuli were indeed influenced by degree of phonotactic well-formedness connected with OCP restrictions. It was shown that the abstract notion of phonological constraints on OCP plays a role in Dutch.

We may conclude from this section on the psychological reality of OCP for native speakers of a language, that in diverse languages, both non-concatenative and concatenative, OCP is a psychological reality. Speakers are unaware of the restrictions they impose upon their judgements. Hence, the restrictions can be classified as abstract realities of phonotactic constraints.

2.3 Hypotheses and Predictions

From the literature discussed it has become apparent that OCP is a phenomenon found in many languages. It is also clear that OCP is not limited to strictly adjacent consonants, but that it can also occur across one or more intervening segments. It has different realizations in different languages, ranging from categorical OCP-voice restrictions in Japanese speech production, to gradient rankings of homorganic consonants in the lexicons of several other languages. Gradience has been argued to be connected with degree of similarity of the homorganic consonants, distance between segments and with prosodic salience. OCP has been shown to be a psychological reality for native speakers of a language with OCP restrictions.

In the present thesis I will investigate OCP in Dutch. OCP [place], [manner] and [voice], as they have come up in the present chapter, will be explored. Based on the literature discussed above I expect to find OCP restrictions in the Dutch language. In accordance with Arabic, Hebrew, Javanese, Japanese, Muna and English, which show evidence of OCP across intervening segments, I expect Dutch to show not only OCP in strictly adjacent consonants, but also across one or more intervening segments. I expect distance between the segments to be of influence. Considering the findings for Hebrew, Arabic, Muna and English, which showed evidence of similarity restrictions, I expect featural similarity to play a role as well. With regard to completely identical phonemes I expect OCP in Dutch to be strictest with identical consonants and less so with other homorganic consonants. Supported by the findings for English, Classical Arabic and other Semitic languages, in which prosodic salience was found to play a role, I expect prosodic salience to be another defining factor for Dutch.

Based on the above the following hypotheses will be checked against the facts of Dutch:

Hypothesis 1: OCP is a constraint active in the Dutch language in adjacent consonants as well as in consonants across segments.

Hypothesis 2: The strictness of OCP is related to distance between two consonants, measured in segments.

Hypothesis 3: Degree of similarity influences the strictness of OCP.

Hypothesis 4: The strictness of OCP is connected with prosodic salience.

In Chapter 4 I will analyse the Dutch lexicon with regard to these four hypotheses. In order to be able to fully understand the analyses in Chapter 4, I will give an inventory of Dutch sounds relevant for the present thesis and introduce the basic phonotactics of Dutch. Chapter 3 will discuss these matters.

3 Dutch Phonology

For a proper understanding of this thesis it is important to understand the basic phonotactics of Dutch. This chapter will introduce the phonology of Dutch and consider the main ingredients of the language that should be taken into account when discussing OCP. I will begin with an inventory of Dutch consonants with their features and categorize them in accordance with the OCP categories examined in this thesis. The present chapter will also contain a description of syllable structure in Dutch, including information on language-specific restrictions on consonant combinations.

3.1 Phoneme inventory and OCP categories

3.1.1 Phoneme inventory

Dutch has 43 sounds, 21 vowel sounds and 22 consonants. The feature matrix below contains all Dutch consonants[10].

(31) feature matrix for Dutch consonants

	p	b	t	d	k	g	f	v	s	z	ʃ	ʒ	x	ɣ	m	n	ŋ	l	r	w	j	h
syllabic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+
sonorant	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-
continuant	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+
nasal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-
approximant	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-
lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
voice	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	+	+	+	-
labial	+	+	-	-	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	+	-	-
dorsal	-	-	-	-	+	+	-	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-
coronal	-	-	+	+	-	-	-	-	+	+	+	+	-	-	-	+	-	+	-	-	-	-

[10] The feature matrix is based on Trommelen 1979, Selkirk 1984, and Booij 1995.

3.1.2 OCP categories

For the present thesis the consonants will be regarded in connection with the three major OCP-restriction areas we came across in Chapter 1, viz. [place], [manner] and [voice]. I will here present the three restriction areas and indicate which consonants are included in the analysis and explain why some are left out.

3.1.2.1 Place of articulation

With reference to place of articulation three place categories will be discussed, viz. labial, [LAB], including /p, b, f, v, m, w/, dorsal, [DOR], including /k, g, x, ɣ, ŋ/ and coronal, [COR], including /t, d, s, z, ʃ, ʒ, n, l/. Glottal /h/ constitutes a place category of its own, and is therefore left out of the analysis; palatal /j/ can also be said to constitute a place category of its own and is therefore left out. /r/ has different place realizations dependent both on individual pronunciation preferences of a speaker as well as on surrounding sounds. Individual /r/ can in general be uvular – coming close to dorsal – or (post-)alveolar – belonging with the coronal category. With regard to the influence of surrounding sounds we find that in an onset cluster /r/ is predominantly uvular or post-alveolar; at word end uvular, (post-)alveolar and palatal /r/ occur; word-medially uvular, (post-)alveolar, palatal and retroflex /r/ are found. Thus /r/ has too many places of articulation to be able to classify the sound for a particular place category. /r/, although often included as a coronal sound in OCP literature, is left out in the present thesis with respect to the feature [place].

3.1.2.2 Manner of articulation

With regard to manner of articulation we distinguish different categories for different manner features. The features that will be discussed in this thesis are [sonorant], [continuant], [nasal] and [lateral]. The analyses for [manner] include proper consonants only, hence /w/ and /j/ are left out on account of the feature [-cons].

3.1.2.3 Voice

With respect to the feature [voice], consonants can be classified into three categories: voiceless obstruent /p, t, k, f, s, ʃ, x/, voiced obstruents /b, d, g, v, z, ʒ, ɣ/ and sonorants /m, n, ŋ, r, l, w, j/. There are two ways of approaching the feature [voice] for sonorants. Since sonorants are more sonorous than obstruents, they may be classified for the feature [voice] with voiced obstruents, and together form the category of voiced consonants, or, they can be said to be unspecified for [voice], and must therefore be left out of the analysis for voicing. Both approaches will be considered below. /h/ is an exception to the general voicing rule and is in both approaches left out of the analysis.

It must be noted that diphthongs are considered to consist of two segments. Van der Hulst (1984) analyses Dutch diphthongs as consisting of a vowel and a glide, /ɛi/ = /ɛj/. On account of the vowel-like quality of the glides /w/ and /j/ and their often vowel-like behaviour, he classifies the glides with the vowels as sharing the feature [-consonantal]. In agreement with this view, diphthongs are in the present thesis considered to consist of two [-consonantal] segments, represented as VV. Thus a word like /rɛik/ 'rich', will be analyzed as CVVC and will be included in the category of CXXC. It is also important to be able to establish the exact adjacency of two consonants separated by a single vowel. Dutch monophthongs can be long or short. Since vowel length in Dutch is always connected with the quality of the vowel – there are no same quality vowels which are phonemically differentiated by length only[11] – I consider any single vowel, long or short, to represent a single vocalic segment, V. The quality and quantity of intervening vowels will be dealt with further in section 4.6 below.

[11] With the exception of vowels in loanwords /ɛ:, ɔ:, i:, y:/.

3.2 Syllable structure and phonotactics

This section will take up the structure of the Dutch syllable and discuss properties of Dutch syllables and words. It will present phonotactic constraints connected with syllable structure, such as The Sonority Sequencing Hierarchy, OCP-[place], and assimilation, as well as a phenomenon connected with word structure, viz. extra-syllabicity.

3.2.1 The syllable

Universally syllables consist of two elements, onset and rhyme; the rhyme is subdivided into nucleus and coda.

The onset is an optional constituent in Dutch syllables. If it is there, it may consist of one or two consonants. Any single consonantal segment of Dutch may constitute an onset, except the velar nasal /ŋ/ (Trommelen 1983). Examples are: VC /e:r/ 'honour', <u>C</u>VC /**k** ɑt/ 'cat', <u>CC</u>VC /**pl** ɑs/ 'pool'. At word begin an extra-syllabic /s/ may be added. This will be dealt with below in § 3.2.2.4.

In Dutch monomorphemic words the rhyme "is limited to two (obligatory) positions, VX, where X is phonologically free (but not null)" (Kager & Zonneveld 1986, p 219), i.e. X is either V or C. The rhyme consists of a nucleus and a coda.

The nucleus is an obligatory constituent in Dutch syllables. It contains at least one element, a vowel, /d **ɑ** k/ 'roof', and at most two elements of which the first is a vowel. The second element of the nucleus may consist of another vowel, as in a long vowel, C<u>VV</u> /z **e:**/ 'sea', or be part of a diphthong C<u>VV</u> /tr **œy**/ 'sweater', or it may be a sonorant, C<u>VC</u>$_{son}$ /b **ɑl** kOn/ 'balcony'.

The syllable coda is again optional in Dutch, CV /la:/ 'drawer'; if it is there, it may consist of any single obstruent: CV<u>C</u> /lɛ **f**/ 'nerve', /0 **k** səl/ 'armpit'. At word end a bi-positional rhyme may be followed by any single consonant (except /h/): /kɑnto: **r**/ 'office', /dro: **m**/ 'dream', /bɑl **k**/ 'beam'. In addition to this a coronal appendix is permitted at word end. The appendix will be dealt with in § 3.2.2.4 on extra-syllabicity.

3.2.2 Properties of Dutch syllables and words

3.2.2.1 The Sonority Sequencing Hierarchy

Important in clusters is the order in which consonants occur. This is connected with the Sonority Sequencing Hierarchy, a universal phenomenon (Clements 1990). Certain sounds are more sonorous than other; vowels have the highest sonority level, obstruents the lowest. The Sonority Index is in (32):

(32) Sonority Index

class	scale
stops	1
fricatives	2
nasals	3
liquids	4
glides	5

This hierarchy determines which sound sequences are possible and which are ruled out. Regular Dutch onset clusters consist of two consonants. Onset clusters must first of all conform to the Sonority Hierarchy, i.e. the second consonant must be more sonorous than the first, but they must also skip at least one category downward in this hierarchy (Trommelen 1988), i.e. the Minimal Sonority Distance (Selkirk 1984) is 2, MSD>2. With the sonority classes and their ranking of (32) as basis this allows the following onset clusters in Dutch: 1–3, plosive – nasal /**kn** ɑp/ 'handsome', 1–4, plosive – liquid /**pr** ɛt/ 'fun', 1–5, plosive – glide /**kw** ɑst/ 'brush', 2–4, fricative – liquid /**fl** ɛs/ 'bottle', 2–5, fricative – glide /**zw** ɛmən/ 'swim' and 3–5, nasal – glide not attested[12].

There are many apparent counterexamples in the language. These can be divided into two types: type 1 contains idiosyncratic irregularities, consisting of onset clusters that do conform to the sonority hierarchy, but do not skip a category: 1–2, plosive – fricative /ps, tʃ, ts, ks/ and 2–3, fricative – nasal /fn, xn, ʃm, ʃn/, and one cluster which makes

[12] Except in some French loanwords.

use of a reverse order for sonority – /wr/ (generally pronounced /vr/ in Dutch). Their total occurrence amounts to only 2.6% of the total number of clusters, 2,531, in the corpus[13]. For the present paper these exceptions are irrelevant and will be ignored. The second type of counterexamples to the rules for Dutch onsets are clusters consisting of /s/ + consonant. This type is connected with extra-syllabicity and will be dealt with below.

The Sonority Sequencing Hierarchy must also be applied in word coda, although the limitations are less strict than those for onset. Consonant coda clusters must diminish in sonority, and must be at least in adjacent categories, MSD>1 (Trommelen 1988 p 378). Possible Dutch codas are therefore: 5-4, glide-liquid, not attested[14], 5-3, glide-nasal, not attested, 5-2, glide-fricative /ni: **ws**/ 'news', 5-1, glide-plosive /no: **jt**/ 'never', 4-3, liquid-nasal /wɑ **rm**/ 'warm', 4-2, liquid-fricative /hɑ **lf**/ 'half', 4-1, liquid-plosive /hɛ **lp**/ 'help', 3-2, nasal-fricative /kɑ **ns**/ 'chance', 3-1, nasal-plosive /pɔ **mp**/ 'pump', 2-1, fricative-plosive /rɑ **sp**/ 'grater'. There are also counterexamples to this constraint; these are related to extra-syllabicity and will be dealt with below.

3.2.2.2 OCP-[place]

A second restriction on clusters is, that they do not share place of articulation. This is the topic of the present thesis and will be treated extensively in Chapter 4.

3.2.2.3 Assimilation

A third property of Dutch related to clusters is assimilation. We distinguish two types:

1. Voice Assimilation: VOI_{ASS} – all obstruent clusters agree in the specification for [voice]. This can be illustrated with examples from polymorphemic words: /zɑ**k**/ + /duk/ → /zɑ**g**duk/ 'handkerchief', /bɔs/ + /**v**ijo:l/ → /bɔ**sf**ijo:l/ 'wood-violet'; VOI_{ASS} is also applied within monomorphemic words: /ɑ**p**t/ 'abbot', /ɑ**b**dɛi/ 'abbey'.

[13] The corpus will be dealt with in 4.2 below.

[14] Except in a single English loanword: /fajl/ 'file'.

2 Nasal Place Assimilation: NAS_{Ass} – any nasal followed by an obstruent shares place of assimilation with that obstruent. This can be illustrated with the preposition 'in' followed by a noun. The /n/ is pronounced differently in accordance with place of articulation of a following consonant. The examples are from Trommelen & Zonneveld 1989: in petto → /I**mp**ɛto/ 'in reserve', in juni → /Iɲ**j**yni/ 'in June', in tranen → /I**nt**ra:nə/ 'in tears', in kas → /Iŋkɑs/ 'cash'. The phenomenon is also found within words: /lɑ**mp**/ 'lamp', /bɑŋ**k**Et/ 'banquet', /hɔ**nt**/ 'dog'.

3.2.2.4 Extra-syllabicity

As indicated above, there are two types of counterexamples to onset cluster restrictions. The second type consists of a large number of word-initial /s/-clusters. Dutch has the possibility to add an /s/-sound at the beginning of a single word-initial consonant or a word-initial consonant cluster. The initial /s/ is known as extra-syllabic /s/. These /s/-clusters do not conform to the restrictions imposed by the SSH, but are still well-formed. We distinguish three types:

1 /s/-clusters which conform to the Sonority Sequencing Hierarchy, but do not skip a category, /**sm** ɑl/ 'narrow', /**sn** Ik/ 'sob';

2 /s/-clusters with reversed sonority order, /**sp** ɛk/ 'bacon', /**sk** əlɛt/ 'skeleton', /**st** a:pəl/ 'pile';

3 a special case with fricative /s/ followed by another fricative, /x/, giving the cluster /sx/[15], /**sx** ɪp/ 'ship'.

The total percentage of /s/-clusters in all word-initial clusters of the Dutch language is 39.2%. With all these initial /s/-clusters in Dutch violating the Dutch constraints for onset clusters as described above, this illustrates the special nature of initial /s/. The only exceptions for /s/ plus consonant are /sr/ (not attested in the language), /sw/

[15] There are also two cases of /sf/ in the language, /sfer/ 'sphere' and /sfɪnks/ 'sphinx', but these are exceptions and will be ignored here.

(developed into /zw/) and the same-place-ànd-manner combinations */ss/, */sz/, */sʃ/, */sʒ/, which will be dealt with in § 4.4.1.3 below on OCP and similarity.

Also in coda clusters we encounter exceptions to the SSH. Dutch has the possibility to add a coronal appendix, consisting of one or two coronal consonants, at word end, /ho:f**t**/ 'head', /do:r**n**/ 'thorn', /pa:r**s**/ 'purple'; resulting clusters do not always conform to the restrictions imposed by the SSH, but are still well-formed. We distinguish two types:

1 clusters with consonants in the same sonority class, /mɑr **kt**/ 'market', /nɪ **pt**/ 'only just';

2 clusters with reversed sonority, /hɛ **ks**/ 'witch', /xɪ **ps**/ 'plaster, /rɛx **ts**/ 'rightward'.

Just as extra-syllabic /s/ in onset, coronal appendix /s/ and /t/ in coda need not adhere to the Sonority Hierarchy.

This sums up the structure of Dutch syllables. We have seen that Dutch words may begin with zero, one, or two consonants, or – with extra-syllabic /s/ – even three, that any syllable has a bi-positional rhyme, that at word end we find zero, one or two consonants, or – with final coronal obstruents – even three or four. Onset clusters avoid same place of articulation (with the exception of clusters with extra-syllabic/s/: /sl, sn, st/). At word begin as well as at word end the Sonority Sequencing Hierarchy must be adhered to. Onsets must skip at least one category downward in the hierarchy (Minimal Sonority Distance, MSD>2), in codas clusters consonants must be at least in adjacent upward sonority categories (MSD>1). Extra-syllabic /s/ at word begin and appendix /s/ and /t/ at word end need not adhere to the SSH. Clusters consisting of nasal + obstruent agree in place of articulation. All obstruent clusters agree in the specification for the feature [voice].

With this picture of the inventory of Dutch phonemes and of Dutch syllable structure, the main ingredients of the language that should be taken into account when discussing OCP, have been dealt with. I will now take up the investigation of OCP in Dutch.

4 OCP in Dutch

After the introduction of OCP and a survey of treatments of the phenomenon in the literature, as well as an introduction to Dutch phonotactics, I will now proceed with the investigation of OCP in Dutch. The four hypotheses summed up in Chapter 2 will be tested. I will first examine whether there are OCP restrictions in Dutch in general, in surface adjacent consonants and/or in consonants across segments – Hypothesis 1: OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments. Secondly, I will examine whether distance plays a role in Dutch – Hypothesis 2: The strictness of OCP is related to distance between two consonants, measured in segments. Subsequently, I will look into the matter of similarity and investigate whether multiple OCP violations increase the strictness of OCP in Dutch – Hypothesis 3: Degree of similarity influences the strictness of OCP. I will also explore the aspect of salience as a defining factor in OCP in Dutch – Hypothesis 4: The strictness of OCP is connected with prosodic salience. The present chapter is divided into 6 parts. Part 4.1 will introduce the database and working method. In section 4.2 I will investigate whether OCP-[place], [manner] and/or [voice] exist in Dutch. Part 4.3 will connect the findings on OCP of section 4.2 with the aspect of distance as introduced in Chapter 2. This will be followed by an investigation into OCP and similarity in part 4.4 and OCP and salience in part in 4.5. In section 4.6 I will discuss two more aspects that might influence OCP, viz. quantity and quality of an intervening vowel, and cluster membership of one or two of the harmonic consonants.

4.1 Database and working method

The Dutch National Expertise Centre CELEX (Centre for Lexical Information) has developed an electronic database with lexical information on contemporary Dutch, English and German (Baayen, Piepenbrock & Gulikers, 1995). The data for the present thesis have been taken from the Celex database on Dutch Phonological Lemmas (DPL). With the help of the Morphological Database all monomorphemic words in the database were extracted; doublets were filtered out. This yielded a type-based corpus of 8,305 monomorphemic Dutch words, with a total number of 41,328 segments. For computations on the corpus I made use of PhonotacTools, a computer programme for research on phonotactics. The programme can extract data, such as occurrence frequencies, transitional probabilities and observed/expected ratios, from phonemically transcribed corpora (Adriaans, 2006). I extracted different segment combinations from the corpus of 8,305 words, such as CC combinations, CVC and CXXC, where C represents any consonant, V represents any single vowel and X represents a vowel or consonant. An example of how e.g. CVC-combinations were extracted is in table (31).

The corpus contains 10,378 CVC combinations; these were extracted from different word types and different positions in the word. The numbers of segments in the words vary from 3 to 11. In the examples in table (31), the extracted combinations are underlined; the respective consonants in the words in phonetics are in bold.

(31) CVC sources in monomorphemic words in Dutch

Monosyllabic words		
CVC	/**k**ɑ**t**/	*cat*
CCVC	/s**t**a**t**/	*state*
CVCC	/**k**ɛ**r**k/	*church*
CCCVC	/st**r**ɑ**k**/	*tight*
CVCCC	/**b**ɑ**r**st/	*crack*
CCVCC	/k**r**ɑ**x**t/	*power*
CCCVCC	/sx**r**ɪ**f**t/	*exercise-book*
Polysyllabic words		
CVCV	/**b**e**b**i/	*baby*
VCVC	/a**d**ə**l**/	*nobility*
CVCVC	/**z**o**m**ər/	*summer*
CVCVC	/zo**m**ə**r**/	*summer*
VCVCC	/a**l**ɑ**r**m/	*alarm*
CCVCCV	/b**l**ɪ**k**jə/	*tin*
VCVCVC	/o**j**ə**v**ar/	*stork*

VCV<u>C</u>V<u>C</u>	/ojə**var**/	*stork*
<u>C</u>V<u>C</u>VCVCCCVC	/**rod**odɛndrɔn/	*rhododendron*
CV<u>C</u>V<u>C</u>VCCCVC	/ro**dod**ɛndrɔn/	*rhododendron*
CVCV<u>C</u>V<u>C</u>CCVC	/rodo**dɛn**drɔn/	*rhododendron*
CVCVCVCC<u>C</u>V<u>C</u>	/rododɛnd**rɔn**/	*rhododendron*

In order to establish whether there is OCP in Dutch, I extracted different strings of segments and analysed these for O/E values. The O/E measure used by Pierrehumbert (1993), is "the ratio of the number of pairs observed in the dictionary (O) to the number of occurrences which would be expected in the absence of any OCP effect (E). O/E ranges from 0 for an absolute effect (an effect so strong that none of the expected occurrences are found) up to 1.0 or more when there is no OCP effect (the combination occurs as often [as, ed], or more, than expected)" (Pierrehumbert 1993). This O/E measure uses the ratio of the observed number of consonants pairs to the number that would be expected if consonants combined at random.
In the present thesis I make use of *contextua*l O/E. The O/E value – the Observed/Expected value – divides the observed value, O – the actual number of occurrences of the segment combination in a particular database – by the expected value, E – the number one expects given that the corpus under scrutiny distributes consonants evenly in accordance with occurrence rates in the context under investigation (Kager & Shatzman, 2007).

