The Effect of Markedness in Phonotactic Speech Segmentation

Lianne van der Boom

MA Thesis Taal, Mens en Maatschappij (Engelse Taal en Cultuur), Utrecht University

Student code: 3282139
Supervisor: prof. René Kager
June 2011
16639 words

Table of contents

1. Introduction

For many centuries people have been fascinated by one of man's greatest accomplishments. It is one of the most important things that separates us from animals. It is Language[1]. We use Language to communicate with each other; to convey our thoughts and feelings to others. A language is very complex, and amazingly the world holds near to 7000 different languages (Anderson, 2005), which are all complex and equally expressive. Yet, despite its complexity, a newborn child generally has no problem acquiring the language of the language community it grows up in. Once the language in question has been acquired, people furthermore have no problems producing and perceiving it, which is quite extraordinary, because speech is a continuous stream of acoustic events in which boundaries between words and sounds cannot be sharply drawn. Human beings are not just able to produce and perceive verbal messages (speaking and listening), they are also able to produce and understand non-verbal messages (writing and reading). In writing, sounds are separated by different symbols. Additionally the words are separated by spaces, sentences by dots, and so on. This makes writing relatively easy to decode into a message compared to spoken language. However, reading and writing need to be explicitly taught to be mastered, while the acquisition of verbal language is an automatic process which only requires a sufficient amount of spoken language input.

This thesis aims to combine two ways of phonological study. On one end by examining how continuous speech is segmented by the listener into a message. Phonology is intrinsically connected to the segmentation process, and psycholinguistic research is often used to gain insight into how the human brain processes phonological information. On the other end this thesis provides a relevant typological approach to the specific phonological area this thesis focuses on, namely phonotactics. Typology compares the linguistic systems of different languages to find out what is common (universal) and what is different (language-specific). The reason why these two different approaches are combined is that certain notions from typological studies may provide a better insight into the segmentation process.

The structure of the remainder of this thesis is as follows. The remaining part of chapter 1 provides information on how typological research and phonology are connected, introducing several important phonological notions that are relevant for the remainder of this thesis. Chapter 2 delves deeper into the segmentation process, in particular by fully reviewing an article of McQueen (1998), in which he describes the role of phonotactic constraints on the

[1] Whenever the term Language has been written with a capital letter, it means that it refers to Language in general as a phenomenon. It is written with a small letter when it refers to a specific language, such as English.

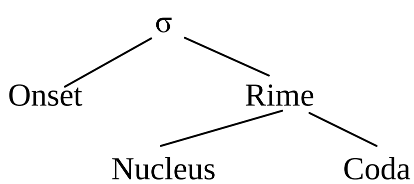
segmentation of continuous speech. Chapter 3 connects segmentation to typology, by fully reviewing an article of Berent, Steriade, Lunnertz and Vaknin (2007), in which they describe how input that violates the phonotactic constraints of the listeners is susceptible to perceptual illusions, and can be explained by the typological notion of *markedness*, which is the prime notion of the typological theory *Optimality Theory* (OT). Chapter 4 proceeds by delveing deeper into this typological approach, first by providing a detailed explanation of the typological theory in question, and second by providing an analysis of phonotactic constraints within OT. Chapter 5 reflects on both McQueen and Berent et al., leading to the formulation of two research questions. Chapter 6 provides the setup and results of two experiments, which were conducted in an attempt to find an answer to both research questions. Chapter 7 provides a general discussion and conclusion.

1.1. Phonology in typology

The grammars of different languages contain language-specific features, which are added to the universal basis of all languages during the language acquisition process (Guasti, 2002). This section provides several examples of language-specific and universal phonological rules and processes. While providing these examples several important phonological notions relevant to this thesis, such as syllable structure, phonotactics, and the sonority hierarchy, are discussed.