The formula applied to establish E is

(32) $E = p(X) * p(Y) * N_{AB}$

The expected value is computed as the probability that consonant X occurs in the initial position of a consonant combination (for example CVC strings), p(X), multiplied by the probability that consonant Y occurs in the final position of a consonant combination (CVC strings), p(Y), multiplied by the total number of consonant combinations in the database,

N_{AB}. Throughout this thesis N_{AB} will vary, since different segments combinations, such as CC, CVC, CXXC, and segment combinations in different positions in the word, such as CC in word onset, word-medial CC and CC in word coda, will be considered. Thus expectancy scores are not absolute; they are dependent on the contexts in which they occur. An example will clarify matters. Calculations for O/E scores on CVC strings, irrespective of position in the word, are based on 10,378 CVC strings. The expected value for e.g. /bVp/ is based on the actual number of initial /b/s in the database of CVC strings, 662, the actual number of /p/s in final position in the string, 442, and the total number of CVC strings, 10,378. Thus the expectancy score for /bVp/ strings is:

(33) E = 0.0638 (662/10,378) x 0.0426 (442/10,378) x 10,378 = 28.2053

The observed number of /bVp/ strings in the corpus is 7, hence the O/E value for /bVp/ is:

(34) /bVp/ O/E: 7 / 28.2053 = 0.2481

The O/E value for /bVp/ in the database of CVC strings irrespective of position in the word is 0.2481. In a specific context, such as for example word-medial CVC, the numbers will differ in accordance with the actual number of occurrences of /b/ in initial position in a word-medial CVC string, and /p/ in final position in a word-medial CVC string, as well as with the total number of word-medial CVC strings, 1,905. The introductions to the analyses below will provide information on the exact nature of the consonant combinations dealt with, as well as on the numbers of segment combinations on which O/E scores are based.

In order to determine O/E scores based on natural classes, it is important to define which consonants belong to a particular class and to reduce the consonants to the feature(s) under investigation. When, e.g. comparing O/E scores for different place categories, it is important to calculate O/E scores based on the place categories only. A means to do this is by coding consonants. When, e.g. all labial consonants in the database, /b, p, v, f, m, w/ are coded /P/ the consonants are reduced to their place feature, irrespective of other features such as sonorancy, continuancy, voice. When totalling the number of consonants sharing a particular (place) feature and calculating O/E rates based on the totalled scores, O/E scores reflect results for [place] only. The O/E score for surface adjacent labial consonant pairs thus calculated is:

(35) PP O/E: 183 / 164.0364 = 1.1156

Calculations of O/E scores in the present thesis are based on this approach.

As indicated above, the perfect O/E rate is 1.0, where observed value and expected value coincide. “A smaller ratio indicates a stronger OCP effect, since it indicates that the observed value is small relative to the expected value” (Berkley, 1994). In the literature on OCP, values below 1.0 have been taken as evidence of restrictions on co-occurrence. Given that total equality of observed and expected gives the perfect score of 1.0 and allowing room for noise and accidental gaps, this thesis takes an O/E rate below 0.75 as an indication of underrepresentation. Consonant combinations with scores above 1.33 (1/0.75), are considered to be evidence of overrepresentation. Anywhere in the present thesis where there is a reference to “normal” O/E scores, it will concern O/E scores between 0.75 and 1.33.
Since it was established in Chapter 2 that OCP is a gradient phenomenon in languages of the world, O/E scores within categories are expected to differ in accordance with aspects influencing gradience in Dutch. In order to ascertain the validity of perhaps minor differences within categories, a measure for evidence in favour of gradience would be desirable. However, whereas a measure for evidence in favour of OCP, viz. O/E < 0.75, was determined, no fixed measure for gradience is included in the present thesis. Differences in scores will be discussed as they arise. For statistical significance, the test outcomes will be computed with chi-square analyses, wherever possible; the probability

rates will be indicated below the tables. An exact indication of how scores have been calculated per feature will be presented in the relevant sections below. It is important that O/E scores are calculated for combinations with a sufficient number of items to be able to draw conclusions. In the present thesis combinations occurring in a set of 1,000 items or more – $N_{AB} > 1{,}000$ – are considered to render a reliable score. Wherever numbers of items below 1,000 have been taken as basis for the calculations this will be indicated. The numbers of items per analyzed category are in the tables.

4.2 Is there OCP in Dutch?

As seen above in Chapter 1, OCP is a phenomenon active in many languages of the world. It does not only concern surface adjacent consonants, but also consonants separated by one or more segments. Three types of OCP relevant to Dutch came up in the literature in Chapter 1, section 1.1, viz. OCP-[place], OCP-[manner] and OCP-[voice]. The situation for Dutch with regard to co-occurrence restrictions on [place], [manner] and [voice] will be scrutinized here. In the present section I will give an analysis of Dutch with regard to surface adjacent harmonic consonants as well as harmonic consonants separated by one segment and harmonic consonants separated by two segments. I will use the terms 'harmony' and 'harmonic' for segments which share a feature under discussion; /p/ and /m/, for instance, share the feature [labial], they are Place-harmonic consonants.

This section will be concerned with Hypothesis 1 – OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments.

The hypothesis will be divided into two separate hypotheses:

Hypothesis 1 (1) – OCP is a constraint active in the Dutch language in surface adjacent consonants

Hypothesis 1 (2) – OCP is a constraint active in the Dutch language in consonants across segments

Hypotheses will be corroborated or rejected based on the O/E scores. Normal O/E scores, 0.75 – 1.33, will be considered as not indicating evidence in favour of OCP; scores below 0.75, reflecting underrepresentation, will be considered evidence in favour of OCP. Scores above 1.33, indicating overrepresentation, will of course also be considered as not indicating evidence in favour of OCP. I will investigate all Place harmonies at different distances in the word in Dutch, followed by Manner harmonies and Voice harmonies at different distances in the word. Hypothesis 1 will therefore be divided into a Place, b Manner and c Voice, as well as into OCP in surface adjacent consonants and OCP in consonants across segments. Hence the six parts of Hypothesis 1 will be:

Hypothesis 1 (a1): OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [place].

Hypothesis 1 (a2): OCP is a constraint active in the Dutch language in consonants across segments for the feature [place].

Hypothesis 1 (b1): OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [manner].

Hypothesis 1 (b2): OCP is a constraint active in the Dutch language in consonants across segments for the feature [manner].

Hypothesis 1 (c1): OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [voice].

Hypothesis 1 (c2): OCP is a constraint active in the Dutch language in consonants across segments for the feature [voice].

Wherever required, relevant phonotactic constraints, as introduced in Chapter 3.2, will be considered.

4.2a OCP-[place]

4.2a.1 OCP-[place] in surface adjacent consonants

In this section I explore whether there is OCP-[place] in Dutch. Restrictions in the language which could be of influence on OCP will be dealt with. Hypothesis 1 (a1) – OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [place] – will be examined here. I extracted CC combinations from the corpus of 8,305 monomorphemic Dutch words, irrespective of position in the word. Thus consonant pairs include CC combinations in word onset, such as /**kl**ɑnt/ 'customer', /**st**ra:t/ and /s**tr**a:t/ 'street', word-medial consonant combinations, /wɛ**rk**ə/ 'work', /vɛ**ns**tər/ and /vɛn**st**ər/ 'window', and CC combinations in word coda, /bɑ**lk**/ 'beam', /dɔ**rs**t/ and /dɔr**st**/ 'thirst'. This yielded a database of 5,628 consonant pairs. This is the total number of pairs on which O/E scores in the table below are based, N_{AB} in the O/E formula in (32). In order to investigate whether there is OCP in Dutch in surface adjacent consonants, I coded the database; all labial consonants, /p, b, f, v, m, w/, were coded P, all dorsals, /k, g, x, ɣ, ŋ/, were coded K, and all coronals, /t, d, s, ʃ, z, ʒ, n, l/, were coded T. Table (35) gives O/E scores for surface adjacent consonant pairs in Dutch.

(35) O/E scores for Place Harmony in Dutch monomorphemic words in surface adjacent consonants

CC Place Harmony in Dutch monomorphemic words	
items	5,628
PP	1.1156
KK	1.3198
TT	1.0666

The table shows that all three categories have normal O/E scores in surface adjacent consonants. This would suggest that strictly surface adjacent consonants are not affected by OCP in Dutch in any of the three place categories. However, as indicated in Chapter 3.2, there are phonotactic constraints on consonant clusters with regard to the feature [place], which might obscure the effects of OCP-[place] by causing strong over- or underrepresentation of certain consonant clusters.

Dutch has onset clusters up to a maximum of three consonants, coda clusters up to a maximum of four consonants (very rare) and word-medial consonant clusters up to a maximum of four consonants (very rare). In order to be able to judge the effect of the above mentioned constraints on consonant clusters I split up the 5,628 consonant clusters into three categories by their position in the word, CC in word onset, /**kl**ɑnt/ ‘customer’, /**st**ra:t/ and /s**tr**a:t/ ‘street’, word-medial CC, /wɛ**rk**ə/ ‘work’, /vɛ**ns**tər/ and /vɛn**st**ər/ ‘window’ , and CC in word coda, /bɑ**lk**/ ‘beam’, /dɔ**rs**t/ and dɔr**st**/ ‘thirst’, and calculated O/E rates per category. The results are in table (36). Chi-square tests were executed on the table in general – based on the actual number of occurrences of the combinations of all three place categories in all three positions – indicating significance between the three place categories; chi-square tests were also executed on the individual place categories, computing significance between the three positions in the word per place category. Chi-square tests per place category for, for instance, the category of labials are based on the actual number of labial pairs, 5 in word onset, 140 word-medially, and 38 in word coda, versus the actual number of other pairs per position in the word, 2700 in word onset, 1716 word-medially and 1029 in word coda. In subsequent tables, similar chi-square results per category will be indicated below the tables, wherever relevant.

(36) O/E scores for Place Harmony in Dutch monomorphemic words in surface adjacent consonants by position in the word

CC Place harmony by position in the word			
position in the word	**word onset**	**word-medial**	**word coda**
items	2,705	1,856	1,067
PP	0.0482	3.0401	2.8598
KK	0	3.4223	2.5748
TT	0.8601	1.2326	1.1226

X^2 166.18 df 4 $p < .0001$

Labial : X^2 189.85 df 2 $p < .0001$
Dorsal : X^2 171.96 df 2 $p < .0001$
Coronal: X^2 286.44 df 2 $p < .0001$

The table shows that labial and dorsal harmony are hardly (labial), or not at all (dorsal) present in onset clusters, whereas O/E scores for coronals are within normal expectation rates for surface adjacent consonants in any position in the word. The O/E scores for word-medial CC and CC in word coda for labial and dorsal surface adjacent clusters are substantially above expectation rates. The differences in O/E rates between coronal word onset clusters on the one hand and coronal clusters word-medially and in word coda position on the other hand indicate a significant (gradient) difference within the category of coronals as well.

Further investigation shows that the O/E score for labial harmony in word onset is based on 5 French loanwords only; there are no native Dutch words with labial harmony in onset clusters. The normal O/E score for coronals is almost entirely due to the phenomenon of extra-syllabicity, as introduced in § 3.2.2.4. When excluding clusters with extra-syllabic /s/ from the analysis, the total number of CC pairs in onset is 1,958, and the O/E score for coronals drops to 0.0912. OCP-effects for coronal consonant pairs in word onset are obscured by extra-syllabicity.

Analysis also shows that all word-medial labial and dorsal harmonies, as well as all labial and dorsal harmonies in word coda, consist of a nasal followed by an obstruent. This points at a phenomenon of Dutch we came across in § 3.2.2.3, viz. nasal place assimilation; NAS_{Ass} – all nasals followed by an obstruent agree in place of articulation with the obstruent. Overrepresentation of labial and dorsal consonant clusters word-medially and in word coda, as well as the significant difference in O/E scores between coronal harmony in word onset and coronal harmony word-medially and in word coda, are thus accounted for. Nasal Place assimilation outranks OCP-[place] in surface adjacent consonants, and obscures any OCP effects in these positions.

Thus, what we have found with regard to OCP-[place] in surface adjacent consonants is very strict OCP for dorsal, labial and coronal surface adjacent consonants other than extra-syllabic /s/ + coronal in onset, and assimilated nasal + obstruent word-medially and in word coda. The faithful realization of extra-syllabic /s/ + consonant outranks OCP constraints and affects coronal harmony in onset, whereas the constraint regulating nasal place assimilation outranks OCP in all three place categories, which affects all word medial consonant clusters and clusters in word coda. The data from onset pairs indicate strong evidence in favour of Hypothesis 1 (a1) – OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [place].

4.2a.2 OCP-[place] in consonants across segments

This paragraph will deal with Hypothesis 1 (a2) – OCP is a constraint active in the Dutch language in consonants across segments for the feature [place]. In order to find out whether the Dutch language is also affected by OCP-[place] restrictions in consonants across segments, table (37) below will give the picture for Place harmony in Dutch across segments in the three Place categories, [LAB], [DOR] and [COR]. The analysis will include consonants separated by a single vowel[16], CVC in the table, and consonants separated by CV, VC or VV (diphthong), CXXC in the table. CVC and CXXC combinations were extracted irrespective of position of the combination in the word, as indicated in the example on CVC combinations in table (32) above. All labial consonants, /p, b, f, v, m, w/, were coded P, all dorsals, /k, g, x, ɣ, ŋ/, were coded K, and all coronals, /t, d, s, ʃ, z, ʒ, n, l/, were coded T. Intermediate segments were filtered out. The numbers of pairs on which O/E scores in the table are based, N_{AB} in the O/E formula, are 10,378 coded CC pairs, based on CVC strings, and 6,302 coded CC pairs, based on CXXC strings.

[16] CCC harmony has been left out on account of the fact that restrictions on these combinations are connected with restrictions on surface adjacent consonants, as discussed in Chapter 3, rather than with possible OCP effects across a segment.

(37) O/E scores for Place harmony in Dutch monomorphemic words across segments

Place harmony in Dutch monomorphemic words across segments		
distance	**CVC**	**CXXC**
items	10,378	6,302
PP	0.4630	0.7174
KK	0.5078	0.8643
TT	0.8709	0.8725

X^2 24.58 df 2 $p < .0001$

Labial : X^2 30.42 df 2 $p < .0001$
Dorsal : X^2 43.43 df 2 $p < .0001$
Coronal : X^2 43.19 df 2 $p < .0001$

It appears from table (37) that labial and dorsal harmonies in CVC-combinations are underrepresented and that the O/E rates for labial CXXC-harmony are below normal as well. It can also be observed that coronals are not affected in either distance category in Dutch. There is thus evidence in favour of Hypothesis 1 (a2) – OCP is a constraint active in the Dutch language in consonants across segments for the feature [place], but only for labials and dorsals.

From the investigation into OCP-[place] above, we may conclude that Hypothesis 1 (1) – OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments for the feature [place] – has been corroborated.

4.2b OCP-[manner]

This section will be concerned with Hypothesis 1 (b) – OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments for the feature [manner]. The Manner features include [consonantal], [sonorant], [continuant], [nasal], [approximant], [lateral]. As indicated above in Chapter 3.1, the category of glides, consisting of /w/ and /j/, is left out of the analysis for [manner] on account of the vowel-like quality of its members, evident from the feature [-CONS]. This also implies that the category for Dutch approximants investigated in this thesis consists of /l/ and /r/ only. These are distinguished from the other sonorants by the feature

[-Nas], hence approximants will be left out of the analysis as well. The feature [consonantal] is irrelevant for the present study of “consonant” co-occurrence of [manner] and will therefore also be left out. The feature [nasal] is only contrastive within the category of [sonorant], and the feature [lateral] is only contrastive within the category of liquids; these two categories will be dealt with in 4.4 on similarity below. This leaves the Manner features [sonorant] and [continuant].

The feature [sonorant] divides all consonants into two categories:

[-SON] includes /p, t, k, b, d, g, f, s, ʃ, x, v, z, ʒ, ɣ/

[+SON] includes /m, n, ŋ, l, r/.

The feature [continuant] also divides all consonants into two categories:

[-CONT] represents /p, t, k, b, d, g, m, n, ŋ/

[+CONT] represents /f, s, ʃ, x, v, z, ʒ, ɣ, l, r/.

4.2b.1 OCP-[manner] in surface adjacent consonants

In order to explore whether there is OCP-[manner] in surface adjacent consonants in Dutch, Hypothesis 1 (b1) – OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [manner] – will be examined here. I calculated O/E scores for the separate manner categories indicated above. O/E rates were calculated over the number of surface adjacent consonant pairs in the corpus, irrespective of position in the word, i.e. 5,628 CC pairs. Table (38) gives O/E rates for surface adjacent consonants in the Manner categories [sonorant] and [continuant].

(38) O/E scores for Manner harmony in Dutch

Manner harmony in Dutch	
	CC
items	5,628
[-SON]	0.7284
[+SON]	0.3096
[-CONT]	0.7001
[+CONT]	0.7795

[sonorant] X^2 1284.73 df 1 <.0001
[continuant] X^2 38.03 df 1 <.0001

It is shown in the table that in the category of sonorancy both categories are underrepresented in the language. In surface adjacent consonants there is a clear co-occurrence restriction for [+SON] and a less strict limitation for [-SON].
With regard to continuants we find O/E rates which are only just within limits for the feature [+CONT], and below limits for the feature [-CONT].

As with restrictions on the feature [place] in surface adjacent consonants, the Manner features could be influenced by other than OCP restrictions in surface adjacent consonants. I will therefore also investigate Manner restrictions on consonant clusters, as introduced in Chapter 3.2.

In order to check the behaviour of OCP [manner] in consonant clusters in connection with position in the word, I divided the category into three separate categories, CC in word onset, /**kl**ɑnt/ 'customer', /**st**ra:t/ and /s**tr**a:t/ 'street', word-medial CC, /wɛ**rk**ə/ 'work', /vɛ**ns**tər/ and /vɛn**st**ər/ 'window', and CC in word coda, /bɑ**lk**/ 'beam', /dɔ**rs**t/ and /dɔr**st**/ 'thirst', and calculated O/E rates per category. The results are in table (39).

(39) O/E scores for CC [manner] by position in the word

CC [manner] by position in the word			
position in the word	word onset	word medial	word coda
items	2,705	1,856	1,067
[-SON]	1.0093	0.8321	1.0478
[+SON]	0	0.7280	1.5686
[-CONT]	0.2890	0.8617	0.7474
[+CONT]	0.7992	0.7643	0.6043

[sonorant] : X^2 202.6 df 2 $p < .0001$
[continuant] : X^2 180.86 df 2 $p < .0001$

[-Son] : X^2 80.11 df 2 $p < .0001$
[+Son] : X^2 176.47 df 2 $p < .0001$
[-Cont] : X^2 96.28 df 2 $p < .0001$
[+Cont] : X^2 183.29 df 2 $p < .0001$

As indicated by the numbers in the table, there are no [+SON] pairs in word onset; [+SON] pairs in word-medial position are restricted; word coda [+Son] pairs have scores above normal. The feature [-SON] has normal scores for consonant pairs in all positions in the word (although word-medial [-SON] pairs score significantly lower than pairs at word begin and word end).

Surface adjacent consonant pairs in the category of [continuant] show restrictions in [-CONT] pairs in word onset and word coda, and in [+CONT] pair in word coda.

As indicated above, results could be influenced by other than OCP constraints. Two phonotactic properties of clusters introduced in Chapter 3.2, i.e. the Sonority Sequencing Hierarchy and extra-syllabicity, will be discussed here.

The Sonority Sequencing Hierarchy

As we saw in Chapter 3, onset clusters as well as coda clusters are restricted by the Sonority Sequencing Hierarchy; initial consonant clusters need to skip at least one category downwards in the sonority hierarchy, MSD>2, coda clusters are minimally one category upwards apart, MSD>1. The Sonority Sequencing Hierarchy, SSH, may be considered a restriction related to OCP. Both regulate an obligatory contour. The SSH is the more restricted variety; not only is it a restriction on the co-occurrence of consonants in the same (or, in onsets, even an adjacent) sonority category, it is also a restriction on the direction applied. Legal Dutch onsets have rising sonority, codas have falling sonority. There are no restrictions on the direction of the sonority category in word-medial CC clusters. The SSH explains the non-occurrence of [+SON] pairs word-initially – with two sonority levels in the category of sonorants it is impossible to skip a category. Since there are only two sonority levels in the category of [-SON] as well, we must infer that the SSH will also outrank OCP [manner] constraints in all [-SON] clusters, so that in this manner category no OCP effects could be visible in the language in onset.

Extra-syllabicity

With the SSH outranking OCP we would expect, as indicated above, that there are limitations on all [αSon] pairs in word onset. However, as seen in table (31) above, the O/E scores for word-initial [-Son] pairs are normal. This is connected with extra-syllabicity.