An example of a universal is the syllable structure of languages (Kaye, 1989). A syllable (σ) consists out of an *onset*, which are the consonants at the beginning of a syllable, and a *rime*, which is the remaining part of the syllable. The rime in turn consists out of a *nucleus*, which is the core of a syllable and occupied by vowels, and of a *coda*, which are the consonants at the end of a syllable. For example, in the syllable *pat*, the/p/ is the onset and /at/ is the rime. Within the rime the /a/ is the nucleus and the /t/ the coda. This is the basic structure of syllables that is universal to all languages, and is visualized in (1):

(1)



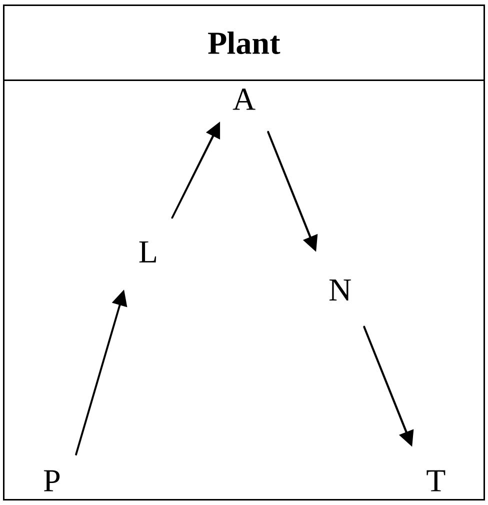
Although the structure in (1) is universal, its contents are not. Languages have different, language-specific phonotactic constraints. Phonotactics concerns itself with the study of syllable structures, in particular whether and what type of consonant clusters may appear in the onset and the coda. For example, Japanese allows no consonant clusters, and the coda is empty. The /n/ may fill the coda (e.g. *hanbaagaa* ‘hamburger’), but this is the only exception (Dupoux et al., 1999). Dutch and English on the other hand allow the coda to be filled with a variety of consonants, and even allow clusters in both the onset and the coda (e.g. *plant*) (Trommelen, 1983 for Dutch; Ostapenko, 2005 for English). Even between these two similar languages there are differences. Dutch allows the onset cluster /kn/ (e.g. *knoop* ‘knot’), while English does not. It is important to stress that this concerns the pronunciation of words. As can be seen in the translation of *knoop*, English likewise allows ‘kn’ in orthography, but the /k/ is not pronounced, while it is in Dutch.

To describe what a language does or does not allow in its syllables is usually described using the sonority hierarchy. A sonority hierarchy is a representation of how one sound is more sonorous (louder) than another. Vowels are the most sonorous, while plosives are the least sonorous. Likewise voiced sounds are more sonorous than voiceless sounds. A detailed sonority hierarchy is provided in (2), which is taken from Selkirk (1984). Generally, in an onset-nucleus-coda syllable, sonority rises from the onset to the nucleus, and falls from the nucleus to the coda. The word *plant* provides an example in (3). In the onset cluster /pl/ sonority starts at 0.5, after which it rises towards 6. It continues to rise towards its nucleus, with /a/ having an index of 10. Sonority falls in the coda /nt/ from 10 via 5 to 0.5.

(2)

Sound	**Sonority index (provisional assignment)**
a	10
e,o	9
i, u	8
r	7
l	6
m,n	5
s	4
v,z, ð	3
f, θ	2
b, d, g	1
p, t, k	0.5

(3)



The sonority hierarchy is arguably similar, but also slightly different between languages. Note for example that the /s/ has its own separate index of 4 while the other voiceless fricatives have an index of 2. The /s/ has interesting properties in English, because this language allows onsets to consist out of three consonants, but only if the first consonant is an /s/ (e.g. *spring,* but not **fpring*). Furthermore, English allows no obstruent-nasal onsets, unless the obstruent is an /s/ (e.g. *snake*, but not **fnake*). Clearly the /s/ causes exceptions in English. One theory is that /sC/ clusters are actually perceived as a single unit (Ostapenko, 2005). The /s/ does not cause exceptions in other languages, making its status the same as the language's other voiceless fricatives, meaning it is assigned the same index. On the other hand the status of the /n/ might be different from other nasals in Japanese, since the /n/ is the only sound that may appear in the coda. Languages can in this way differ in their specific sonority hierarchy, but they also have rules on how great the sonority differences have to be between two consecutive members of a syllable. In English, if the first member of the syllable is a /d/, then the next member needs to have an index of 7 or higher in Selkirk's hierarchy, because any combination of /d/ plus a sound with an index less than 7 is illegal in English. Russian however, even allows a cluster like /db/. If the first member is a /d/ in Russian, the second member needs to have an index of 1 or higher.