As indicated in Chapter 3.2, Dutch has the possibility to add an /s/-sound at the beginning of a single word-initial consonant or a word-initial consonant cluster. At word end there can be a coronal appendix, which need not adhere to the sonority hierarchy. The normal O/E scores for [-SON] pairs in word onset and word coda are entirely due to these extra-syllabic consonants outranking SSH constraints. When we exclude extrasyllabic /s/ from the onset CC database, no fricative-plosive combinations are found, whereas plosive-fricative clusters are restricted to borrowings from other languages; in word coda there are no [-SON] pairs of which not at least one segment is a coronal appendix. Thus, we may conclude that restrictions in the Manner feature [sonorant] in literally surface adjacent consonants in onset are very strict.

With the SSH outranking OCP [sonorant] in surface adjacent consonants in onset and extra-syllabicity outranking the SSH in onset and coda, we may infer that syllable structure constraints obscure OCP [manner] effects in surface adjacent consonants in onset and coda for the feature [sonorant].

This leaves the category of [continuant]. In the category of [continuant], the normal score for word-initial [+CONT] pairs can be explained by an idiosyncratic /s/-cluster, viz. /sx/[17]. We find 247 initial /sx/ pairs out of a total number of 723 CC [+Cont] pairs. When leaving out the 247 /sx/ clusters, O/E scores drop to 0.6385, which indicates underrepresentation, just as onset [-Cont] pairs are underrepresented.

With restrictions on onset and coda surface adjacent consonant pairs in the feature [continuant], and restrictions on [+SON] pairs word-medially, we find evidence in favour of Hypothesis 1 (b1) – OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [manner].

4.2b.2 OCP-[manner] in consonants across segments

The second part of Hypothesis 1 (b), Hypothesis 1 (b2) – OCP is a constraint active in the Dutch language in consonants across segments for the feature [manner] – will be addressed here. Table (40) below will give the picture for Manner harmony in Dutch across segments. The analysis will cover consonants separated by a single vowel, CVC in the table, and consonants separated by CV, VC or VV (diphthong), CXXC in the table. The table includes the features [sonorant] and [continuant].

[17] /sx/ is a remnant of Middle Dutch (c.1100 – c.1500). It developed into /s/ in Late Middle Dutch in word-medial and word-final position. Initial /sx/ was retained (Van Loon, 1986).

(40) O/E scores for [sonorant] and [continuant] across segments by position in the word

[sonorant] and [continuant] across segments by position in the word		
	CVC	CXXC
items	10,378	6,302
[-SON]	0.7975	1.0966
[+SON]	0.8555	1.0048
[-CONT]	0.8571	1.0022
[+CONT]	0.8587	1.0450

[sonorant] : X^2 985.46 df 1 p < .0001
[continuant] : X^2 27.03 df 1 p < .0001

[-Son] : X^2 1297.80 df 1 p < .0001
[+Son] : X^2 301.87 df 1 p < .0001
[-Cont] : X^2 116.57 df 1 p < .0001
[+Cont] : X^2 3.35 df 1 p 0.0672

Table (40) shows normal scores for the categories of [sonorant] and [continuant]. This indicates that there is no evidence in favour of Hypothesis 1 (b2) – OCP is a constraint active in the Dutch language in consonants across segments for the Manner features [sonorant] and [continuant]

With OCP-[αCONT] in Dutch in surface adjacent consonants in word onset and word coda, and OCP-[+SON] word-medially, and with no OCP-[manner] across segments, Hypothesis 1b – OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments for the feature [manner] – is only partly corroborated.

4.2c OCP-[voice]

4.2c.1 OCP-[voice] in surface adjacent consonants

As with the features [place] and [manner], the investigation into OCP-[voice] will take two hypotheses into consideration. Hypothesis 1 (c1) – OCP is a constraint active in the Dutch language in surface adjacent consonants for the feature [voice]– will be considered first. Table (31) below will give the general picture for Voice harmony in surface adjacent consonant in Dutch. [voice] is a distinguishing feature within the category of obstruents; these can either be voiced or voiceless. There are two ways of approaching the feature [voice] for sonorants. Since sonorants are more sonorous than obstruents, they may be connected for the feature [voice] with voiced obstruents, and together form the category of voiced consonants, or, since there is no distinction within the category of sonorants for [voice], they can be said to be unspecified for [voice], and must therefore be left out of the analysis for voicing. The second approach will be considered in connection with similarity in section 4.4. below. Table (41) will give the general picture for the division of [voice] in Dutch including sonorants. I extracted CC combinations, irrespective of position in the word. Subsequently, consonants /p, k, t, f, x, s, ʃ/ were coded for [-VOI], consonants /b, g, d, v, ɣ, z, ʒ, m, ŋ, n, l, r, j, w/ were coded for [+VOI]. O/E scores were thus calculated on the basis of the feature [voice] only. O/E rates were calculated over the number of surface adjacent consonant pairs in the corpus, 5,628.

(41) O/E scores for [voice] harmony in Dutch monomorphemic words in surface adjacent consonants

[voice] harmony in Dutch monomorphemic words in surface adjacent consonants	
distance	CC
items	5,628
[-VOI]	0.9804
[+VOI]	0.9748

Both combinations have O/E rates as expected. It looks as if there is no OCP effect in a two way division of voicing among all consonants. As with [place] and [manner] I will investigate the phonotactic constraints of Chapter 3.2 for [voice] as well. Division into three categories in connection with position in the word leads to:

(42) O/E scores for Voice harmony in Dutch monomorphemic words in surface adjacent consonants by position in the word

Voice harmony in Dutch monomorphemic words in surface adjacent consonants by position in the word			
CC	initial	medial	final
items	2,705	1,856	1,067
[-VOI]	1.2125	1.3241	1.0505
[+VOI]	1.3801	1.1690	1.5199

X^2 347.31 df 2 p < .0001

The O/E scores for voiceless consonant clusters are normal or above normal in all positions in the word. It is evident from table (42) that in no position in the word is there underrepresentation of voice harmony in surface adjacent consonants in monomorphemic words. This is again in accordance with a phonotactic phenomenon discussed in Chapter 3.2, viz. Assimilation of Voice: VOI_{Ass} – all obstruent clusters agree in voicing. There is no OCP-[voice] in Dutch in surface adjacent consonants. Hypothesis 1 (c1) – OCP is a constraint active in Dutch in surface adjacent consonants for the feature [voice] – must be rejected.

4.2c.2 OCP-[voice] in consonants across segments

In order to find out whether the Dutch language is affected by OCP [voice] restrictions in consonants across segments, I analyzed [voice] harmony in Dutch across segments. The analysis will cover consonants separated by a single segment, CVC in the table, and consonants separated by CV, VC or VV (diphthong), CXXC in the table.

(43) O/E scores for Voice harmony in Dutch monomorphemic words across segments

Voice harmony in Dutch monomorphemic words across segments		
distance	CVC	CXXC
items	10,378	6,302
[-VOI]	0.8283	0.9638
[+VOI]	0.9489	1.0058

X2 720.9 df 1 P < .0001

All combinations have O/E rates as expected.

It is evident that there is no OCP effect in a two way division of voicing among all consonants.

This concludes the initial investigation on the existence of OCP in Dutch. The question "*is there OCP in Dutch?*" can be answered as follows. Firstly, for surface adjacent consonants, there is OCP-[place], viz. [labial] and [dorsal], there is OCP [manner], viz. [continuant] and [+Son], there is no OCP-[voice]. Secondly, across segments we find an OCP effect for [place], viz. [labial] and [dorsal], and not for [manner] or for [voice]. We can posit that Hypothesis 1 (1) – OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments – is corroborated for [place], partly corroborated for [manner] and rejected for [voice].

4.3 OCP and Distance

This section, as well as subsequent sections in the present chapter, will deal with influences on OCP and on the gradient nature of OCP, as observed in Chapter 2. Since OCP [voice] was shown not to be active in the Dutch language, this section will contain a two-way division, viz. an investigation into a Place and b Manner. I will here take on Hypothesis 2 – the strictness of OCP is related to distance between two consonants, measured in segments.

Hypothesis 2 is again divided into two hypotheses:

Hypothesis 2 (a): the strictness of OCP-[place] is related to distance between two consonants, measured in segments.

Hypothesis 2 (b): the strictness of OCP-[manner] is related to distance between two consonants, measured in segments.

4.3a OCP-[place] and distance

In order to investigate the influence of distance between two elements on OCP-[place] in Dutch, we take a closer look at tables (36) and (37), here combined into (44). Since it was established in section 4.2a above that word-medially and in word coda OCP is outranked by nasal place assimilation, and therefore not visible in the language, O/E scores for surface adjacent consonants regard onset pairs only.

(44) O/E scores for Place harmony in Dutch monomorphemic words by distance between the segments

Place harmony in Dutch monomorphemic words by distance between the segments			
distance	CC(onset)	CVC	CXXC
items	2,705	10,378	6,302
PP	0.0482	0.4630	0.7174
KK	0	0.5078	0.8643
TT	0.8601	0.8709	0.8725

		X^2 144.55 df 4 p < .0001
Labial	:	X^2 107.53 df 2 p < .0001
Dorsal	:	X^2 93.11 df 2 p < .0001
Coronal	:	X^2 45.73 df 2 p < .0001

As can be seen from the table, coronal pairs have normal O/E scores in all three categories. It is clear that dorsal and labial harmony in onset clusters are avoided in the language. In CVC combinations there is underrepresentation in these two place categories. There is a substantial difference in O/E scores in labials and dorsals between harmonic surface adjacent consonants and harmonic consonants across one segment, viz. 0.42 and 0.50 respectively, as well as in consonants across one segment and two segments, viz. 0.25 and 0.36. Dorsal consonant pairs reach normal scores across two segments.

The investigation on OCP-[place] in connection with the aspect of distance gives a clear picture. Firstly, it can be established that strictly surface adjacent labial and dorsal consonants in onset are strongly affected by OCP-[place]. Secondly, for dorsals and labials in CVC harmonies, there are fairly strict OCP limitations. Thirdly, these limitations become weaker when the segments are four places apart. Strictly surface adjacent onset consonants show the strongest limitation, the number of intervening segments proportionally lessens the effect in labials and dorsals. We may infer that distance is a defining factor in OCP-[place]. Hypothesis 2 (a) – The strictness of OCP-[place] is related to distance between two consonants, measured in segments – has been corroborated.

4.3b OCP-[manner] and distance

In order to investigate the influence of distance between two elements on OCP-[manner] we take a closer look at the findings in section 4.2b above. It was established that OCP in surface adjacent consonants was only visible in the language in the feature [continuant] on account of other constraints outranking OCP-[αSON] and therefore obscuring possible OCP effects. Table (45) below shows manner harmony in connection with distance between the segments.

(45) O/E scores for Manner harmony in Dutch

Manner harmony in Dutch			
	CC	CVC	CXXC
items	5,628	10,378	6,302
[-SON]	X	0.7975	1.0966
[+SON]	X	0.8555	1.0048
[-CONT]	0.7001	0.8571	1.0022
[+CONT]	0.6385*	0.8587	1.0450

[sonorant] : X^2 985.46 df 1 $p < .0001$
[continuant] : X^2 27.03 df 1 $p < .0001$

[-Son] : X^2 1297.80 df 1 $p < .0001$
[+Son] : X^2 301.87 df 1 $p < .0001$
[-Cont] : X^2 226.38 df 1 $p < .0001$
[+Cont] : X^2 5.88 df 1 p 0.0529

* this score is based on [+Cont] pairs excluding /sx/; (see footnote 17)

Although O/E scores across segments for the feature [sonorant] are normal. It must be noted that in both [sonorant] categories across segments O/E scores proportionally diminish with distance between the segments. Both continuancy categories show underrepresentation in surface adjacent consonant pairs. O/E scores across segments for the feature [continuant] are normal. In both [continuant] categories across segments O/E scores proportionally diminish with distance between the segments. For both manner categories there is increased leniency with increased distance. Hypothesis 2 (b) – The strictness of OCP [manner] is related to distance between two consonants, measured in segments – has been corroborated.

This concludes the investigation on the influence of distance on OCP-[place] and [manner]. The influence of distance on OCP effects is present in Dutch. With regard to [place], we find that in labial and dorsal harmony the influence of OCP proportionally diminishes with distance. With regard to [manner] we find an effect of distance on the features [sonorant] and [continuant]. Thus distance is a defining factor in the phenomenon of OCP-[place] and [manner] in Dutch and Hypothesis 2 – the strictness of OCP is related to distance between two consonants, measured in segments – has been corroborated.

4.4 OCP and Similarity

It appeared from Chapter 2 that the gradient nature of OCP was not only connected with distance between the segments, but also with similarity between two segments. The similarity matrix, developed by Pierrehumbert, Frisch and Broe (1997), as discussed in Chapter 2, section 2.4.1, is based on the assumption that the more similar two consonants are, the stricter the OCP effect will be. I will explore whether featural similarity influences OCP in Dutch, thus accounting for gradient OCP effects. This section will deal with Hypothesis 3 – degree of similarity influences the strictness of OCP.
As indicated above in § 2.1.4.1 on similarity, segments can be marked "+" or "-" for a particular feature, or they can be unspecified for that feature. The similarity matrix is based on shared classes of marked features only. Since the theoretical background for this thesis is Optimality Theory, marked features relevant for the similarity matrix, will here be considered marked features in OT. Hence the analysis in this section will refer to

multiple feature markedness, rather than be expressed as similarity. I will investigate [place] in 4.4.1 and [manner] in 4.4.2.

4.4.1 Place

4.4.1.1 [place] and [sonorant]

In this section I investigate whether Place features are influenced by additional feature markedness. I examine the three Place features, [labial], [dorsal] and [coronal], and explore if and how they are influenced by the Manner features [sonorant], [continuant] and [voice]. Since OCP-[place] in surface adjacent consonants is only visible in Dutch in word onset and is practically absolute for the features [labial] – O/E 0.04, and [dorsal], 0 instances, I will leave out surface adjacent consonants and concentrate on OCP across segments. Tables (46), (47) and (48) below will give the picture for Place harmonies in Dutch, also analyzed by the feature [sonorant]. I have calculated O/E scores per place category. Table (46) will start with [labial] and [sonorant], followed by [dorsal] and [sonorant] in (47) and [coronal] and [sonorant] in (48). I extracted all CVC and CXXC combinations from the CELEX database and coded consonants for both the features [place] and [sonorant]; intermediate segments were filtered out. Thus, all combinations violate the constraint OCP-[place] and some violate OCP [sonorant] as well. It is expected that combinations which violate both constraints show lower O/E rates than the ones that violate the place constraint only. In the table below, labial obstruent *p* includes /f, v, p, b/, while labial sonorant *m* represents /w[18], m/. *pp* and *mm* show double markedness of OCP-[LAB] and OCP-[αSON], *pm* and *mp* share the feature [labial] only.

[18] The glide /w/ was left out of the analysis for manner before. However, since it was included in the analysis for [place], the sound is here included as a labial sonorant.

(46) O/E scores for double feature violation of [labial] and [sonorant]

double feature violation of [labial] and [sonorant]		
distance	CVC	CXXC
items	10,378	6,302
pm	0.3572	0.9004
mp	0.5504	0.4945
pp	0.4786	0.6156
mm	0.4850	1.0836

X^2 15.94 df 3 P 0.0012

pm/mp : X^2 29,98 df 2 P < .0001
pp/mm : X^2 17.45 df 2 P 0.0002

It can be seen from the table that all labials across one segment are underrepresented. O/E scores range between 0.35 and 0.55. There is no dividing line between the two categories. In labial CXXC combinations there is no dividing line between single and double feature violation either. It is clear that a double OCP violation of OCP-[LAB] and OCP-[αSON] does not influence OCP-[labial] across segments.

The situation for dorsals is in table (47) below. Dorsal obstruent *k* includes / k, g, x, ɣ/, while dorsal sonorant *N* represents /ŋ/. *kk* and *NN* show double markedness of OCP-[DOR] and OCP-[αSON], *kN* and *Nk* share the feature [dorsal] only.

(47) O/E scores for double feature violation of [dorsal] and [sonorant]

double feature violation of [dorsal] and [sonorant]		
distance	CVC	CXXC
items	10,378	6,302
kN	0.5638	0.9347
Nk	0	0
kk	0.5133	0.9017
NN	0	0.9453

X^2 [19] *kN/Nk* : x
kk/NN : X^2 49.26 df 2 P <.0001

[19] It is clear from the table that for several cells in this category, the numbers of items are too low for a X^2-square test to be valid. Although, the validity of the outcome cannot be statistically established, the category is still included in the analysis.

It can be seen from the table that there are no dorsal CVC combinations beginning with a sonorant, so no valid chi-square analysis can be computed. Since there is no notable difference in O/E scores between dorsal combinations with [place] violation only and those with double violation, OCP-[DOR] and OCP-[αSON], and since in CXXC combinations, O/E rates are either absent or normal, we may say that there is no gradient OCP effect based on a double OCP feature violation of OCP-[DOR] and OCP-[αSON].

The outcome for coronals is in table (48) below. Coronal obstruent *t* includes /s, ʃ, z, ʒ, t, d/, while coronal sonorant *n* represents /l, n/. *tt* and *nn* show double markedness of OCP-[COR] and OCP-[αSON], *tn* and *nt* share the feature [coronal] only.

(48) O/E scores for double feature violation of [coronal] and [sonorant]

double feature violation of [coronal] and [sonorant]		
distance	CVC	CXXC
items	10,378	6,302
tn	1.1182	1.1447
nt	1.3957	1.0367
tt	0.5589	0.6688
nn	0.4596	0.9897

X^2 182.55 df 3 P <0.0001

tn/nt X^2 10.49 df 2 P 0.0053
tt/nn X^2 65 df 2 P <.0001

In CVC combinations in coronals in the table we observe a clear difference between O/E rates of combinations with a single violation, [coronal] only, and those with a double violation, OCP-[COR] and OCP-[αSON]. The combinations with double feature markedness are underrepresented. CXXC is restricted in the category of obstruents with double violation only. There is a clear indication of the effect of OCP-[sonorant] within the category of coronals across one segment.

From the above it is evident that labial and dorsal place harmony are not affected by double feature markedness, whereas coronal harmony is. [sonorant] is a secondary OCP feature influencing coronal harmony across one segment and coronal obstruent harmony across two segments.

4.4.1.2 [place] and [continuant]

The following tables give surveys of the O/E rates of homorganic consonant pairs separated by one and by two segments, with and without violation of the feature [continuant]. I will again treat the three place categories separately. In table (49) below, labial *p* includes [-CONT] /p, b, m/, labial *f* represents [+CONT] /f, v, w/. *pp* and *ff* show double markedness of OCP-[LAB] and OCP-[ɑCONT], *pf* and *fp* share the feature [labial] only.

(49) O/E scores for double feature violation of [labial] and [continuant]

double feature violation of [labial] and [continuant]		
distance	CVC	CXXC
items	10.378	6.302
pf	0.5427	0.5685
fp	0.2896	0.7665
pp	0.4392	0.6184
ff	0.1861	0.5782

X^2 13.44 df 3 p 0.0038

pf/fp X^2 20.37 df 2 P < .0001
pp/ff X^2 22.97 df 2 P 0.0029

The table shows for labials, that there is no distinction between the category with a single OCP-[place] violation and the category with double OCP violation of OCP-[LAB] and OCP-[ɑCONT]. In CXXC combinations there is no dividing line between the two categories, either. It is clear that double OCP markedness of [labial] and [continuant] does not affect OCP.

The situation for dorsals is in table (50) below. Dorsal *k* includes [-CONT] /k, g, ŋ/, dorsal *x* represents [+CONT] /x, ɣ/. *kk* and *xx* show double markedness of OCP-[DOR] and OCP-[ɑCONT], *kx* and *xk* share the feature [dorsal] only.

(50) O/E scores for double feature violation of [dorsal] and [continuant]

double feature violation of [dorsal] and [continuant]		
distance	CVC	CXXC
items	10.378	6.302
kx	0.3231	0.4237
xk	0.4195	0.6175
kk	0.6063	1.1840
xx	0.6212	0.6444

X^2 9.2 df 3 p 0.0267

kx/xk : X^2 5.13 df 2 P 0.0769
kk/xx : X^2 48.53 df 2 P <.0001

It can be seen from the table that, in dorsal CVC combinations, O/E scores for double violation of OCP-[DOR] and OCP-[ɑCONT] are higher than O/E scores for single violation. In CXXC combinations, we see the same – the double violation of OCP-[DOR] and OCP-[-CONT] even has a normal O/E score. We can say that double OCP violation of OCP-[DOR] and OCP-[ɑCONT] does not influence OCP.

Table (51) below contains the results for coronal harmony in connection with the feature [continuant]. Coronal *t* includes [-CONT] /t, d, n/, coronal *s* represents [+CONT] /s, ʃ, z, ʒ, l, /. *tt* and *ss* show double markedness of OCP-[COR] and OCP-[ɑCONT], *ts* and *st* share the feature [coronal] only.

(51) O/E scores for double feature violation of [coronal] and [continuant]

double feature violation of [coronal] and [continuant]		
distance	CVC	CXXC
items	10.378	6.302
ts	0.9434	1.0356
st	1.1585	0.8563
tt	0.6457	0.8533
ss	0.7415	0.8161

X^2 35.7 df 3 p < .0001

ts/st : X^2 34.32 df 2 P < .0001
tt/ss : X^2 42.94 df 2 P < .0001

It is clear from the table that the only categories that are underrepresented in CVC combinations are the double feature violations of OCP-[COR] and OCP-[αCONT]. This indicates that there is an effect of double feature markedness of the feature [continuant] on coronals.