These relevant phonological notions provide a general overview of how phonology forms a very interesting field of study from a typological point of view. Another important area within phonology is speech segmentation. Kaye (1989) argued how phonological processes are necessary in the segmentation of continuous speech. He proposes an experiment, in which listeners have to listen to a stretch of speech in which all effects of phonological processes have been removed. He beliefs such an auditory stimulus would be incomprehensible when played back at normal speed. He did not perform such an experiment, but other researchers have examined what kind of phonetic and phonological information plays a role in the segmentation of speech. Chapter 2 delves deeper into this process.

2. Speech segmentation

Incoming speech needs to be segmented by the listener into a comprehensible message. Speech segmentation on the basis of phonetic and phonological information occurs at many different levels, namely at the sound, syllable, word, phrase, and sentence level. These levels provide the many different factors that have a role in this process. This chapter first provides a general overview of these factors, showing how intricate the segmentation process is. Section 2.1 moves on to the specific role of language-specific phonotactic constraints within this process.

At the sound level, the listener needs to distinguish one sound from another. Strange (1995) argued that listeners need to rely on *cues.* Cues are acoustic properties which characterize speech sounds (phones). Every phone has a number of cues. First language learners need to learn which cues are important for each individual sound in their native language, and rely on them to help decode the spoken message. For example, vowels differ in their value of the first two formants F1 and F2, their duration, and many more properties. To distinguish between the English /ɪ/ and /i/ (as pronounced in the words *bit* and *beat* respectively) a speaker of Southern English relies mainly on duration, while a speaker of Scottish relies mainly on the value of the F1 (Escudero & Boersma, 2004).

The characteristics of separate phones plays only a small role in the segmentation of speech into separate words. For example, in English the word-initial voiceless plosives /p/, /t/, and /k/ are aspirated, while these are not aspirated in word-final position (Lehiste, 1960). If a listener encounters an aspirated /t/ the aspiration is a cue, signalling the start of a word. However, not all words start with a /p/, /t/, or /k/, and there are cases in which these phones are aspirated in the initial positions of stressed syllables, whether they are word-initial or not (McQueen, 1998), making additional sources of information necessary. The prosodic features of a sentence and the phonotactic constraints of a language provide this additional information.

Prosodic features are expressed by stress and intonation. Word-stress languages follow a set of rules concerning which syllable carries main stress. For example, despite a few exceptions English words never have main stress on the final syllable. When listeners encounter a stressed syllable in English, they can be reasonably sure that it is not followed by a word boundary, unless the word only has one syllable (Cutler & Norris, 1988). Additionally, word-final syllables are lengthened (Beckman & Edwards, 1990; Klatt, 1975), providing an additional cue signalling an upcoming word boundary. The intonation of a sentence is partly expressed in the stress contour of its individual words, but not entirely. In a declarative sentence the pitch is higher at the beginning of the sentence than at the end (Rietveld & van Heuven, 2001). The melody of a sentence may help the listener to determine sentence boundaries.

As mentioned before, phonotactic constraints deal with the structure of the syllable itself. These constraints can aid the listener to determine syllable, and possible word boundaries. For example, English does not allow /mr/ to occur in either the onset or the coda of its syllables. When an English listener encounters this cluster he or she is sure that it marks a syllable boundary, and possibly even a word boundary.

As has become evident, the segmentation of speech is a very complex process that nevertheless causes few problems for its listeners. A question that remains is how big the role of each of these cues is on the entire process. One thought is that there is one central mechanism that can function independently, but is aided by phonetic and phonological information (Norris, 1994; Norris et al., 1995, 1997). McQueen (1998) discusses this mechanism, and the role of metrical information as an aid. Moreover, he investigated the specific role of phonotactic constraints as an aid to this mechanism, which is of interest to this thesis. His article is fully reviewed in the following section.

2.1. Phonotactics in speech segmentation

This section reviews McQueen's (1998) article. He notes that, although a speech signal is packed with all kinds of acoustic cues, many of those cues are ambiguous. He is particularly interested in the role of phonotactic constraints in the segmentation of continuous speech, and conducted three experiments to find an answer.

In his introduction McQueen discusses *Shortlist* (Norris, 1994; Norris et al., 1995, 1997), which is a competition model of spoken word recognition. In this model candidate words are activated according to the segmental material provided by the input. Next, Shortlist tends to settle on an optimal parse of the input, which has to result into a logical message. McQueen provides an example input to illustrate its workings. The string *enjoyable over-indulgence* activates a set of candidates. *Enjoy, enjoyable, joy, below, over-indulgence, indulgence,* and *dull* are only several of the activated candidates. *Enjoyable* and *over-indulgence* will win over the other candidates, since they can join forces to account for the complete input. If the candidates *enjoy a below ?? indulgence* were to win, the segment [vər] would be unaccounted for. As a competition model Shortlist provides a solution to the segmentation of speech in which the boundaries between words may not be acoustically cued.