From the above it is apparent that labial and dorsal place harmony are not affected by double feature markedness of [place] and [continuant]. There is an effect for coronal harmony with regard to the feature [continuant].

So far, we have no effect of [sonorant] or [continuant] on the Place categories of labials and dorsals; there is an effect of both the features [sonorant] and [continuant] on coronals.

4.4.1.3 [place], [sonorant] and [continuant]

This paragraph will take the investigation one step further. I will analyse whether degree of similarity of more than two features gradiently influences OCP. Based on the similarity matrix we would expect more violations of feature markedness to increase OCP restrictions. Thus, if the similarity matrix can be applied, there will be a gradient difference between a single violation, [place] only, a double violation, [place] and [sonorant] or [place] and [continuant], and a triple violation, [place], [sonorant] and [continuant]. The three place categories will again be treated separately.
In table (52) below, labial *p* includes [-SON] and [-Cont] /p, b/, labial *f* represents [-SON] and [+CONT] /f, v/ and labial *m* represents [+SON] and [-CONT] /m/. Thus, the combinations *pp, ff* and *mm* share three violations of feature markedness, *pf* and *fp* show double feature markedness of OCP-[LAB] and OCP-[αSON], *pm* and *mp* share double feature markedness of OCP-[LAB] and OCP-[αCONT]. The combinations *fm* and *mf* only violate OCP-[LAB]. X^2 tests are based on three categories, viz. triple violation, double violation, single violation.

(52) O/E scores for multiple feature violation of [labial], [sonorant] and [continuant]

multiple feature violation of [labial], [sonorant] and [continuant]		
distance	CVC	CXXC
items	10,378	6,302
pp	0.6024	0.6082
ff	0.1864	0.5786
mm	0.4348	0.7953
pf	0.5291	0.7550
fp	0.2663	0.4731
pm	0.3882	0.8258
mp	0.2271	0.1224
fm	0.3116	1.1766
mf	0.5722	0.0698

triple violation :	X^2	8.21	df 3	P 0 .0419
double violation :	X^2	17.22	df 4	P 0 .0018
single violation :	X^2	22.72	df 2	P < .0001

Table (52) shows that CVC combinations in labials are all affected by OCP. The triple violation in labial fricatives, '*ff*', has the lowest score for labials, which would confirm strictest OCP for triple violation; however, '*pp*' combinations, also triple violation, have the highest O/E score of all, so similarity cannot be the issue here. In the table we can see no dividing line between categories based on the number of violations in CVC combinations, nor in CXXC combinations. Labials are not affected by multiple markedness.

The picture for dorsal similarity is in table (53) below. Dorsal *k* includes [-SON] and [-CONT] /k, g/, dorsal *x* represents [-SON] and [+CONT] /x, ɣ/ and dorsal *N* represents [+SON] and [-CONT] /ŋ/. Thus, the combinations *kk, xx* and *NN* share three violations of feature markedness, *kx* and *xk* show double feature markedness of OCP-[DOR] and OCP-[αSON], *kN* and *Nk* share double feature markedness of OCP-[DOR] and OCP-[αCONT], the combinations *xN* and *Nx* only violate OCP-[DOR].

(53) O/E scores for multiple feature violation of [dorsal], [sonorant] and [continuant]

multiple feature violation of [dorsal], [sonorant] and [continuant]		
distance	CVC	CXXC
items	10,378	6,302
kk	0.6345	1.2652
xx	0.6207	0.6449
NN	0	0.9453
kx	0.3390	0.6696
xk	0.4085	0.4431
kN	0.6357	1.2621
Nk	0	0
xN	0.4424	0.3039
Nx	0	0

triple violation : X^2 51.53 df 3 P < .0001
double violation : X^2 8.15 df 4 P 0 .0862
single violation : x

The table shows that there is no dorsal harmony across one segment of *NN, Nk* and *Nx* in the corpus. Hence, no significance can be established for single violation of *kN/Nx*; moreover O/E rates for double markedness in dorsals are not significant. I will therefore leave out the category of dorsals in the present analysis.

The picture for coronals is in table (54) below. Coronal *t* includes [-SON] and [-CONT] /t, d/, coronal *s* represents [-SON] and [+CONT] /s, ʃ, z, ʒ/, coronal *n* represents [+SON] and [-CONT] /n/, and coronal *l* represents [+Son] and [+Cont] /l/. Thus, the combinations *tt, ss* and *nn* and *ll* share three types of feature markedness, *ts, st, nl* and *ln* show double feature markedness of OCP-[COR] and OCP-[αSON], *tn, nt, sl* and *ls* share double feature markedness of OCP-[COR] and OCP-[αCONT], the combinations *sn, ns, tl* and *lt* only violate OCP-[COR].

(54) O/E scores for multiple feature violation of [coronal], [sonorant] and [continuant]

multiple feature violation of [coronal], [sonorant] and [continuant]		
Distance	CVC	CXXC
Items	10,378	6,302
tt	0.4563	0.7621
ss	0.5348	0.3079
nn	0.1861	0.7309
ll	0.1024	0.1401
ts	0.6160	0.9373
st	0.6622	0.7439
nl	0.5131	2.7749
ln	1.0648	0.8752
tn	0.7845	1.4769
nt	1.2489	0.6016
sl	1.3268	1.4597
ls	1.2998	1.5758
sn	1.3359	0.8528
ns	1.5138	0.6377
tl	1.1614	0.7755
lt	1.4880	1.1386

3 violations : X^2 56.89 df 1 $p < .0001$
2 violations : X^2 78.48 df 1 $p < .0001$
1 violation : X^2 26.51 df 1 $p < .0001$

The table indicates that coronal CVC combinations with triple OCP violations have the lowest O/E scores. The scores for double feature markedness of OCP-[COR] and OCP-[αSON] – with the exception of /ln/ – also show underrepresentation, whereas the scores for double feature markedness of OCP-[COR] and OCP-[αCONT] are normal. In CXXC combinations scores for triple markedness are all below normal or only just within limits. There is no dividing line between the other categories. There is clearly a gradient effect connected with multiple markedness on coronal CVC strings. As indicated above in table (48), CVC combinations in coronals are affected by the feature [sonorant]. A violation of OCP-[αCONT] was found to affect coronal OCP across a single segment in table (43) above. In the present analysis, however, no OCP effect for continuant is visible. We must conclude that OCP-[COR] is influenced by additional markedness of OCP-[αSON] and that, contrary to earlier conclusions, [continuant] is not an additional feature influencing OCP-[COR], but an additional feature influencing the conjoined markedness of OCP-[COR]&OCP-[αSON].

With regard to multiple OCP violations of the place features [labial], [dorsal], [coronal], combined with the manner features [sonorant] and [continuant], it may be stated that similarity, or multiple OCP violations of [place] and [sonorant], influences OCP in coronals. There is a gradient effect related to degree of similarity in the features [COR], [αSON], [αCONT]. It is also the case that multiple OCP violations have no effect on labials or dorsals. We may conclude that multiple markedness in connection with coronal place harmony is an aspect that influences OCP effects in Dutch. The relevant features for Dutch are [sonorant] and [continuant].

4.4.2 Manner

4.4.2.1 [sonorant] and [continuant]

As indicated above, there was a gradient effect of similarity on the place feature [coronal]. In this paragraph the aspect of similarity among manner categories only will be investigated. I will start with double feature markedness of [sonorant] and [continuant], irrespective of [place]. I have taken the entire corpus of monomorphemic words as basis for the calculations. I extracted all CVC and CXXC combinations and coded consonants for the features [sonorant] and [continuant]; intermediate segments were filtered out. Consonant pairs were coded in accordance with the number of manner violations, i.e. violation of both OCP-[αSON] and OCP-[αCONT], indicated as OCP ** in the table, violation of either OCP-[αSON] or OCP-[αCONT], indicated as OCP *, or no OCP [manner] violation at all. The results are in table (55):

(55) O/E scores for double feature violation of [sonorant] and [continuant]

double feature violation of [sonorant] and [continuant]		
	CVC	CXXC
items	30,378	6,302
OCP **	0.9833	0.9720
OCP *	1.1110	1.0685
no OCP	1.0500	1.0954

X^2 582.01 df 2 $p < .0001$

It can be seen in the table that all O/E scores are normal. This would suggest that there is no influence of similarity on the manner features [sonorant] and [continuant]. However, as OCP restrictions on the place features were found in the place feature [COR] only, it might be the case that this general picture should be adapted for a single feature. I therefore split up the three categories into their respective combinations and calculated O/E scores per separate category. The results are in table (56) below. In the table, *p* includes [-SON] and [-CONT] /p, b, k, g, t, d/, *f* represents [-SON] and [+CONT] /f, v, x, ɣ, s, ʃ, z, ʒ/, *n* includes [+SON] and [-CONT] /m, n, ŋ/ and *l* represents [+SON] and [+CONT] /r, l/. The pairs *pp, ff, nn* and *ll* share double feature markedness of [sonorant] and [continuant], the combinations *pf, fp, nl, ln, pn, np, fl* and *lf* share a single feature, either [sonorant] or [continuant]; the pairs *pl*, *lp*, fn and *nf* do not share [manner] features. Table (56) below shows the O/E rates for single and double feature markedness per feature category.

(56) O/E scores for double feature violation of [sonorant] and [continuant]

double feature violation of [sonorant]] and [continuant]		
distance	CVC	CXXC
items	10,378	6,302
pp	0.8203	1.0908
ff	0.8013	0.8823
nn	0.6756	0.5588
ll	0.2924	0.8393
pf	0.7715	1.0780
fp	0.7908	0.9241
nl	1.0096	2.4748
ln	1.2427	0.8212
pn	0.8422	1.1475
np	1.1533	0.6876
fl	1.2408	1.3195
lf	1.4768	1.1263
pl	1.3819	0.5171
lp	1.3719	1.0190
fn	1.0137	1.1254
nf	1.0317	0.7377

As can be seen from the table, in CVC strings all combinations but two, *nn* and *ll*, and in CXXC strings only *nn* (double OCP-[manner] violation), *np* (single OCP-violation) and *pl* and *nf (*no OCP violation)*,* have normal O/E scores. It is clear that the effect of double markedness only occurs within the category of [+SON] and that there is no gradient effect on OCP-[manner] in general with double feature markedness of [sonorant] and [continuant] versus single feature markedness or no feature markedness.

4.4.2.2 [+Son] and [nasal]

As seen above in table (56), the double violations in consonants across segments of OCP-[+SON][-CONT], *nn*, and OCP-[+SON][+CONT], *ll,* are restricted in Dutch. It was also established in table (45) above that O/E scores for [αSON] CVC strings are within normal limits. I will here investigate the additional influence of the feature [nasal] on sonorants. Since the feature [nasal] is only contrastive within the category of sonorants, OCP-[αNAS] is here treated as a secondary OCP-marker. Table (57) gives the results for the combinations of consonants based on the division of CVC and CXXC strings in the corpus coded by sonority category.

(57) O/E scores for OCP restrictions within sonorants

OCP restrictions within sonorants		
	CVC	CXXC
items	10,378	6,302
nn	0.6756	0.5588
ll	0.2924	0.8393
nl	1.0096	2.4748
ln	1.2427	0.8212

X^2 96.47 df 3 P < .0001

It is clear from the table that nasals are restricted both across one segment and across two. Liquids are limited across one segment, not across two segments. The O/E scores for strings combining liquid and nasal are normal or above normal at both distances. This indicates that the feature [nasal] separates the nasals [+NAS] from the liquids [-NAS] in the category of sonorants and is responsible for the limitations found within the category of sonorants.

There was no OCP-[sonorant] in consonants across one segment in Dutch, as seen in table (45) above, although scores within the normal range differed significantly with distance. This has now been explained by the additional influence of OCP-[nasal] on OCP-[+Son].

4.4.2.2.1 [-Nas] and [lateral]

As indicated in Chapter 2, section 2.3.2, the feature [lateral] has been shown to influence OCP within the category of liquids, both in English (Berkley 2000) and in Latin (Berkley 2000, Van der Torre 2003). I expect the feature [lateral] to also influence liquids in Dutch. Table (58) shows the results for single and double markedness of liquids.

(58) O/E scores for double feature violation of [liquid] and [lateral]

double violation of liquid and [lateral]		
distance	CVC	CXXC
items	10,378	6,302
ll	0.1024	1.2438
rr	0.0448	0.8706
lr	0.5308	1.2116
rl	0.5079	0.8773

X^2 31.38 df 3 $p < .0001$

As can be seen from the table, the combinations with lateral harmony in consonants across segments score at least 0.39 lower than the mixed pairs, both across one segment and across two. It is evident that multiple markedness of OCP-[-NAS]&OCP-[ɑLAT] influences OCP. OCP-[ɑLAT] can be considered an additional feature influencing OCP-[-NAS].

4.4.2.3 [sonorant] and [voice]

As noted above in section 4.2c, no OCP-[voice] was found in Dutch. However, [voice] could still be present as a secondary feature. Since the category of [+SON] is unspecified for [voice], sonorants are left out. This section investigates the influence of [voice] in double feature markedness of [-SON] and [voice]. An analysis of a combination of the features [-SON] and [voice] is in table (49) below. [-VOI] represents voiceless obstruents, [+VOI] represents voiced obstruents.

(59) O/E scores for double feature violation of [sonorant] and [voice]

double feature violation of [-Son] and [voice]		
distance	CVC	CXXC
items	10,378	6,302
[-VOI]	0.8283	0.9921
[+VOI]	0.9685	1.5230

X^2 14.82 df 1 $p < .0001$

The table shows normal scores and one score indicating overrepresentation. Since there is no underrepresentation, it is clear that there are no OCP restrictions based on the feature [voice] in connection with [-SON]. OCP [voice] is not a constraint active in the Dutch language as a secondary markedness feature in OCP.

4.4.3 Total similarity

Thus far we have established that multiple feature markedness plays a limited role in Dutch. In connection with [place], multiple feature markedness plays a role in coronals, not in labials and dorsals. In connection with [manner], multiple feature markedness plays a role within the category of sonorants and liquids. This leaves us with the final issue that came up in relation to similarity in § 2.1.4.1, viz. total similarity versus partial similarity. This section will deal with the aspect of similarity in identical segments. It was established in the literature, discussed in Chapter 2, that in some languages identical segments are exempt from OCP restrictions. Whereas, in agreement with the theory on similarity, one would expect strictest OCP for totally similar segments and less strict OCP for less similar segments combinations, this appeared not to be the case in several languages. I will here examine the situation in detail for Dutch. The table below shows the scores per labial, dorsal and coronal CVC and CXXC combination in the corpus. The general O/E scores for labial, dorsal and coronal CVC and CXXC, taken from table (37) above, are indicated in row 4 in the tables. These are the average O/E scores when all place harmonic consonants are included, i.e. identical as well as partly similar pairs.

(60) O/E scores for total identity labial, dorsal and coronal

labial total identity		
distance	CVC	CXXC
general picture	0.4630	0.7174
/pp/	0.8574	0.6782
/bb/	1.3437	2.7036
/ff/	0.1903	0.8147
/vv/	0.4042	0.8113
/mm/	0.4348	0.7953
/ww/	0.5123	0.8432

dorsal total identity		
distance	CVC	CXXC
general picture	0.5078	0.8643
kk	0.6494	1.3148
gg	0	0
xx	1.2476	0.7851
GG	0	0
NN	0	0.9453

coronal total identity		
distance	CVC	CXXC
general picture	0.8709	0.8725
tt	0.4921	0.8917
dd	0.3841	0.7604
ss	0.6863	0.3048
zz	0	0
nn	0.1861	0.7309
ll	0.1024	0.1401

O/E scores in the table for CVC labial plosives, /pp/ and /bb/, show normal and above normal O/E rates. The totally similar other CVC combinations in labials have scores that indicate underrepresentation in various degrees. When four places apart, plosive /pp/ harmony has the lowest score of all, the other scores are normal or show overrepresentation.

With regard to dorsals, we may notice that dorsal fricative /xx/ across one segment has normal O/E scores and dorsal plosive /kk/ across one segment is underrepresented. The scores across two segments that are present are normal.

Coronal harmony across one segment is underrepresented in all identical consonant pairs, coronal harmony across two segments is restricted in four out of six categories as well.

The normal and above normal scores for labial /pp/ and /bb/ would seem to indicate that total similarity in labials is accepted in Dutch. However, with the other four labial consonants being restricted, total similarity cannot be the issue here.

With regard to dorsals, it is difficult to draw conclusions when so many categories are not represented at all. With only two out of five dorsal consonants present as harmonic pairs, I will leave dorsals out of the analysis here.

Coronals show a different picture. Whereas the general picture for coronals across one segment indicates no OCP restrictions at all, totally similar coronals are all restricted. This is of course connected with the findings on multiple markedness above. It was established that coronal OCP was gradiently influenced by multiple feature markedness, with coronal sonorants as most restricted categories on account of triple and quadruple violations of OCP-COR, OCP-[+SON], OCP-[αNAS] and OCP-[αLAT]. Total similarity implies multiple markedness, so coronals can be said to be influenced by featural multiple markedness. There are no exceptions for coronal identical segments.

This concludes the investigation on similarity as a defining factor on (gradient) OCP in Dutch. The influence of similarity is found in Dutch, although not as a general rule for all features. Homorganic coronal consonants are influenced by additional feature similarity; the relevant features are [sonorant] and [continuant]; the manner category of [sonorant] is influenced by the feature [nasal], the category of liquids is influenced by the feature [lateral]. With regard to total similarity, we may state that total similarity is no exception to OCP regulations in Dutch. Based on the findings for coronals, sonorants and liquids, we may conclude that Hypothesis 3 – degree of similarity influences the strictness of OCP – has been corroborated for Dutch.

4.5 OCP and Prosodic Salience

So far I have treated distance and similarity, or multiple markedness, as defining factors in the gradient nature of OCP. I will now address a third aspect that has been shown to be of influence on the gradient nature of OCP, viz. salience. The present section will be concerned with Hypothesis 4 – the strictness of OCP is connected with prosodic salience. Since I will explore three aspects of prosodic salience, viz. position in the word, word stress, and sonority, Hypothesis 4 will be divided into three separate hypotheses:

Hypothesis 4 (1): the strictness of OCP is connected with position in the word.

Hypothesis 4 (2): the strictness of OCP is connected with word stress.

Hypothesis 4 (3): the strictness of OCP is connected with sonority.

Since the influence of distance has been established as a defining factor in gradient OCP, the aspect of distance will in this section be left out; only CVC strings will be examined.

4.5.1 Position in the Word

As indicated above in § 2.4.3, Berkley (2000) noticed for English three degrees of salience with respect to position in the word: strictest OCP was found in initial closed syllables, followed by CVC strings across initial syllable and finally CVC strings within final closed syllable. I expect the same left-to-right effect of gradience to be valid for Dutch. In order to investigate the subject I divided the CVC strings in the subdatabase of 3,818 polysyllabic[20] words in the corpus into CVC with initial C in word onset – irrespective of tautosyllabic or heterosyllabic strings, /**kɑn**to:r/ ‘office’, /b**rɑn**dən/ ‘burn’, /**k**a:**t**ər/ ‘tom-cat’, /s**t**a:**p**əl/ ‘pile’, word-medial CVC – irrespective of tautosyllabic or heterosyllabic strings, /a**mɑn**dəl/ ‘almond’, /xlɪns**tər**ən/ ‘shine’, and CVC with final C in word coda, /kɑn**t**o:**r**/ ‘office’, /bɔrs**t**ə**l**/ ‘brush’. I extracted all CVC strings and coded consonants in accordance with place of articulation. Since the feature [sonorant] was established to be an additional feature influencing OCP-[COR] in Dutch, I split up the category of coronals into obstruents, coronal [-SON], and sonorants, coronal [+SON]. Hence there are four place categories, i.e. labial, PP, dorsal, KK, coronal [-SON], TT, and coronal [+SON], NN. Intermediate vowels were filtered out. I then calculated O/E rates by position in the word. Table (61) below shows the results for the three separate categories based on position in the word. The numbers of items per category are in the table.

[20] There is a difference in the ways in which O/E scores for English by Berkley and for Dutch in the present thesis were established. I will come back to this in section 5.2.1.

(61) O/E scores for CVC harmony by position in the word

CVC position in the word			
position	wd onset	word-medial	wd coda
items	4,340	1,908	1,546
PP	0.4789	0.2574	0.2472
KK	0.5753	0.1186	0.3554
TT	0.6109	0.4646	0.0580
NN	0.5034	0.3981	0.5142

	X^2 128.26	df 6	p <.0001
PP:	X^2 78.93	df 2	p <.0001
KK:	X^2 36.45	df 2	p <.0001
TT:	X^2 31.95	df 2	p <.0001
NN:	X^2 6.23	df 2	p 0.0444

It can be seen from the table, that O/E scores for labial, dorsal and coronal [-SON] harmony with initial C in word onset are higher than those for harmony further down in the word. For labials there is no difference between word-medial harmony and harmony with final C in word coda. Dorsals have higher O/E scores for harmony with final C in word coda than for word-medial harmony. Coronal obstruent CVC is avoided in word coda, coronal sonorants are most restricted word-medially, whereas they have equal O/E scores for word begin and word end.

We may conclude that, in general, harmony is permitted more easily in CVC strings with initial C in word onset than in CVC strings further down in the word. There is a clear indication of the effect of position in the word on OCP. Hypothesis 4 (1) – the strictness of OCP is connected with position in the word – has been corroborated for Dutch.
It must be noted that the effect is contrary to what was expected, based on Berkley's findings for English. More salience in Dutch entails less strict OCP. This coincides with other findings in the literature. With regard to word initial position, it has been established that "...the phonological behaviour of segments is different in syllable initial position as opposed to syllable-final position... In general, languages have more distinct syllable onsets than they do syllable codas" (Ohala and Kawasaki 1984, p 115). "It is wellknown that contrasts are best realized in *perceptually salient* positions... (Kager 1999, p 408). The issue will be discussed further in section 5.2.1.

4.5.2 Word Stress

Another aspect of prosodic salience claimed to be of influence in OCP is word stress. This section deals with Hypothesis 4 (2) – the strictness of OCP is connected with word stress. In order to check expectations I took the subdatabase of 3,818 polysyllabic monomorphemic words and coded consonants in accordance with place of articulation, as well as their being part of a stressed syllable or unstressed syllable in the word. Intermediate vowels were filtered out. The results are in table (62).

(62) O/E scores for harmony in stressed syllables and in unstressed syllables

CVC harmony in stressed and unstressed syllables		
	stressed	unstressed
items	5,866	3,361
pp	0.6024	0.5215
kk	0.7231	0.5520
tt	0.6383	0.5489
nn	0.6401	0.4837

	X^2	89	df 3	P 0.8272
PP:	X^2	34.67	df 1	P <.0001
KK	X^2	11.91	df 2	P 0.0006
$TT_{[-SON]}$	X^2	48.73	df 2	P <.0001
$TT_{[+SON]}$	X^2	33.12	df 2	P <.0001

It can be seen in the table that all categories are underrepresented and that all combinations are in favour of harmony in stressed syllables.

From the significant differences between stressed and unstressed place harmonic CVC strings we may conclude that stress is a contributing factor in OCP in Dutch in labial, dorsal and coronal harmony. Hypothesis 4 (2) – the strictness of OCP is connected with stress – is corroborated.

It can be noted that the effect of OCP is again contrary to the effect in English, where stricter OCP was found in stressed syllables. In Dutch, there is a preference for OCP in unstressed syllables, rather than stressed, again implying: more salience – less strict OCP. I will come back to this in section 5.2.1.

4.5.3 Sonority

As stated in Chapter 2, Elmedlaoui (1995) found for Classical Arabic and other ancient Semitic languages that rising sonority in a homorganic CVC pair was preferred over falling sonority. In accordance with this I expect the same to be applicable to Dutch. This paragraph will explore Hypothesis 4 (3) – the strictness of OCP is connected with sonority. In order to check the effect of sonority on OCP in Dutch I took the subdatabase of 10.378 CVC strings as basis and coded consonants in accordance with place of articulation, as well as with rising, falling or equal sonority. Intermediate vowels were filtered out. O/E scores were calculated per category of rising and falling sonority[21]. Since rising or falling sonority is connected with sonority distance across sonorancy classes, I have here taken the category of coronal obstruents and sonorants together and calculated O/E scores over the entire category. The results are in table (63).

(63) O/E scores for CVC harmony and sonority

CVC harmony and sonority		
	rising	falling
	4,879	4,050
PP	0.4570	0.4446
KK	0.7199	0.6554
TT	0.8730	0.9748

X^2 7 df 2 p 0.0302

PP: X^2 2.15 df 1 p 0.1426
KK: X^2 3,92 df 1 p 0.0477
TT: X^2 0.43 df 1 p 0.512

As expected, both labials and dorsals are underrepresented in both categories. It is clear from the table that there is no significant difference for labials and dorsals between harmonic CVC strings with rising or falling sonority. Coronals are within normal range for both categories. It looks as if there is a gradient difference for coronal consonant strings in favour of harmony in CVC combinations with falling sonority. However, the chi-square

[21] Since homorganic CVC strings within the same sonority class share the features [place], [sonorant] and [continuant] – a matter which was dealt with in section 4.4 on multiple markedness – the category of equal sonority is not included in the present analysis.

measure indicates p 0.512, hence the result cannot be considered significant. This means that there is no evidence in favour of Hypothesis 4 (3) – the strictness of OCP is connected with sonority.

This concludes the investigation on salience as a secondary influence on the gradient nature of OCP. The three aspects of salience discussed in the literature on OCP in Chapter 2, i.e. position in the word, word stress and sonority, were examined for Dutch. It was found that position in the word influences labial, dorsal and coronal OCP. The more salient position in the word, viz. at word onset, is more permissive of harmony than the less salient positions. With regard to word stress it was established that stress is a defining factor in labial, dorsal and coronal OCP in Dutch. CVC strings in stressed syllables show less OCP restrictions than CVC strings in unstressed syllables. The aspect of sonority was found not to influence OCP in Dutch. With evidence in favour of the influence of salience in position in the word and word stress, we may conclude that Hypothesis 4 – The strictness of OCP is connected with prosodic salience – is corroborated.

4.6 Other influences on OCP

It was established in the previous sections in this chapter that there is OCP in Dutch, that distance plays a role, that multiple OCP markedness plays a role and that position of two harmonic consonants in the word as well as word stress influence OCP. An additional aspect that could influence the effect on OCP is the quantity and/or quality of an intervening vowel. Whereas autosegmental phonology separated consonants from vowels by appointing them to separate tiers underlyingly, Optimality Theory operates with surface-based phonology. OT does not necessarily separate the two categories. It might well be the case that vowel features should be included in an OCP analysis. This will be investigated below in § 4.6.1. I expect qualities of a vowel separating two consonants to influence the strictness of OCP. Another aspect that might be of influence on OCP is related to multiple feature markedness, treated in section 4.4 above on OCP and similarity. I will investigate whether multiple markedness of an OCP constraint combined with a non-OCP constraint, viz. cluster markedness, influences OCP. These two aspects will be discussed in the present section. I will start with the quantity and quality of vowels separating two harmonic consonants in § 4.6.1. This will be followed by an investigation into the role of cluster membership as a possible influence on the strictness of OCP in § 4.6.2. Two hypotheses will be dealt with:

Hypothesis 5: OCP across segments is influenced by intervening vowels

Hypothesis 6: OCP across segments is influenced by multiple markedness of an OCP constraint with cluster membership.

4.6.1 Intermediate Vowels

This section will be concerned with vocalic segments separating two consonants. Two aspects will be considered, viz. the quantity of the intervening vowel and the quality of the intervening vowel. With regard to the quantity of the intervening vowel, I refer to Chapter 4.3 on distance as a defining factor on OCP. Although I have so far treated short and long vowels as single segments, and diphthongs as two segments, I will here temporarily depart from this view and analyze the three categories as separate categories distinguished by length of the vowel. Berkley (2000) stated for English that the OCP effect is strongest across a short vowel and gets weaker with distance. In connection with this, short vowels could be less permissive of harmony than long vowels, which in their turn could be less permissive of harmony than diphthongs. With respect to the quality of the intervening vowel, I expect vowels with the same place of articulation as the surrounding consonants to be less permissive of consonant harmony than vowels with a different place of articulation. Since I will examine two different aspects of the intervening vowels, Hypothesis 5 will be divided into two separate hypotheses:

Hypothesis 5 (1): OCP across segments is influenced by the quantity of an intervening vowel.

Hypothesis 5 (2): OCP across segments is influenced by the quality of an intervening vowel.

4.6.1.1 Vowel quantity

As indicated above I expect the effect of distance between two consonants to be connected with the quantity of an intermediate vowel. In order to be able to calculate scores I extracted CVC combinations, irrespective of position in the word, as indicated in the example on CVC combinations in table (31) above. Subsequently, all labial consonants, /p, b, f, v, m, w/, were coded P, all dorsals, /k, g, x, ɣ, ŋ/, were coded K, and all coronals, /t, d, s, ʃ, z, ʒ, n, l/, were coded T. Since I believe restrictions connected with intervening vowel(s) not to be connected with the feature [sonorant] I took the categories of coronal obstruents and sonorants together. Intermediate vowels were coded for vowel length in three categories, short vowels /ɛ, ɑ, ɔ, ɪ, ʉ, ə/, long vowels /i, e, a, o, y, u, ø/ and diphthongs /ɛi, œy, ɑu/. O/E scores were thus calculated on the basis of the Place features and quantity of the intermediate vowel. O/E rates were calculated over the number of tokens with the particular segment combination in the corpus. The number of tokens per combination is in row three in the table below. It must be noted that the number of consonant combinations across a diphthong is below 1,000, the number indicated in 4.1 as a sufficient number of items to render a reliable score. The category is nevertheless included in the analysis.

(64) O/E scores for vowel length and OCP

vowel length and OCP			
CVC	short	long	diphthong
items	5,982	4,394	714
PP	0.4810	0.4469	0.7178
KK	0.4388	0.6293	0.4976
TT	0.8726	0.8691	0.9406

	X^2	9.63	df 4	p 0.0471
PP:	X^2	13.21	df 2	p 0.0014
KK:	X^2	1.51	df 2	p 0.47
TT:	X^2	30.17	df 2	p <.0001

It can be seen that coronals have normal O/E scores in all three vowel categories, with a slightly higher O/E score for harmony across diphthongs. There is no significant difference between the O/E scores for harmony across short vowels and across long vowels in labials; there is a difference in scores between labial monophthongs and diphthongs. It is evident from the chi-square computation on dorsals, that the figures for dorsals are not significant. I will therefore ignore dorsals with regard to the outcome of this table.

It is clear that there is no effect of single vowel length on labials and coronals. Both the differences between labial and coronal harmony across monophthongs and across diphthongs can be considered an indication of a gradient difference between the two categories. As seen above in section 4.3 on distance, which was established to be a factor influencing labial OCP in Dutch, it is clear that distance between two consonants separated by a single consonant, long or short, increases when diphthongs separate the consonants, just as there is an effect of distance with two intervening segments of which one is a consonant. Since vowel length in monophthongs does not influence OCP in Dutch, we may conclude that diphthongs must be regarded as consisting of two segments. Accepting the view that diphthongs are two-segmental, Hypothesis 5 (1) – OCP across segments is influenced by the quantity of an intervening vowel – must be rejected.

4.6.1.2 Vowel quality

As an introduction to the aspect of vowel quality I refer to a phenomenon present in developing grammars of children acquiring (the Dutch) language, viz. Consonant Harmony. Consonant Harmony, henceforth also referred to as CH, is a process in which one consonant assimilates to another. Fikkert and Levelt (2005) attend to the quality of the intervening vocalic segment on CH. They did longitudinal research on developmental grammars and found that in the first stage of Consonant Harmony in Dutch all words are harmonic, i.e. completely coronal, labial or dorsal. The place feature is on the entire word. Words consist either of coronal consonants and front, i.e. coronal vowels, or of labial consonants and round, i.e. labial vowels, or dorsal consonants and back, i.e. dorsal vowels. The low vowels can appear with any combination of consonants, and are considered neutral vowels. Pater and Werle, (2001), found the same in the developing grammar of Trevor acquiring the English language. Based on these findings on Consonant Harmony, a phenomenon contrary to OCP, I expect OCP to be influenced by an intervening vowel. Since OCP is about dissimilation, I expect that when two consonants share place of articulation, the intervening vowel preferably possesses a different place of articulation. Hence, I expect front vowels to be combined with back consonants and vice versa, whereas labial consonants are expected to avoid rounded vowels. I expect to find the neutral vowels /a/, /ɑ/ and /ə/ with all three place combinations.

In order to be able to calculate scores I extracted CVC combinations and coded all labial consonants P, dorsals K, and coronals T. Intermediate vowels were coded in accordance with four places of articulation, i.e. front unrounded /i, ɪ, e, ɛ/, front rounded /y, ʉ, ø/, back rounded /o, ɔ, u/, and neutral /ɑ, a, ə/. It must be noted that the number of items for CVC combinations across a front rounded vowel are below 1,000.

(65) O/E scores for OCP across vowels of different quality

OCP across intermediate vowel with different quality				
CVC	front unround	front round	back	neutral
items	3,282	877	2,181	4,036
PP	0.4547	0.4878	0.6147	0.3865
KK	0.6502	0.4893	0.8484	0.3023
TT	0.8719	0.8730	0.9222	0.8461

X^2 29.56 df 6 $p < .0001$

PP: X^2 20.74 df 3 p 0.0001
KK: X^2 17.15 df 3 p 0.0007
TT: X^2 7.06 df 3 p 0.07

It can be seen from the table that coronals have normal O/E scores in all four categories, with the highest score across back vowels. With labials, harmony across back vowels has the highest score. The difference in O/E scores for dorsals between harmony across a back vowel and a front or neutral vowel is in favour of harmony across a back vowel. The lowest scores in labials and dorsals are across neutral vowels.

Even though coronals prefer back vowels, as expected, the fact that labials prefer rounded back vowels to neutral or front vowels and dorsals prefer back vowels to front vowels, is not what we expect if OCP extends over vowels. It is clear that assimilation of place in labials and dorsals is preferred over dissimilation when including the vowel in the analysis. Moreover, where neutral vowels were expected to be the least affected by harmony across a vowel, they turn out to have the lowest harmony scores in all three place categories. We must therefore conclude that Hypothesis 5 (2) – OCP across segments is influenced by the quality of an intervening vowel – must be rejected.

4.6.2 Multiple Markedness

We saw above in section 4.4 that multiple feature-markedness of e.g. OCP-[COR] and OCP-[ɑSON] played a role in the strictness of OCP. We also recognized the psychological reality of OCP in a perception experiment in English with CCVC words beginning with sC (Coetzee, 2005)[22]. Coetzee (2008) mentions a conjoined constraint prohibiting a combinations of initial /s/ with CVC of two identical consonants, *skɑk, *spɑp. It became evident in § 4.5.1 that Dutch is more permissive of harmony across segments at word onset than further down in the word. It was not clear how the difference in O/E scores between CVC within initial syllable, O/E [Lab] 0.3483 and [Dor] 0.5051, and CVC across initial syllable, O/E [Lab] 0.5110 and [Dor] 0.6848, should be interpreted. In the present section I will combine the aspects indicated above and investigate the syllable as domain of OCP, rather than the word, and whether multiple markedness of one OCP constraint together with cluster markedness, influences the effect of OCP in monomorphemic words in Dutch. I will consider the difference between a possible influence of cluster membership versus single consonants in general, as well as a difference in a possible influence connected with cluster-membership and position of the CVC string in the word. Also the aspect of initial extra-syllabic /s/ will be taken into account. This section will deal with Hypothesis 6 – OCP is influenced by multiple markedness of an OCP constraint with cluster markedness.

[22] See Chapter 2, section 2.1.5

4.6.2.1 OCP syllable

In order to establish whether OCP is more restricted within syllables than across syllables in Dutch, I used the subdatabase of all polysyllabic words in the corpus, 3,818 polysyllabic monomorphemic words. From this subdatabase I extracted all CVC strings and split them up into CVC within syllable and CVC across syllable, and coded them in accordance with place of articulation. Intermediate vowels were filtered out. I then calculated O/E rates for the two categories. The results are in table (66) below.

(66) O/E scores for CVC within syllables and across

CVC within syllable and across		
items	2,801	4,920
CVC	within syll	across syll
PP	0.3573	0.4853
KK	0.4895	0.5234
TT	0.4672	0.5592
NN	0.7270	0.2883

	X^2 89.25	df 3	p <.0001
PP:	X^2 31.65	df 1	p <.0001
KK:	X^2 2.38	df 1	p 0.1229
TT:	X^2 5.9	df 1	p 0.0151
NN:	X^2 47.15	df 1	p <.0001

The table shows that O/E rates for labial, dorsal and coronal [-SON] harmony across a single segment within syllables are more restricted than the O/E scores across syllable. This would suggest stricter OCP within syllable than across. It is also possible that another aspect influences the differences in scores between harmony within syllable and harmony across syllable, viz. it could be the case that the difference is connected with cluster membership of (one of the two) consonants. This will be investigated in the next section.

4.6.2.2 The influence of clusters

In order to investigate the influence of cluster membership on OCP in CVC strings, I extracted all CVC combinations from the corpus and coded consonants in accordance with place of articulation, and as single consonant or cluster member. Vowels were filtered out. O/E scores for CVC harmony of two single consonants versus CVC harmony with at least one of the two consonants as part of a cluster are in table (57) below. In the table 'single C' indicates two single consonants separated by a vowel, including CVC words, /**k**ɑ**t**/ 'cat', VCVC words and combinations, /a:**d**ə**l**/ 'nobility', /fɑ**m**i**l**i/ 'family', CVCV words and combinations, /**b**e:**b**i/ 'baby', /**k**o:**k**ən/ 'cook', /**f**ɑ**m**ili/ 'family'; 'cluster CC' indicates that either the initial or the final consonant of a place-harmonic CVC pair is part of a cluster, /s**t**ɑ**l**/ 'stables', /s**t**a:**p**əl/ 'pile', /**l**ɑ**s**t/ 'burden', /**l**ɑ**s**tɪg/, /lɑs**t**ɪ**g**/ 'troublesome', or that both consonants are cluster members, /k**r**ɑ**n**t/ 'newspaper', /bɪl**j**ɑ**r**tə/ 'play at billiards'. CVC strings were extracted irrespective of position in the word, without discrimination between tautosyllabic or heterosyllabic combinations, /ɑ d**r**ɛ**s**/ 'address', /ɑt **m**i **r**al/ 'admiral'. I calculated O/E scores on the basis of the numbers of CVC combinations thus extracted from the corpus per place category, i.e. labials, PP, dorsals, KK, coronal obstruents, TT and coronal sonorants, NN. The numbers of items per category are in row three in the table.

(67) O/E scores for CVC harmony of single consonants and with cluster member(s)

CVC with and without cluster member		
	single C	with cluster
items	5,451	4,927
PP	0.5524	0.3153
KK	0.6287	0.3650
TT	0.6768	0.4048
NN	0.4207	0.4894

X^2 67.72 df 3 P <.0001

PP: X^2 66.93 df 1 p <.0001
KK: X^2 17.83 df 1 p <.0001
TT: X^2 44.38 df 1 p <.0001
NN: X^2 6.74 df 1 p 0.0094

It can be seen from the table that in all categories, except coronal sonorants, restrictions on pairs of two single consonants are smaller than restrictions on pairs of which one element is a cluster member. The high score for coronal sonorant harmony within syllable may be connected with the position of sonorants in the syllable. In onset they are often part of a cluster, /pl, bl, fl, vl, kl, xl, sl, sn/, in the rhyme they may be part of the nucleus, which may be followed by any other consonant (provided the SSH is adhered to), hence CVC combinations with sonorants more readily combine in cluster situations than as single consonant. This implies that the general picture suggests that clusters influence OCP effects. In order to gain a better insight into the exact nature of the influence of clusters, I will deal with the matter in more detail below. It was established above in § 4.5.1 that in CVC strings OCP on strings with initial C in word onset was less strict than OCP further down in the word. I will here investigate the influence of cluster membership at three different places in the word, viz. CVC with initial C in word onset, word-medial CVC and CVC with final C in word coda. I will examine harmonies with single consonants only and harmonies with at least one consonant as cluster member in monosyllabic and polysyllabic words.

Monosyllables

For the purpose of examining the situation for Dutch I divided the CELEX corpus of 8,305 words into monosyllabic and polysyllabic words. In order to be able to calculate scores I extracted CVC combinations with initial C in word onset and coded consonants in accordance with place of articulation as well as with + or - cluster membership. Intermediate vowels were filtered out. O/E scores were thus calculated on the basis of the place features combined with cluster membership. Table (71) deals with monosyllables. It must be noted that the number of consonant combinations in monosyllabic CVC words is just below 1,000. The category is nevertheless included in the analysis.

(68) O/E scores for CVC without and with cluster in monosyllables

CVC in monosyllables		
	CVC	cluster
	954	1,295
PP	0.6870	0.2655
KK	0.8548	0.1924
TT	0.8926	0.4522
NN	0.5373	0.4366

X^2 27 df 3 P <.0001

PP: X^2 24.66 df 1 p <.0001
KK: X^2 15.49 df 1 p <.0001
TT: X^2 27.27 df 1 p <.0001
NN: X^2 1.25 df 1 p 0.2636

It can be seen from table (68) that single CVC in monosyllables is only restricted in labials and in coronal sonorants; dorsals and coronal obstruents with single CVC have normal O/E scores. When one of the two consonants is a cluster member or both are cluster members all categories show severe restrictions. There is a substantial difference between single CVC and CVC with cluster membership in labials, dorsals and coronal obstruents. The scores for coronal sonorants are not significant. We may safely say that cluster membership influences OCP in monosyllables.

Polysyllables

OCP with initial C in word onset

Table (69) shows O/E rates for polysyllables at word onset.

(69) O/E scores for CVC without and with cluster member(s) at word onset in polysyllables

CVC with cluster at word onset in polysyllables		
	single CVC(V)	(C)(C)CVCC(C)(V) with cluster
items	2,200	1,187
PP	0.5651	0.4084
KK	0.6705	0.4946
TT	0.7067	0.3630
NN	0.3846	0.9909

X^2 79.43 df 3 P <.0001

PP: X^2 24 df 1 p <.0001
KK: X^2 2.89 df 1 p 0.0891
TT X^2 15.86 df 1 p <.0001
NN: X^2 35 df 1 p <.0001

It can be seen from table (69) that CVC in polysyllables is restricted in all categories except coronal sonorants with cluster member. When one of the two consonants is a cluster member or both are cluster members three out of four categories show stricter OCP than with single consonants. The difference between the O/E scores is significantly in favour of harmony with single consonants. We may conclude that cluster membership influences OCP in polysyllables in labial, dorsal and coronal [-SON] CVC with initial C in word onset.

When we consider investigations on the double markedness of OCP and cluster membership, we must take extra-syllabic /s/ into account. It is apparent from the literature that double markedness of *sC at word onset combined with OCP-[place] operates on CVC strings in English (Coetzee 2007). For Dutch the same phenomenon – no /s/ at word onset ànd OCP-[LAB] – is claimed to be connected with a prominent place of AlignLabial in the constraint ranking for Dutch (Kager and Shatzman 2007). Having established above that cluster membership influences OCP I will here investigate whether there is a difference between "regular" clusters at word onset and clusters with extra-syllabic /s/. In order to check the state of affairs for Dutch I extracted from the CELEX corpus all CC(C)VC strings at word onset and divided them into two categories, viz. CC (not sC) at word onset and sC(C) at word onset. I coded all CVC combinations in accordance with place of articulation; intermediate vowels were filtered out. Thus O/E rates are based on place and + or – extra-syllabicity only. It must be noted that the number of combinations with extra-syllabic /s/ is below 1,000.

(70) O/E scores for CC(C)VC at word onset without and with extra-syllabic /s/

CC(C)VC at word onset with and without /s/ cluster		
	regular CCVC	sC(C)VC
items	1,305	867
PP	0.5795	0.0442
KK	0	0.4895
TT	0	0.6264
NN	0.6027	0.4438

	X^2	77.29	df 3	p <.0001
PP:	X^2	9.44	df 1	p 0.0021
KK:	X^2	27.91	df 1	p <.0001
TT:	X^2	26.38	df 1	p <.0001
NN:	X^2	9.05	df 1	p 0.0026

The table shows mixed outcomes. Only labials and coronal sonorants are represented in both categories. Dorsals and coronal obstruents are only found with extra-syllabic /s/. The difference between "regular" labial clusters and labial clusters with extra-syllabic /s/ is 0.56 in favour of harmony with "regular" clusters. It is clear that labial harmony with the first segment in an sC cluster is far more restricted than labial harmony with the first segment in a regular cluster. We may conclude that multiple markedness of OCP labial and extrasyllabic /s/, as well as OCP coronal [+SON] and extrasyllabic /s/ outrank OCP and cluster membership in general. The ranking cannot be established for dorsals or for coronal obstruents.

Word-medial OCP

In order to check the situation for word-medial harmony, I have taken all word medial CVC strings and coded consonants according to place of articulation as well as to their being a single consonant or part of a cluster. Of the total number of word-medial CVC combinations – 1,853 strings – I found only 10 harmonic labial and 1 harmonic dorsal single CVC pairs, and 3 harmonic labial and 1 harmonic dorsal pairs with cluster member. As already noted above in § 4.5.1 on position in the word, OCP is very powerful word-medially; with so little harmony there is no proof of an extra effect of cluster membership.

OCP with final C in word coda

In order to check the situation for word final harmony, I have taken all word final combinations including CVC, i.e. words ending in VCVC, /aː**d**ə**l**/ 'nobility', /hoːri**z**ɔ**n**/ 'horizon', and words ending in CCVC, /kɑs**t**eː**l**/ 'castle', /kalɛn**d**ə**r**/ 'calender', CVCC, /eːlə**m**ɛ**n**t/ 'element'. CCVCC combinations at word end, /ɑk**s**ɛ**n**t/ 'accent', are also included. The number of items in the corpus with these combinations amounts to 1,548, i.e. 524 "single" CVC and 1,048 for CVC with cluster member. These numbers are too low to give reliable O/E scores. It may be noted that the total number of labial and dorsal harmonic CVC pairs at word end is 8 for labial, all with single C at word end, and 5 for dorsal, CVC only. This indicates that word final OCP is very strict and more so, when combined with word final cluster membership.

This concludes the investigation on the influence of clusters across a single segment. It is clear that there is an influence of clusters on labial, dorsal, as well as on coronal [-SON] CVC with the first element in word onset. CVC strings in word coda show that single consonant pairs are less restricted than pairs of which one segment is part of a cluster, in both labials and dorsals. From the above it is clear that double markedness of OCP-[place] with cluster markedness increases OCP restrictions. We may deduce that

Hypothesis 6 – OCP across segments is influenced by multiple markedness of an OCP constraint with cluster membership – is corroborated.

4.7 Conclusion

This concludes the investigation on OCP in Dutch based on the literature discussed in Chapter 2. Expectations based on the literature in Chapter 2 led to four hypotheses. Two additional hypotheses were included. These six hypotheses were considered and it was shown that Dutch behaved as expected in certain aspects on OCP and differently in other aspects. The results for Dutch in connection with the six hypotheses are as follows:

Hypothesis 1:

OCP is a constraint active in the Dutch language in surface adjacent consonants as well as in consonants across segments.

The hypothesis was corroborated for [place] in labial and dorsal word-initial surface adjacent consonants, as well as in consonants across segments in all positions in the word. It was corroborated for [manner] in surface adjacent consonants for the feature [continuant].

In connection with Hypothesis 1 it was established that extra-syllabicity outranks OCP-[COR] in word-initial surface adjacent consonants, and that Nasal Place Assimilation outranks OCP-[place] in word medial and word final surface adjacent consonants in all three place categories. With regard to OCP-[manner] it was found that the Sonority Sequencing Hierarchy outranks OCP in all surface adjacent consonant pairs at word begin and at word end.

Hypothesis 2:

The strictness of OCP is related to distance between two consonants, measured in segments.

The hypothesis was corroborated for [place] for OCP [labial] and [dorsal]. Hypothesis 2 was also corroborated for [manner] in the features [sonorant] and [continuant], although the feature [sonorant] had initially not been recognized as a feature category influenced by OCP in general. It was established that coronals as a general category are exempt from OCP restrictions in relation to distance.

Hypothesis 3:

Degree of similarity influences the strictness of OCP.

The hypothesis was corroborated for [place] for coronals – they are influenced by additional [sonorant] and [continuant] – as well as for [manner] for sonorants, additionally influenced by [nasal] and [lateral]. The place categories of labials and dorsals are not affected by degree of similarity, nor is the manner category of [continuant].

Hypothesis 4:

The strictness of OCP is connected with prosodic salience.

The hypothesis was corroborated for position in the word, as well as for stress for labials, dorsals and coronals. It was established that there is no influence of sonority on OCP in Dutch.

Hypothesis 5:

OCP across segments is influenced by intervening vowels.

No evidence in favour of this hypothesis was found. The hypothesis was rejected.

Hypothesis 6:

OCP across segments is influenced by multiple markedness of an OCP constraint with cluster membership.

The hypothesis was corroborated for [place] for the categories of labials, dorsals and coronal obstruents.

This completes the investigation on OCP in Dutch. The next chapter will connect the present findings for Dutch with findings from the literature, discuss differences and similarities between the two, and suggest which aspects of OCP may be considered universal and which are language-specific.

5 Results, OT analysis and Discussion

The aim of the present thesis is to establish whether OCP is a universal phenomenon and part of Universal Grammar, and which aspects influencing the gradient nature of OCP can be said to be part of Universal Grammar as well. In order to find evidence for the universal nature of OCP I investigated OCP in Dutch, based on Dutch language data. I examined both if there is OCP at all in Dutch and I analysed several possible influences on the strictness of OCP. In this chapter I review the hypotheses formulated for Dutch and summarize the results. The outcome for Dutch will be interpreted in terms of OT rankings of constraints in accordance with the data encountered. The first part of this chapter will also contain conclusions based on OCP in Dutch.

The second part of this chapter will be devoted to comparing the findings for Dutch with findings from the literature discussed in Chapter 2, and to connecting the two. Conclusions on universal aspects of OCP based on both Dutch and other languages will be dealt with.

5.1 Results and conclusions for Dutch, OT analysis

5.1.1 Results

As an introduction to the discussion, the outcome of the investigation on OCP in Dutch is represented in the matrix below. The leftmost column in the matrix covers the place features [LAB], [DOR] and [COR], the two manner features including all consonants, viz. [sonorant] and [continuant], and the feature [voice]. The analysis of Chapter 4 covered OCP in general in surface adjacent consonants and in consonants across segments, columns 2 and 3, as well as aspects influencing the gradient nature of OCP, including similarity, distance, salience, vowels and multiple markedness of OCP and cluster membership, columns 4-16. A "+" in a cell indicates evidence of (influence on) the relevant OCP category, a "-" in a cell indicates that there is no evidence of (influence of) OCP in the relevant category or that no evidence can be established on account of other constraints obscuring possible OCP effects. Impossible combinations, irrelevant areas, and invalid outcomes are also indicated by "-".

(71) OCP matrix for Dutch

	OCP surf adj	OCP acr segm	Influence on gradient OCP												
			similarity						dist	salience			vowels		M>2
			[son]	[cont]	[son] [cont]	[voi]	[nas]	[lat]		pos	stress	sonor	length	place	OCP +clus
[LAB]	+	+	-	-	-	-	-	-	+	+	+	-	-	-	+
[DOR]	+	+	-	-	-	-	-	-	+	+	+	-	-	-	+
[COR]	-	-	+	+	+	-	-	-	+	+	+	-	-	-	+
[+SON]	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-
[-SON]	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
[+CONT]	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-
[-CONT]	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-
[αVOI]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

It can be seen from table (71) that the row for [αVOI] has "-" only, indicating that OCP-[voice] constraints are not active in Dutch. Apart from this one row there is no row with "-" only, which implies that all other features in column 1 show evidence of OCP constraints in Dutch. Hypothesis 1 – OCP is a constraint active in the Dutch language in adjacent consonants as well as in consonants across segments – is visualized in column 2, "OCP in surface adjacent consonants", and column 3, "OCP in consonants across segments". There is OCP-[LAB] and OCP-[DOR], both in surface adjacent consonants and in consonants across segments; no OCP-[COR] is visible in the general picture on account of extra-syllabicity outranking the OCP constraint. OCP effects for the feature [sonorant] are obscured by the SSH; there is OCP-[continuant] in surface adjacent consonants.

The results for Hypothesis 2 – The strictness of OCP is related to distance between two consonants, measured in segments – can be seen in column 10, "distance". All OCP-restricted combinations show diminishing effects of co-occurrence restrictions with larger distance between the segments.

Hypothesis 3 – Degree of similarity influences the strictness of OCP – is represented in columns 4-9, “similarity”. It can be seen that coronals are influenced by additional markedness of [sonorant] and/or [continuant], sonorants are influenced by the features [nasal] and [lateral]. The feature [voice] does not participate in additional markedness.

Hypothesis 4 – The strictness of OCP is connected with prosodic salience – is expressed in columns 11-13, “salience”. Labials, dorsals and coronals are affected by position in the word and by word stress. No influence of sonority is present in Dutch.

With regard to Hypothesis 5 – OCP across segments is influenced by intervening vowels – we may note that both column 14 on vowel length and column 15 on vowel place have “-“ only, indicating no influence of these aspects on the gradient nature of OCP in Dutch.

Finally, the results for Hypothesis 6 – OCP across segments is influenced by multiple markedness of an OCP constraint with cluster membership – can be found in column 16, “M>2”. Multiple markedness of OCP-[LAB], OCP-[DOR] and OCP-[COR] with an additional (non-OCP) constraint prohibiting cluster membership is present in Dutch.

5.1.2 Conclusions for Dutch

OCP is a constraint type present in the Dutch language. There is OCP-[place] and OCP-[manner].

In all feature categories in which OCP is active in the language, either as a singular constraint or by means of conjoined constraints, there is an influence of distance between the segments on the strictness of OCP. Wherever there is no OCP in surface adjacent consonants in Dutch there is no OCP across segments. Wherever there is OCP in surface adjacent consonants there is an influence of distance on OCP across segments. Even when scores are within the normal range, there are gradient differences connected with distance between the segments. It is not clear that this is also the case for the feature [sonorant] since other constraints on surface adjacent consonants obscure possible OCP effects. However, with all other feature categories following the same pattern, we may deduce that, although not visible in the language, OCP constraints in surface adjacent consonants are present for the feature [sonorant] as well.

With regard to similarity it may be noted that the predictions of the similarity metric, as formulated by FBP (1997), do not apply in Dutch. There is evidence in Dutch of manner features influencing coronals in the category of place – [sonorant] and [continuant] – and of manner features influencing sonorants in the category of manner – [nasal] and [lateral], however, the features influencing OCP in Dutch are a restricted number of specific constraints conjoining with specific other constraints. Multiple markedness does not imply counting numbers of shared or non-shared features, but denoting which particular feature interacts with which particular other feature.

In the place categories prosodic salience plays a role in Dutch. OCP constraints for all three place categories are affected by position in the word and word stress; there is no influence of sonority. For all three categories the more salient harmonic strings are less prone to OCP than the less salient harmonic strings.

Vowel quality and quantity appear to be irrelevant factors on the (gradient) influence of OCP in Dutch. It is clear that OCP is present in surface adjacent consonants as well as in consonants across segments, and also that intermediate vowels are invisible to the OCP restrictions.

Multiple markedness of an OCP constraint with cluster membership influences all three place categories in Dutch.

5.1.3 OT analysis

In this section I will present an OT account of the findings on OCP in Dutch as discussed above.

5.1.3.1 OT on OCP in general[23]

Having answered the question "*is there OCP in Dutch*" affirmatively, an OT constraint ranking for Dutch must express the fact that there is virtually absolute OCP-$C_{LAB}C_{LAB}$ and OCP-$C_{DOR}C_{DOR}$ in Dutch in surface adjacent consonants, as well as less strict OCP-$C_{[\alpha CONT]}$ $C_{[\alpha CONT]}$, but also that no OCP-$C_{COR}C_{COR}$ and OCP-$C_{[\alpha SON]}C_{[\alpha SON]}$ effect is visible, and that there is no OCP-[αVOI]. This means that a general OT constraint, such as OCP-$C_{[\alpha PLACE]}C_{[\alpha PLACE]}$ – adjacent consonants with identical place features are prohibited – must be split up into its derived constraints OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$, OCP-$C_{COR}C_{COR}$, as already suggested by McCarthy (1988) and applied in OT by Coetzee and Pater (2008). Similarly, in order to be able to account for differences within the categories, the OCP constraints on [manner] will also have to be split up into their respective subcategories, for example OCP-$C_{[+SON]}C_{[+SON]}$ and OCP-$C_{[-SON]}C_{[-SON]}$. For the facts found in relation to consonant clusters in the place categories we must include constraints connected with extra-syllabic /s/ at word begin, and with nasal place assimilation for adjacent nasal + obstruent combinations. For the manner categories we must include constraints connected with the Sonority Sequencing Hierarchy, as well as with the exceptional position of the onset cluster /sx/. For the exceptional /sx/ cluster I make use of a constraint applied by Adriaans & Kager (2010) in segmentation, viz. a contiguity constraint for high probability biphones, "CONTIG-IO(*sx*) 'Sequence *sx* should be preserved*'*" (Adriaans & Kager, p 6). Extrasyllabicity can of course also be expressed by contiguity. However, I here express extrasyllabicity with the constraint IDENT-sC_{ONS} – word-initial sC clusters must be realized. The constraints required for the facts connected with the general question "*Is there OCP in Dutch?*" are:

---

[23] I here refer to OCP in surface adjacent consonants only; OCP across segments will be dealt with below in the section on OCP and distance.

(72) OCP constraints for Dutch

OCP-$C_{[\alpha PLACE]}C_{[\alpha PLACE]}$	– surface adjacent consonants with identical place features are prohibited
OCP-$C_{LAB}C_{LAB}$	– surface adjacent labial consonants are prohibited
OCP-$C_{DOR}C_{DOR}$	– surface adjacent dorsal consonants are prohibited
OCP-$C_{COR}C_{COR}$	– surface adjacent coronal consonants are prohibited
OCP-$C_{[\alpha SON]}C_{[\alpha SON]}$	– surface adjacent consonants with identical [sonorant] features are prohibited
OCP-$C_{[+SON]}C_{[+SON]}$	– surface adjacent [+SON] consonants are prohibited
OCP-$C_{[-Son]}C_{[-SON]}$	– surface adjacent [-SON] consonants are prohibited
OCP-$C_{[\alpha CONT]}C_{[\alpha CONT]}$	– surface adjacent consonants with identical [continuant] features are prohibited
OCP-$C_{[+CONT]}C_{[+CONT]}$	– surface adjacent [+CONT] consonants are prohibited
OCP-$C_{[-CONT]}C_{[-CONT]}$	– surface adjacent [-CONT] consonants are prohibited
OCP-$C_{[\alpha VOI]}C_{[\alpha VOI]}$	– surface adjacent consonants with identical [voice] features are prohibited
SSH_{ONS}	– in onset CC the SSH must be adhered to
NAS_{Ass}	– all nasals followed by an obstruent agree in place of articulation with the obstruent
VOI_{Ass}	– all surface adjacent obstruents agree in the feature [voice]
IDENT-sC_{ONS}	– initial sC clusters must be realized faithfully
CONTIG-IO(sx)	– the sequence /sx/ must be realized faithfully

For the facts answering the general question: *Is there OCP in Dutch?*, I start from a constraint ranking as below. The constraint rankings for surface adjacent consonants are in (73) OCP-[place], (74) OCP-[manner] and (75) OCP-[voice]:

(73) OCP-[place]

IDENT -sC_{ONS}, NAS_{Ass} >> OCP- $C_{LAB}C_{LAB}$, OCP- $C_{DOR}C_{DOR}$, OCP- $C_{COR}C_{COR}$ >> IDENT-[PLACE]

This constraint ranking ascertains faithful realization of word initial sC_{COR}, and nasal place assimilation word-medially and word-finally. The constraint ranking with OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$ and OCP-$C_{COR}C_{COR}$ outranking IDENT-[PLACE] implies that co-occurrence of labials, dorsals and coronals (other than sC_{COR}) is avoided in surface adjacent consonants.

The IDENT-[PLACE] constraint following OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$ and OCP-$C_{COR}C_{COR}$ is responsible for the faithful realization of all other labial, dorsal and coronal consonants. This ranking creates a categorical restriction, i.e. it predicts absolute non-occurrence of labial, dorsal and coronal consonant pairs, and free co-occurrence of non-adjacent labial, dorsal and coronal consonants. As we saw above, however, OCP restrictions are often gradient, rather than absolute. The fact that coronal consonant pairs do occur in Dutch suggests a more prominent position in the ranking for the faithfulness constraint IDENT-[PLACE]. As indicated above, there is gradience in well-formedness and there are different ways of accounting for gradient effects of OCP, such as comparative tableaux in OT and weighting of constraints in Harmonic Grammar (§ 2.1.5.2). For the present I will concentrate on OT and I will assume the differences in the strictness of OCP to be explained by comparative OT tableaux, as illustrated in tableau (21) above. This means that the position of the faithfulness constraints IDENT-[PLACE] and IDENT-[MANNER] are prominent in the ranking and that underrepresentation, rather than non-occurrence of consonant combinations, is captured by the relative ranking of the OCP constraints *below* IDENT-[PLACE] and IDENT-[MANNER]. This gives a revised version of (73) OCP-[place]:

(73, revised) OCP-[place]

> IDENT -sC_{ONS}, NAS_{ASS} >> OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$ >> IDENT-[PLACE] >> OCP-$C_{COR}C_{COR}$

This ranking is consistent with the facts found for Dutch, viz. co-occurrence of coronals due to the faithful realization of sC_{COR}, harmony in surface adjacent word-medial and word-final nasal + obstruent in all three place categories, no co-occurrence of dorsals or labials in word onset, and restricted co-occurrence of coronals (other than /sC/) in word onset. Subsequent rankings in this section will express the gradient nature of co-occurrence restrictions in a similar way, starting with OCP-[manner] in surface adjacent consonants in (74).

(74) OCP-[manner]

CONTIG-IO(sx) >> SSH_{ONS} >> IDENT-[SON], IDENT-[CONT] >> OCP-$C_{[\alpha CONT]}C_{[\alpha CONT]}$ >> OCP-$C_{[\alpha SON]}C_{[\alpha SON]}$

This constraint ranking ascertains adherence to the SSH in word onset. It allows co-occurrence of consonants sharing the feature [continuant] and of consonants sharing the feature [sonorant]. The constraint OCP-[ɑSON] being outranked by OCP-[ɑCONT] predicts that restrictions on consonant pairs sharing the feature [sonorant] are less strict, if present at all, than those sharing the feature [continuant].

(75) OCP-[VOICE]

VOI_{ASS} >> IDENT-[ɑVOI] >> $C_{[\alpha VOI]}C_{[\alpha VOI]}$

With the VOI_{ASS} constraint outranking IDENT-[ɑVOI], this ranking predicts that there are no surface adjacent consonant pairs with different specifications for the feature [voice].

The three constraint rankings suggested in (73, revised), (74) and (75) comprise the general OCP results for surface adjacent consonants in Dutch.

5.1.3.2 OT on OCP and distance

As indicated above in Chapter 4, OCP is active in adjacent consonants as well as in consonants across segments. The OCP effects were found to be gradient. The gradient nature of OCP is not only influenced by the relative strength of the OCP constraints themselves, but it is also influenced by several aspects influencing OCP. In connection with this, further modifications to the constraint rankings are required. The following rankings will deal with aspects influencing the strictness of OCP.

It was established that distance between segments plays a role in the strictness of OCP. The constraints expressing OCP-restrictions connected with distance are an elaboration on OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$, OCP-$C_{COR}C_{COR}$, etc, as introduced above. What must be expressed in the constraint ranking is that we find strictest OCP in surface adjacent segments in labials, dorsals and coronals, and decreasing OCP effects across one and two segments, as well as strictest OCP in surface adjacent [αCont] consonant pairs, and diminishing strictness of OCP-[αCONT] and OCP-[αSON] across segments. The remaining part of the ranking will be the same as in general OCP above.

I suggest the following constraints to capture the influence of distance. I here give the constraints connected with distance for labials only; they should be considered to represent similar constraints for the other feature categories.

(76) Constraints for distance

OCP-$C_{LAB}C_{LAB}$	– surface adjacent labial consonants are prohibited
OCP-$C_{LAB}V(:)C_{LAB}$	– labial consonants separated by a vowel are prohibited
OCP-$C_{LAB}VC_{LAB}$	– labial consonants separated by a short vowel are prohibited
OCP-$C_{LAB}V{:}C_{LAB}$	– labial consonants separated by a long vowel are prohibited
OCP-$C_{LAB}X_{\alpha}X_{\beta}C_{LAB}$	– labial consonants separated by two segments are prohibited
OCP-$C_{LAB}X^{>2}C_{LAB}$	– labial consonants separated by more than two segments are prohibited

(77) Constraint ranking for Place and distance:

IDENT-SC_{ONS}, NAS_{ASS} >> OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$ >> IDENT-[PLACE] >> OCP-$C_{COR}C_{COR}$ >> OCP-$C_{LAB}V(:)C_{LAB}$, OCP-$C_{DOR}V(:)C_{DOR}$ >> OCP-$C_{LAB}X_{\alpha}X_{\beta}C_{LAB}$, OCP-$C_{DOR}X_{\alpha}X_{\beta}C_{DOR}$ >> OCP-$C_{COR}V(:)C_{COR}$, OCP-$C_{COR}X_{\alpha}X_{\beta}C_{COR}$, OCP-$C_{\alpha PL}X^{>2}{}_{\alpha PL}$

The constraint ranking with OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$, OCP-$C_{COR}C_{COR}$ outranking all other OCP constraints implies strictest OCP for surface adjacent consonants. The offence of violating OCP-$C_{LAB}V(:)C_{LAB}$, OCP-$C_{DOR}V(:)C_{DOR}$ is less severe than a violation of the higher ranked constraints OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$. Violations of the second constraint in the ranking are therefore permitted more easily than violations against the topmost OCP constraint, hence violations against OCP-$C_{LAB}X_{\alpha}X_{\beta}C_{LAB}$, OCP-$C_{DOR}X_{\alpha}X_{\beta}C_{DOR}$ matter even less. The present ranking predicts that coronal OCP across segments, and restrictions on consonants across more than two segments in any place category, are least likely to occur. This is consistent with the facts found for Dutch: no coronal OCP across segments, and no restrictions on labial and dorsal harmony beyond two segments.

(78) Constraint ranking for Manner and distance:

> SSH$_{\text{ONS}}$ >> IDENT-[SON], IDENT-[CONT] >> OCP-$C_{[\alpha\text{CONT}]}C_{[\alpha\text{CONT}]}$ >> OCP-$C_{[\alpha\text{SON}]}C_{[\alpha\text{SON}]}$ >> OCP-$C_{[\alpha\text{CONT}]}V(:)C_{[\alpha\text{CONT}]}$, OCP-$C_{[\alpha\text{SON}]}V(:)C_{[\alpha\text{SON}]}$ >> OCP-$C_{[\alpha\text{CONT}]}\ X^{>1}C_{[\alpha\text{CONT}]}$, OCP-$C_{[\alpha\text{SON}]}\ X^{>1}C_{[\alpha\text{SON}]}$ >> ...

As in the general picture in (74) above, this constraint ranking ascertains adherence to the SSH in word onset, as well as faithful realizations of [αCONT] and [αSON] pairs. The hierarchical ranking of OCP-$C_{\alpha\text{CONT}}C_{\alpha\text{CONT}}$ >> OCP-$C_{\alpha\text{CONT}}V(:)C_{\alpha\text{CONT}}$ >> OCP-$C_{\alpha\text{CONT}}\ X^{>1}C_{\alpha\text{CONT}}$ predicts that there is diminishing strictness in co-occurrence rates for consonants across a single vowel and consonants across more than one segment, when they share the feature [continuant]. The same applies to consonants sharing the feature [sonorant].

5.1.3.3 OT on OCP and multiple feature markedness

Constraints connected with multiple feature markedness are composed of combinations of constraints on single feature markedness. This can be dealt with in OT by conjoined constraints. The constraint ranking must capture the facts that coronals are restricted when in the same sonorancy and/or continuancy category, that sonorants are restricted when sharing the feature [nasal] and that liquids are restricted when sharing the feature [lateral]. The restrictions within the category of sonorant need in the present thesis not be expressed by conjoined constraints. Since the feature [nasal] is restricted to the category of sonorants and here splits up the category into two categories only, viz. nasals, [+NAS], and liquids, [-NAS], and the feature [lateral] is restricted to liquids, the double markedness in sonorants is implied in the single OCP-[αNAS] and OCP-[αLAT] constraints. I here propose the following (conjoined) constraints in connection with multiple markedness. The constraints based on coronals must be taken to represent similar constraints covering other place features.

(79) (Conjoined) OCP-constraints for multiple feature markedness

OCP-$C_{Cor}C_{Cor}$&OCP-$C_{[\alpha Son]}C_{[\alpha Son]}$	– adjacent coronals from the same sonorancy class are prohibited
OCP-$C_{Cor}C_{Cor}$&OCP-$C_{[-Son]}C_{[-Son]}$	– adjacent [-Son] coronals are prohibited
OCP-$C_{Cor}C_{Cor}$&OCP-$C_{[+Son]}C_{[+Son]}$	– adjacent [+Son] coronals are prohibited
OCP-$C_{Cor}C_{Cor}$& $C_{[\alpha Cont]}C_{[\alpha Cont]}$	– adjacent coronals from the same continuancy class are prohibited
OCP-$C_{Cor}C_{Cor}$&OCP-$C_{[-Cont]}C_{[-Cont]}$	– adjacent [-Cont] coronals are prohibited
OCP-$C_{Cor}C_{Cor}$&OCP-$C_{[+Cont]}C_{[+Cont]}$	– adjacent [+Cont] coronals are prohibited
OCP-$C_{[\alpha Nas]}C_{[\alpha Nas]}$	– adjacent sonorants with the same specification for [nasal] are prohibited
OCP-$C_{[+Nas]}C_{[+Nas]}$	– adjacent nasals are prohibited
OCP-$C_{[-Nas]}C_{[-Nas]}$	– adjacent liquids[24] are prohibited
OCP-$C_{[\alpha Lat]}C_{[\alpha Lat]}$	– adjacent sonorants with the same specification for [lateral] are prohibited

The constraint rankings covering the findings on OCP in Dutch in connection with multiple feature markedness are in (80) for place and (81) for manner.

(80) Place and similarity

OCP-$C_{Lab}C_{Lab}$, OCP-$C_{Dor}C_{Dor}$ >> Ident-[Place] >> OCP-$C_{Cor}C_{Cor}$ >> OCP-$C_{Lab}V(:)C_{Lab}$, OCP-$C_{Dor}V(:)C_{Dor}$, OCP-$C_{Cor}V(:)C_{Cor}$&OCP-$C_{[\alpha Son]}V(:)C_{[\alpha Son]}$ >> OCP-$C_{Lab}X_{\alpha}X_{\beta}C_{Lab}$, OCP-$C_{Dor}X_{\alpha}X_{\beta}C_{Dor}$, OCP-$C_{Cor}X_{\alpha}X_{\beta}C_{Cor}$&OCP-$C_{[\alpha Son]}X_{\alpha}X_{\beta}C_{[\alpha Son]}$ >> OCP-$C_{Cor}V(:)C_{Cor}$, OCP-$C_{Cor}X_{\alpha}X_{\beta}C_{Cor}$, OCP-$C_{\alpha PL}X^{>2}{}_{\alpha PL}$ ...

This constraint ranking limits co-occurrence of labial and dorsal pairs across a single segment, as well as co-occurrence of homorganic coronal pairs with the same specification for the feature [sonorant]. The low position in the ranking of singleton OCP-

[24] As indicated above in § 3.1.2.2, glides are not included in the manner categories, so [-Nas] is here restricted to liquids.

$C_{COR}V(:)C_{COR}$ predicts that other coronals than those sharing the feature [sonorant] will co-occur freely.

(81) Manner and similarity

> SSH$_{ONS}$, OCP-[αNAS], OCP-[αLAT] >> IDENT-[SON], IDENT-[CONT] >> OCP-[αCONT] >> OCP- [αSON] >> OCP-$C_{[αNAS]}V(:)C_{[αNAS]}$, OCP-$C_{[αNAS]}X_αX_βC_{[αNAS]}$, OCP-$C_{[αLAT]}V(:)C_{[αLAT]}$ >> OCP-$C_{[αCONT]}V(:)C_{[αCONT]}$, OCP-$C_{[αSON]}V(:)C_{[αSON]}$, OCP-$C_{[αLAT]}X_αX_βC_{[αLAT]}$ >> OCP-$C_{[αNAS]}X^{>2}C_{[αNAS]}$, OCP-$C_{[αLAT]}X^{>2}C_{[αLAT]}$, OCP-$C_{[αCONT]}X^{>1}C_{[αCONT]}$, OCP-$C_{[αSON]}X^{>1}C_{[αSON]}$ >> ...

As in the general picture in (74) above, this constraint ranking ascertains adherence to the SSH in word onset, categorical avoidance of surface adjacent consonants sharing specifications for the feature [nasal] and surface adjacent consonants sharing the feature [lateral], as well as restrictions on the co-occurrence of surface adjacent consonants sharing the feature [continuant]. The ranking indicates that nasals are equally restricted across one segment and across two segments, and that the OCP effect on liquids disappears across more than one segment.

5.1.3.4 OT on OCP and salience

For Dutch it appeared that two elements of prosodic salience, viz, position in the word and word stress influence OCP. What must be expressed in a constraint ranking are the influence of position in the word and word stress on OCP-[LAB], OCP-[DOR] and OCP-[COR]. With no harmony across segments for the manner features in general, manner features are left out. I take constraints on labials as examples for all three place categories. I again propose a conjoining of OCP-constraints with other constraints. The constraints connected with salience are:

(82) Constraints on salience

AlignLeftLab[25]	–	all labials must be left-aligned; this implies that labial consonants not in initial position in the word receive one violation mark for each segment position removed from the initial position.
OCP-$C_{LAB}V(:)^{+stress}C_{LAB}$	–	two labials consonants separated by a stressed vowel are prohibited
OCP-$C_{LAB}V(:)^{-stress}C_{LAB}$	–	two labials consonants separated by an unstressed vowel are prohibited

The constraint rankings covering the findings on OCP in Dutch in connection with salience are in (83) for position in the word and in (84) for word stress.

(83) Constraint ranking covering position in the word:

OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$ >> Ident-[Place] >> OCP-$C_{COR}C_{COR}$&OCP-$C_{[\alpha SON]}C_{[\alpha SON]}$ >> AlignLeftLab&OCP-$C_{LAB}V(:)C_{LAB}$, AlignLeftDor&OCP-$C_{DOR}V(:)C_{DOR}$ >> OCP-$C_{COR}V(:)C_{COR}$&OCP-$C_{[\alpha SON]}V(:)C_{[\alpha SON]}$ >> OCP-$C_{LAB}X_{\alpha}X_{\beta}C_{LAB}$, OCP-$C_{DOR}X_{\alpha}X_{\beta}C_{DOR}$, OCP-$C_{COR}X_{\alpha}X_{\beta}C_{COR}$&OCP-$C_{[\alpha SON]}X_{\alpha}X_{\beta}C_{[\alpha SON]}$ >> OCP-$C_{COR}V(:)C_{COR}$, OCP-$C_{COR}X_{\alpha}X_{\beta}C_{COR}$, OCP-$C_{\alpha PL}X^{>2}{}_{\alpha PL}$ ...

This constraint ranking covers the difference in the strictness of OCP between word-initial harmony and harmony further down in the word. The more violation marks there are in the category of AlignLeft, the stricter OCP is applied. More leniency word-initially than further down in the word is brought out by this ranking.

[25] The AlignLeft constraint applied here is used for a different purpose than the AlignLabial constraint introduced by Kager and Shatzman (2007), to account for the difference between the strictness of OCP between OCP immediately at word begin and OCP after initial extra-syllabic /s/.

(84) Constraint ranking covering stress, (combined with general OCP and distance of (92)):

> $\text{OCP-C}_{\text{LAB}}\text{C}_{\text{LAB}}$, $\text{OCP-C}_{\text{DOR}}\text{C}_{\text{DOR}}$ >> IDENT-[PLACE] >> $\text{OCP-C}_{\text{COR}}\text{C}_{\text{COR}}$ >>
> $\text{OCP-C}_{\text{LAB}}\text{V(:)}^{\text{-stress}}\text{C}_{\text{LAB}}$, $\text{OCP-C}_{\text{DOR}}\text{V(:)}^{\text{-stress}}\text{C}_{\text{DOR}}$, $\text{OCP-C}_{\text{COR}[\alpha\text{SON}]}\text{V(:)}^{\text{-stress}}\text{C}_{\text{COR}[\alpha\text{SON}]}$
> >> $\text{OCP-C}_{\text{LAB}}\text{V(:)}^{\text{+stress}}\text{C}_{\text{LAB}}$, $\text{OCP-C}_{\text{DOR}}\text{V(:)}^{\text{+stress}}\text{C}_{\text{DOR}}$,
> $\text{OCP-C}_{\text{COR}[\alpha\text{SON}]}\text{V(:)}^{\text{+stress}}\text{C}_{\text{COR}[\alpha\text{SON}]}$ >> $\text{OCP-C}_{\text{LAB}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{-stress}}\text{C}_{\text{LAB}}$,
> $\text{OCP-C}_{\text{DOR}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{-stress}}\text{C}_{\text{DOR}}$, $\text{OCP-C}_{\text{COR}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{-stress}}\text{C}_{\text{COR}}$ >> $\text{OCP-C}_{\text{LAB}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{+stress}}\text{C}_{\text{LAB}}$,
> $\text{OCP-C}_{\text{DOR}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{+stress}}\text{C}_{\text{DOR}}$, $\text{OCP-C}_{\text{COR}}\text{X}_{\alpha}\text{X}_{\beta}^{\text{+stress}}\text{C}_{\text{COR}}$ >> $\text{OCP-C}_{\alpha\text{PL}}\text{X}^{>2}\text{C}_{\alpha\text{PL}}$ >> ...

With the OCP restrictions across "-stress" segments outranking the OCP restrictions across "+stress" segments, the constraint ranking ensures stricter OCP across segments for labials, dorsals and [αSon] coronals in unstressed syllables than in stressed syllables.

5.1.3.5 OT on OCP and cluster membership

Double markedness again requires conjoining of constraints. For Dutch there must be a ranking that is consistent with severe restrictions on $\text{OCP-C}_{\text{Lab}}\text{V(:)C}_{\text{Lab}}$ and $\text{OCP-C}_{\text{Dor}}\text{V(:)C}_{\text{Dor}}$ in initial syllables when one of the two consonants is part of a cluster. Severest restrictions are found word-initially with extra-syllabic /s/. The constraints, both for cluster membership in general and for extra-syllabic /s/ in particular, are:

(86) cluster membership and OCP

*CC	– adjacent consonants are prohibited
$\text{*CC\&OCP-C}_{\text{LAB}}\text{V(:)C}_{\text{LAB}}$	– labial consonants across a segment are prohibited when one consonant is part of a cluster
*sC_{ONS}	– no sC at word begin
$\text{*sC}_{\text{ONS}}\text{\&OCP-C}_{\text{LAB}}\text{V(:)C}_{\text{LAB}}$	– labial consonants across a segment are prohibited when one is part of a word-initial /s/-cluster

A constraint ranking for OCP restrictions connected with cluster membership is in (87) below:

(87) OCP and cluster membership

> OCP-$C_{LAB}C_{LAB}$, OCP-$C_{DOR}C_{DOR}$, CONTIG-IO(sx), IDENT-SC_{ONS} >> IDENT-[PLACE] >> *sC_{ONS} &OCP-$C_{LAB}V(:)C_{LAB}$ >>, *sC_{ONS} &OCP-$C_{DOR}V(:)C_{DOR}$ >> *CC&OCP-$C_{LAB}V(:)C_{LAB}$, *CC&OCP-$C_{DOR}V(:)C_{DOR}$, *CC&OCP-$C_{COR[\alpha SON]}V(:)C_{COR[\alpha SON]}$ >> OCP-$C_{LAB}X_{\alpha}X_{\beta}C_{LAB}$, OCP-$C_{DOR}X_{\alpha}X_{\beta}C_{DOR}$, OCP-$C_{COR[\alpha SON]}X_{\alpha}X_{\beta}C_{COR[\alpha SON]}$ >> OCP-$C_{\alpha PL}X^{>2}C_{\alpha PL}$ >> ... >> *sC_{ONS}

The prominent position of the conjoined constraint *sC_{ONS} &OCP-$C_{LAB}V(:)C_{LAB}$ ensures that labial harmony across a single segment, with the initial consonant as part of a word-initial cluster with extra-syllabic /s/, is avoided. The avoidance of labial harmony across a single segment, with the initial consonant not being part of a word-initial cluster with extra-syllabic /s/, is less strict. The faithfulness constraints CONTIG-IO(sx) and IDENT-sC_{ONS}, obscure any possible effect of a difference between the conjoining of dorsal or coronal OCP With *sC_{ONS} or *CC_{ONS}. The single constraint *sC_{ONS} has a position way down in the ranking, so that no effect of this single constraint is visible and sC in onset is freely allowed in Dutch.

5.2 OCP and the languages of the world

This section will compare Dutch with other languages and determine common ground for universal and language-specific aspects of OCP.

5.2.1 Dutch and other languages

It was established in the literature that OCP is a phenomenon found cross-linguistically. There is evidence of OCP-[place] in diverse languages, such as Classical Arabic, (McCarthey 1986, 1988), Javanese, (Yip 1989), English, (Berkley 1994, 2000), Muna, (Coetzee and Pater 2008). It was also established for Dutch that there are co-occurrence restrictions for labial consonants across segments (Kager & Shatzman 2007). The present thesis contains information on OCP in Dutch. The language-specific version of OCP in Dutch includes OCP-[LAB] and OCP-[DOR], as there is in Arabic and English, both in adjacent consonants and in consonants across segments. There is no OCP-[COR] as a singular constraint in Arabic and English, nor is there in Dutch across segments. Latin, however, is claimed to show evidence of a singular OCP-[COR] constraint (Berkley 2000). With regard to Manner it must be noted that there is evidence of OCP-[manner] in English and Latin for the feature [lateral] (Berkley 2000, Van der Torre 2003), as there is in Dutch. There are no indications in the literature consulted of OCP-restrictions in the features [sonorant] and [continuant]. It may be the case that no OCP-[αSON] or OCP-[αCONT] have been found for any of the languages discussed, or that no separate investigations have been conducted. In the literature the [manner] features are only present as secondary markers of OCP restrictions. In Dutch, OCP-[manner] is present in surface adjacent [αCONT] and [αNAS] pairs, as well as in [αNAS] combinations across consonants.

OCP-[voice] was found to exist in Japanese (Yip 1988); there is no OCP-[voice] in Dutch. This general picture indicates that the phenomenon of OCP is not a regular phenomenon for all features, neither in Dutch, nor in the other languages referred to, but that it ìs a phenomenon present in different languages.

In the literature the gradient nature of OCP was found to be connected with several aspects, which will be discussed here. The influence of distance was claimed to be a defining factor on the gradience of OCP in Classical Arabic, Hebrew, Jordanian (Frisch 1997) and English (Berkley 1994). Distance is also one of the aspects found to influence gradient OCP in Dutch. Co-occurrence restrictions on surface adjacent consonants are stricter than co-occurrence restrictions on consonants separated by one or more segments. The strictness of OCP diminishes with distance between the consonants, measured in segments. We find that, wherever there is no OCP in surface adjacent consonants in Dutch there is no OCP across segments. Wherever there is OCP in surface adjacent consonants there is an influence of distance on OCP across segments. It must be noted that OCP constraints can be obscured by other constraints so that it will appear as if no OCP constraints are present for surface adjacent consonants, whereas there are for consonants across segments. Evidence of this can be found in Dutch. Although scores are within the normal range for e.g. the features [sonorant] and [continuant] across segments, there is still a gradient difference between harmony across one segment and harmony across two segments in both the features [sonorant], with obscured OCP in surface adjacent consonants, and [continuant], with OCP in surface adjacent consonants.

The issue of the influence of additional feature markedness was brought up by McCarthy (1988) when he divided the category of coronals into coronal obstruents and coronal sonorants. Frisch, Broe and Pierrehumbert (1997) devised the similarity metric to account for gradient differences in Arabic. The metric was later applied to English by Berkley (Berkley 2000). As noted above, in sections 4.4 on similarity, the similarity metric as formulated by FBP (1997) cannot be applied in Dutch. Multiple markedness in Dutch is not a matter of counting numbers of shared or non-shared features, but of denoting which particular feature interacts with which particular other feature. This is consistent with the investigation on the difference in the influence of secondary features on OCP in Arabic and Muna by Coetzee and Pater (2008). Arabic shows equal influence of the features [sonorant], [continuant] and [voice], whereas in Muna there is influence of [sonorant] and [continuant], not of [voice]. These findings suggest that the similarity metric should be abandoned in its current form. It may be the case that the metric has predictive value for Arabic (FBP) and English (Berkley 2000), for which claims have been made, because these languages happen to be influenced to the same extent by multiple features. However, the fact that multiple feature markedness applies in Dutch, but is restricted to coronals and sonorants, as well as findings for the difference between Muna and Arabic by Coetzee and Pater suggest a language-specific conjoining of feature markedness constraints rather than the fact being dependent on the number of shared constraints. Multiple markedness in OT is the language-specific conjoining of specific OCP constraints with specific other OCP-constraints.

For Dutch, as for Arabic and English, the effect of similarity is most readily recognized in the category of coronals. There are limited or no co-occurrence restrictions when considering the entire category of coronals, there are limitations within the categories of coronal obstruents and coronal sonorants in all three languages. This was claimed to be connected with size of category. The similarity matrix produces smaller sized categories, the greater the similarity the smaller the number of segments involved. This suggests that small enough categories, such as labials and dorsals in Dutch, English and Arabic – and coronals in Latin – do not need additional restrictions for limitations on co-occurrence within the category to work. It also implies that a language with a considerable number of e.g. labials and more limited numbers of dorsals and coronals, will split up the category of labials into subcategories, which then adhere to OCP restrictions. By splitting up a category by means of additional feature restrictions, smaller sized categories result. On the other hand, Muna has a small enough coronal category and still shows signs of restrictions within the class of coronal sonorants and within the class of coronal obstruents, but not among these two classes (Coetzee & Pater, p 329). If similarity would be connected with size of category this would imply that a language with more or less equally sized categories, such as labials and coronals in Muna[26], would show similar restrictions connected with multiple feature markedness for both place categories. There is no evidence of this in the languages investigated. Since in most languages investigated coronals are the only place category affected by OCP restrictions when split up into two categories, not when considered as a whole, this would suggest that it is not the *size* of a category which determines the strictness of OCP, but that it must be connected with the *nature* of the category.

What is it then that makes coronals exceptional?

When considering the Dutch language we find that the distribution of coronals in general is freer in Dutch than the distribution of dorsals and labials. As indicated above in 3.2, Dutch allows extra-syllabic /s/ at word begin and a coronal appendix at word end. Moreover, coronals are used for inflections of all types, e.g. plural /s/ in /ɑpəls/ 'apples',

[26] Muna has 8 labials and 10 coronals (Coetzee & Pater, p 329)

plural /n/ in /pe:rən/ 'pears', in comparison /gro:tst/ 'largest', verb inflection 2nd and 3rd person singular present tense /t/, /wo:nt/ 'lives, plural /n/ /wo:nən/ 'live', present participle /wo:nənt/ 'living', past tense /d, t/ /wo:ndə/ 'lived', /wɑstə/ 'washed', past participle /t/ /gəwo:nt/ 'lived'. This makes coronals the unmarked place category in Dutch. It is evident from the literature, (Paradis and Prunet 1991, Lombardi 2003), that coronals are the least marked place category universally. This might explain the relative freedom of co-occurrence of coronals within the entire category as opposed to stricter co-occurrence avoidance in the marked categories of dorsals and labials.

I assume that OCP is active in marked categories. Conjunction of the features [coronal] and [ɑSON] creates two marked categories, viz. coronal obstruents and coronal sonorants. Multiple markedness is not a means to reduce category *size*, but it results in two marked categories. As suggested by the data from Muna there are restrictions within the universally marked categories of labials and dorsals, but not within the unmarked category of coronals, even though category sizes are practically equal; there are also restrictions within the marked categories of coronal obstruents and coronal sonorants. Conjoining of OCP constraints creates (more) marked categories within which OCP constraints are active. This implies that the similarity metric as devised by FBP, should be abandoned; the aim of multiple OCP markedness is not reduction of category size; conjunction of OCP-constraints creates marked categories.

Multiple markedness other than multiple feature markedness was not discussed in the literature. The only aspect possibly connected with the issue of OCP and cluster membership is the suggestion by Berkley that OCP is syllable-based, rather than word based. Berkley did not find evidence for this in English. I did not find evidence for syllable based OCP in Dutch either. In the present thesis OCP in Dutch is connected with multiple markedness of OCP and cluster membership. In unmarked situations co-occurrence of consonants is less restricted, in marked situations OCP is more restricted. Universally a CV syllable is the unmarked syllable, hence restrictions on consonant co-occurrence in e.g. CVCV combinations are more lenient than restrictions on combinations marked by cluster membership. In monosyllabic words co-occurrence of homorganic consonants in CVC words are far less restricted than homorganic CVC combinations of which one of the two consonants is part of a cluster. Single consonants are less marked than consonant clusters, hence less prone to OCP. As above, in the place category of coronals, marked by sonorancy and continuancy, more markedness here also implies more severe OCP restrictions.

Prosody is an aspect that was examined for English by Berkley (2000). When comparing Dutch and English we find that, whereas in English the more salient position in the word, word-initial syllable, and the more salient syllable in the word, the stressed syllable, apparently show strictest OCP, Dutch displays the reverse: more salient position in the word and more salient syllable on account of stress, are more permissive of harmony.

It must be noted that there is a difference in the way in which O/E scores were established between Berkley's approach for English and the present analysis for Dutch. Berkley examined *di*–syllabic words. With O/E scores based on disyllabic words only, there is no comparison between word-initial, word-medial and word-final harmony. Since, in disyllabic words, harmony across syllables is by nature restricted to harmony across initial syllable, and word-final harmony is equal to harmony in the second syllable, rather than necessarily at the end of the word, no gradient difference can be established with the English disyllabic subdatabase for initial, medial and final harmony. In order to avoid these obscuring restrictions, the present thesis on OCP in Dutch analyses all *poly*syllabic words, i.e. all words containing 2 to 4 syllables.
It must also be noted that Berkley's interpretations are based on average results, in which differences between categories do not show. The average difference is determined by the average O/E scores. An example will illustrate this (Berkley 2000, p 190).

(88) O/E ratios across initial syllable and within final syllable in English disyllables

Table 7.4. O/e ratios for consonant pairs separated by exactly one segment, a short vowel, with C1 in the word onset and C2 in the next syllable's onset, in English disyllables

		C_2				examples:
		labial	cor.obs.	cor.son.	dorsog.	*puppet* pʌ.pət]
	labial	0.49	1.17	1.25	0.71	*tennis* [tɛ.nIs]
	cor.obs.	1.32	0.74	0.95	1.46	
C_1	cor.son.	1.38	1.11	0.53	1.29	
	dorsog.	1.09	0.88	1.18	0.57	

N=1071; Chi square = 88.884, p <.001)

Table 7.5. O/e ratios for consonant pairs separated by exactly one segment, a short vowel, with C1 in the second syllable's onset and C2 in the second syllable's coda, in English disyllables

		C_2				examples:
		labial	cor.obs.	cor.son.	dorsog.	*puppet* [pʌ.pət]
	labial	0.36	1.08	1.16	0.79	*tennis* [tɛ.nIs]
	cor.obs.	1.5	0.75	1.06	1.08	
C_1	cor.son.	0.96	1.10	0.85	1.26	
	dorsog.	0.68	1.42	0.89	0.32	

N=1702; Chi square = 90.466, p <.001)

Berkley's analysis of the above is: "the o/e ratios for second-syllable onset-coda pairs (table 7.5) are closer to 1 than those for onset-onset pairs by an average of .05" (Berkley 2000, p 191). She does not differentiate between categories, but averages scores. Her conclusion based on the tables above is that harmony is permitted more easily in word final position in English, whereas I would say it is clear that labials and dorsals are *less* restricted in word final position in English disyllables by at least 0.13, and that coronal obstruents are affected equally in both positions, the exception here is found in coronal sonorants. With such different approaches to the same figures, it is not really possible to compare Dutch and English OCP effects on salience.

I suggest that an explanation for more leniency in Dutch in salient positions is again connected with markedness. The least marked position in the word is the initial position. As indicated above on position in the word, this coincides with other findings from the literature. The outcome with regard to stress is comparable. Cross-linguistically stress languages share a number of properties. With respect to OCP the relevant property is the *"culminative property*: It is typical for stress languages that morphological or syntactic constituents (stem, word, phrase, etc.) have a single prosodic peak" (Kager p 143). This implies that all words have a prosodic peak, i.e. all words have a stressed syllable, but not all words have unstressed syllables as well (cf. monomorphemic words), entailing that the prosodic peak in a word is unmarked. This explains why the consonant combinations in unmarked positions, i.e. word-initial position and stressed syllables, are more permissive of harmony than consonant combinations in marked positions, i.e. further down in the word and in unstressed syllables.

Berkley's analysis of OCP in connection with salience sometimes fits in better with the present analysis than with her approach. "However, if stressed syllables are more prosodically salient than unstressed syllables, and thus more subject to OCP, it is unexpected to find that unstressed word onsets exhibit a stronger OCP effect than stressed word onsets" (Berkley 2000, p 209). Berkley's findings are here contrary to her expectations, but coincide with the analysis of OCP in connection with markedness.
The analysis for Dutch reveals that OCP is indeed connected with salience, i.e. the most salient position in the word and the most salient syllable – in the present analysis the least marked position and the least marked syllable – are more permissive of harmony than the less salient position and syllable. I suggest the findings for English be reconsidered in accordance with this view on salience influencing OCP.

5.2.2 Final conclusion

It may be noted that Dutch fits in perfectly with other languages with respect to a part, not all, of the phenomena under investigation. This indicates that the phenomenon of OCP is not a general phenomenon for all features in Dutch, as it is not for the other languages referred to, but that it is a phenomenon covering different areas in different languages.
The common denominator is the phenomenon itself. The fact that diverse languages manifest the same co-occurrence restrictions suggests a common basis in UG. Although there will be languages without OCP restrictions this does not mean that the constraints are not present in the innate universal constraint set. It only means that other constraints will outrank OCP constraints and that any possible OCP-effects will be obscured. Moreover, proof of the psychological reality of the phenomenon in both concatenative and non-concatenative languages suggests that there is a basis in which this psychological reality is built. The present analysis of OCP in Dutch provides additional evidence in favour of UG containing a family of OCP constraints, which is basically present in all languages of the world.

The aspect of distance is a phenomenon connected with OCP and shared by all languages investigated. This implies that distance between segments is inherent to OCP. Wherever there is OCP in surface adjacent consonants in a language, there may or may not be OCP across segments. When a language has no OCP restrictions in adjacent segments, there will be no OCP across segments. I therefore assume that distance is a universal aspect of OCP.

The aspect of similarity was extensively dealt with above. Similarity, or multiple markedness does not depend on the number of features shared by two consonants, but it is the language-specific conjoining of one specific OCP-constraint with another specific OCP-constraint. The effect is the formation of marked forms out of previously unmarked or less marked categories. Marked forms are always language-specific choices in favour of faithfulness constraints outranking markedness constraints. Combinations of marked forms, such as OCP-[COR], with other marked forms, such as OCP-[αSON], show multiple violations. More violations (or more severe violations) increase ill-formedness, thus restricting forms with more markedness more easily than less marked forms. Even though single marked forms will be part of the language, and violations of the relevant markedness constraints are accepted on account of other constraints (such as IDENT-PLACE) outranking the violated constraints, forms incurring more than one violation or violating more than one constraint are more marked and therefore less desirable, they will therefore be more easily defined as less harmonic. A single constraint can conjoin with virtually any other single constraint, so an OCP-constraint can conjoin with another OCP-constraint, or it can conjoin with some other markedness constraint, resulting in a marked situation, where an unmarked (or less marked) situation is at hand. This is what we find in OCP and cluster membership in Dutch in which an OCP-constraint conjoins with a non-OCP-constraint. The fact that two constraints are conjoined is not connected with general OCP restrictions. The restrictions on OCP-Cor[αSon] and on OCP and cluster membership are language-specific realizations of the conjoining of two constraints, resulting in (more) active OCP in a marked category. Conjunction of constraints is part of the make up of universal grammar. I therefore consider multiple markedness an aspect of OT, rather than it being inherently connected with OCP.

The influence of prosodic salience is present in English and in Dutch. I did not find any other indication in the literature discussed of evidence either in favour of the influence of position in the word and/or word stress on OCP or against such influence. Position in the word, as well as word stress, are again connected with markedness. As with multiple (OCP) markedness, additional markedness with regard to position in the word and word stress results in more marked categories. This is not inherently connected with OCP.

Summing up we may say that OCP is a universal phenomenon restricting the co-occurrence of similar consonants in each other's vicinity. The strictness of OCP is always connected with distance between harmonic segments.

In most languages OCP is a gradient phenomenon, with strictest OCP in marked situations and less strict OCP effects in unmarked or less marked situations. On the one hand, there are facets of OT inherently connected with markedness, such as labial and dorsal place features, on the other hand, there are influences resulting in markedness in universally unmarked situations, such as multiple OCP markedness, conjunction of an OCP-constraint with cluster membership, position in the word and word stress.

OCP is a universal phenomenon, universally influenced by distance between two segments; gradience in OCP is connected with the language-specifically defined playing field for OCP. OT provides the means to denote markedness, thus allowing languages to define their playing fields.

References:

Adriaans, Frans W. & René W.J. Kager (2010). Adding Generalization to Statistical Learning: The Induction of Phonotactics from Continuous Speech. In: *Journal of Memory and Language* **62**: 311-331

Adriaans, Frans W. (2006). *PhonotacTools – manual*. Utrecht Institute of Linguistics OTS, Utrecht University

Alderete, John D. (1997). Dissimilation as Local Conjunction. *Proceedings of the North East Linguistics Society* **27**:17-32

Alderete, John D. & Stefan A. Frisch (2006). Dissimilation in Grammar and the Lexicon. In: Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge, Cambridge University Press

Baayen, Harald, Richard Piepenbrock & Leon Gulikers (1995). *The CELEX Lexical Database.* Philadelphia: Linguistics Data Consortium, University of Pennsylvania.

Berent, Iris & Joseph Shimron (1997). The Representation of Hebrew Words: Evidence from the Obligatory Contour Principle. *Cognition* **64**: 39-72

Berkley, Deborah M. (1994a). Variability in Obligatory Contour Principle Effects. In: *Proceedings of the 30th Annual Meeting of the Chicago Linguistic Society, part two: The Parasession on Variation and Linguistic Theory,* p 1-12

Berkley, Deborah M. (1994b). The OCP and Gradient Data. In: *Studies in the Linguistic Sciences* **24**: 59-72

Berkley, Deborah M. (2000). *Gradient Obligatory Contour Principle Effects*. Ph.D. Dissertation, Northwestern University

Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English.* Harper and Row, New York

Clements, George N. (1990). The Role of the Sonority Cycle in Core Syllabification. In: Kingston, John & Mary E. Beckman (eds.) *Papers in Laboratory Phonology 1.* Cambridge University Press

Coetzee, Andries (2005). The OCP in the Perception of English. In S. Frota, M. Vigario, M.J. Freitas (eds.) *Prosodies.* New York: Mouton de Gruyter, p 223-245

Coetzee, Andries (2009). Grammar is both Categorical and Gradient. In S. Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation.* London: Equinox Publishers, p 9-42. ROA-864 [Pre-print version]

Coetzee, Andries W. (2007). Grammaticality and Ungrammaticality in Phonology. In *Project Muse,* p 218-257

Coetzee, Andries W. & Joe Pater (2008). Weighted Constraints and Gradient Restrictions on Place Co-occurrence in Muna and Arabic. *Natural Linguistic Theory* **26:** 289-337

Elmedlaoui, Mohamed (1995). Géométrie des Restrictions de Cooccurrence de Traits en Sémitique et en Berbère: Synchronie et Diachronie. *Canadian Journal of Linguistics* **40**: 39-76

Fikkert, Paula & Clara C. Levelt (2005). 'How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar'. In: P. Avery, B. Elan Dresher & K. Rice (eds.) *Contrast in phonology: Perception and Acquisition*. Berlin: Mouton

Frisch Stefan A. (1996). *Similarity and Frequency in Phonology*, Ph.D. Dissertation, Evanston, Illinois

Frisch, Stefan A., Michael B. Broe & Janet B. Pierrehumbert (1997). Similarity and Phonotactics in Arabic, unpublished manuscript, Indiana University, ROA-223

Frisch, Stefan A. & Bushra Zawaydeh (2001). The Psychological Reality of OCP-Place in Arabic. *Language* **77**: 91-106

Frisch, Stefan A., Janet B. Pierrehumbert & Michael B Broe (2004). Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory* **22:** 179-228

Goldsmith, John A. (1976). *Autosegmental Phonology,* Ph.D. Dissertation*,* Indiana University, Bloomington, Indiana

Hulst, Harry G. van der (1984). *Syllable Structure and Stress in Dutch*, Foris Publications, Dordrecht

Ito, Junko & Armin Mester (1998) Markedness and Word Structure: OCP Effects in Japanese. Santa Cruz: University of California, Santa Cruz, ROA-255

Kager, René W.J. & Wim Zonneveld (1986). Schwa, Syllables and Extrametricallity in Dutch. *The Linguistic Review* ***5**,* 1985-1986, p 197-221

Kager, René (1999) *Optimality Theory,* Cambridge University Press

Kager, René & Keren Shatzman (2007). Phonological Constraints in Speech Processing. In: Los, B. & M. van Koppen (eds.) *Linguistics in the Netherlands 2007*, 100-111. Amsterdam: John Benjamins

Kawahara, Shigeto, H. Ono & K. Sudo (2005). Consonant Cooccurrence Restrictions in Yamato Japanese. In Timothy Vance (ed.) *Japanese/Korean Linguistics* **14**: 27-38. Stanford: CSLI Publications.

Kenstowicz, Michael J. (1994). *Phonology in Generative Grammar.* Wiley-Blackwell

Legendre, Géraldine, Y. Miyata & Paul Smolensky (1990). Harmonic Grammar: a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In: *Proceedings of the twelfth annual conference of the cognitive science society,* p 388-395. Cambridge MA: Lawrence Erlbaum

Leben, W.R. (1973). *Suprasegmental Phonology*. Ph.D. Dissertation, MIT, Cambridge, Massachusetts. Distributed by Indiana University Linguistics Club, Bloomington. Published 1980, Garland Publishing, Inc., New York and London

Lombardi, Linda (2002). Coronal epenthesis and markedness. In: *Phonology* **19**: 219-251.

Loon, Jozef van (1986). *Historische fonologie van het Nederlands.* Acco, Leuven

McCarthy, John J. (1986). OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* **17**: 207-263

McCarthy, John J. (1988). Feature Geometry and Dependency: A Review. *Phonetica* **43**:84-108

Ohala, John J. & Haruko Kawasaki (1984). Prosodic phonology and phonetics. In Ewen C. & E. Kaisse (eds.) *Phonology Yearbook* **1**, p 113-127. Cambridge University Press.

Padgett, Jaye (1995). *Stricture in Feature Geometry,* CSLI Publications, Stanford, CA

Paradis, Caroline & Jean-François Prunet (1991). *The Special Status of Coronals. Internal and External Evidence.* San Diego: Academic Press.

Pater, Joe & Adam Werle (2001). *Typology and Variation in Child Consonant Harmony.* University of Massachusetts, Amherst. http:/www.umass.edu/linguist/faculty/pater_werle.pdf

Pierrehumbert, Janet (1992) Dissimilarity in the Arabic Verbal Roots, *Proceedings of the 23rd Meeting of the Northeastern Linguistic Society,* Graduate Student Association, University of Massachusetts, Amherst. 367-381

Pierrrehumbert, Janet (1993). Dissimilarity in the Arabic Verbal Roots. In: A. Schafer (ed.) *Proceedings of the North East Linguistics Society* **23**:367-381, Amherst GLSA

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar.* Rutgers University Center for Cognitive Science Technical Report **2**, New Brunswick

Selkirk, Elisabeth O. (1984). On the Major Class Features and Syllable Theory. In: Aronoff, Mark & Richard T. Oehrle (eds.) *Language Sound Structure.* The MIT Press, Cambridge Massachusetts

Shatzman, Keren B. & René Kager (2007). A Role for Phonotactic Constraints in Speech Perception. In J. Trouvain & J.J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences,* p 1409-1412. Dudweiler, Germany: Pirrot.

Tesar, Bruce & Alan Prince (2003/2007). Using Phonotactics to Learn Phonological Alternations. In: *The Proceedings of CLS 39 (2003), Vol. II: The Panels*, 209-237. ROA-620

Torre, Erik Jan van der (2003). *Dutch Sonorants: the Role of Place of Articulation in Phonotactics*. Ph.D. Dissertation, Leiden University

Trommelen, Mieke T.G & Wim Zonneveld (1979). *Inleiding in de Generatieve Fonologie.* Dick Coutinho, Muiderberg

Trommelen, Mieke T.G. (1983). *The Syllable in Dutch.* Foris Publications, Dordrecht

Trommelen, Mieke T.G. (1988). De structuur van lettergrepen. In: Broecke, Marcel P.R. van den (ed) *Ter Sprake: Spraak als betekenisvol geluid in 36 thematische hoofdstukken.* Foris Publications, Dordrecht

Weijer, J. van de (2003). Consonant Variation within Words. In: Archer, D., P. Rayson, A. Wilson & T. McEnery (eds.) *Proceedings of the Corpus Linguistics 2003 Conference. University Centre for Computer Corpus Research on Language Technical Papers*, Vol. 16, p 184-190. Lancaster University, United Kingdom

Yip, Moira (1988). The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. *Linguistic Inquiry* **19**, 65-100

Yip, Moira (1989). Feature Geometry and Cooccurrence Restrictions. *Phonology* **6**, 349-374