McQueen likewise discusses one of his earlier experiments, in which listeners had more trouble finding the word *mess* when imbedded in the string [dəmɛs], the onset of the longer word *domestic*, than in the string [nəmɛs], which is not the onset of a longer word (McQueen et al., 1994). This provides additional evidence that word recognition is a process of lexical competition. Other types of experiments, such as cross-modal priming studies (Gow & Gordon, 1995, among others), recognition memory experiments (Wallace, Stewart & Malone, 1995; Wallace et al., 1995), and phonological priming studies (Slowiaczek & Hamburger, 1992, among others) have likewise provided additional evidence supporting competition models like Shortlist.

McQueen continues by arguing that lexical competition models do not provide a complete account of the segmentation process. Several studies have indicated that listeners also use metrical information. For example, English listeners found it harder to spot the word *mint* in the StrongStrong sequence [mɪnteɪv], than in the StrongWeak sequence [mɪntəv] (Cutler & Norris, 1988), arguably because segmentation starts at the onsets of strong syllables, but not at the onsets of weak syllables. In the StrongStrong sequence the /t/ of *mint* would be segmented as the onset of the second syllable instead of as the coda of the first syllable, while this does not happen in the StrongWeak sequence. Dutch listeners responded in the same way in a different experiment (Vroomen & de Gelder, 1995), providing additional evidence that metrical information is used in the segmentation process.

Several experiments attempted to find out whether lexical competition and metrical information work in tandem. It has turned out that they do. In experiments in which the metrical information was kept constant while the second syllable activated many or few lexical candidates, reaction times were slower when the second syllable activated more possible candidates (Norris et al., 1995; Vroomen & de Gelder, 1995). At this point, many believe that competition remains the central mechanism, but is aided by other sources of boundary information, such as metrical information (Norris et al., 1997). McQueen now turns to his discussion whether phonotactic information acts as an additional source.

McQueen provides the example string *the team require*. In the Shortlist model, words like *tea* and *acquire* are activated, but soon penalized because English does not allow the cluster /mr/ to appear in either the onset or the coda, providing an additional cue to the competition model that there must be a syllable, and potential word, boundary between *team* and *require*.

Other experiments have already shown that listeners are sensitive to phonotactic constraints. Even children as young as nine months show a preference to speech material meeting the phonotactic constraints of their first language (Jusczyk et al., 1993). When

English adults are provided with an ambiguous segment between /l/ and /r/, and are forced to make a choice which one it is, they are more likely to label it as an /r/ in /t?i/, and as an /l/ in /s?i/, because /tri/and /sli/ are legal, while /tli/ and /sri/ are illegal in English (Massaro & Cohen, 1983). Additional experiments, such as phoneme monitoring experiments (Pitt & Samuel, 1995, among others) have offered the same hint that phonotactic constraints play a role in speech segmentation, but McQueen wants to find out for sure, and conducted his own experiments.

Experiment 1 was a word-spotting task. Native speakers of Dutch were presented with auditory disyllabic strings, of which some contained embedded Dutch words. The subjects' task was to spot these words and press a button when they had. If phonotactic constraints aid in the segmentation of continuous speech, then they should have no problem spotting the word *pil* when aligned with a clear phonotactic boundary, such as /lv/ (e.g. [pɪl.vrem], but not *[pɪ.lvrem] or *[pɪlv.rem], because /lvr/ is an illegal onset cluster and /lv/ an illegal coda cluster in Dutch). On the other hand, if *pil* is misaligned with a clear phonotactic boundary, such as /mr/, then it should be hard to spot it (e.g. [pɪlm.rem], but not *[pɪ.lmrem] or *[pɪl.mrem] because /lmr/ and /mr/ are illegal onset clusters in Dutch). The listeners were also presented with sequences in which the target words were in final position. The metrical structure of the strings was likewise manipulated to find out whether phonotactic and metrical cues work as independent or as complementary cues. Each target word appeared four times in the experiment: Once in a StrongStrong, Aligned, once in StrongStrong, Misaligned, once in a StrongWeak or Weak Strong (depending on whether the target was in initial position (StrongWeak) or in final position (WeakStrong)), Aligned, and once in a StrongWeak or Weak Strong, Misaligned setup. The results are provided in table 1, which is copied from the article:

Measure	**Target position**	**Metrical structure**	**Aligned**	**Misaligned**
Errors	Initial	StrongStrong	32%	57%
		StrongWeak	38%	59%
	Final	StrongStrong	21%	56%
		WeakStrong	19%	63%
RT	Initial	StrongStrong	766	828
		StrongWeak	750	809
	Final	StrongStrong	535	629
		WeakStrong	499	614

Table 1. Mean percentage missed targets (errors) and mean reaction times for correct detection (RT, in ms), measured from target-word offset, in experiment 1.

The results show that the error rate was significantly higher when the target words were misaligned, as was expected. This effect is even stronger with the words in final position, by having a lower error rate for the aligned targets. The reaction times show a similar effect, being significantly faster when the target words were aligned, as was expected. In general target words in final position were detected faster than target words in initial position. Moreover, the significance between the aligned and misaligned reaction times was stronger in the final targets, than in the initial targets. These results provide a clear picture that listeners use the phonotactic constraints of their language as a cue in speech segmentation. McQueen additionally found no interaction between the different metrical structures, and argues that this provides evidence that metrical and phonotactic information provide complementary cues.

Although these results seem to be clear, McQueen notes that there are other factors that could have caused these results, such as the duration of the targets. The final targets were longer than the initial targets. Likewise the aligned targets were longer than the misaligned targets. A quick analysis of the effect of duration revealed that the effect of phonotactic alignment was still significant, but duration is not the only factor in play. McQueen decided to conduct an additional experiment as a control to find out whether his results are due to phonotactic alignment, or were caused by other acoustic cues.

Experiment 2 was a running lexical decision task. The targets of experiment 1 were excised from their contexts and presented as monosyllabic words to the subjects. The subjects had to press a button whenever they heard a real word. The participants of this experiment did not participate in experiment 1. If the results of experiment 1 were due to phonotactic alignment, then there should be no significant differences between the targets excised from the aligned or misaligned contexts. If however the results of experiment 1 are due to acoustic differences between the targets across contexts, then the same pattern of results should emerge. The results are provided in table 2, which is likewise copied from the article:

Measure	**Target position**	**Metrical structure**	**Aligned**	**Misaligned**
Errors	Initial	StrongStrong	19%	32%
		StrongWeak	27%	33%
	Final	StrongStrong	33%	32%
		WeakStrong	22%	25%
RT	Initial	StrongStrong	477	530
		StrongWeak	506	550
	Final	StrongStrong	379	399
		WeakStrong	349	344

Table 2. Mean percentage missed targets (errors) and mean reaction times for Yes responses to targets (RT, in ms), measured from target-word offset, in experiment 2.

An analysis of the data reveals that there are significant differences between both the error rates and the reaction times for the initial targets, indicating that the results of experiment 1 are due to acoustic differences between the targets across contexts, rather than due to phonotactic alignment. In contrast, this effect was not found among the final targets, indicating that the results of experiment 1 are due to phonotactic alignment for these targets. An additional analysis comparing the results of experiments 1 and 2 still reveals a significant effect on alignment, making McQueen keep to his conclusion that the results of experiment 1 are due to phonotactic alignment.

McQueen conducted a third and final experiment in an attempt to find out what contextual information is used to produce the alignment effect. However, his argumentation for conducting this experiment does not become really clear. It is likewise unclear what the results of this experiment tell us. Finally, this experiment is also irrelevant for the topic of this thesis. These are the reasons why I will not discuss this experiment in detail. McQueen finishes by concluding that phonotactic cues are used in speech segmentation, which is an interesting and important notion to my experiments, which will be discussed in chapter 5.

3. The strength of perceptual illusions caused by markedness

Speech segmentation is an automated process that requires no effort from the listener. However, when listeners encounter speech that does not conform to their grammar, for example by listening to a foreign language, this may give rise to perceptual illusions (Dupoux et al., 1999). In this chapter I will review an article by Berent, Steriade, Lunnertz and Vaknin (2007). In their article these authors link previous literature on perceptual illusions to the typological notion of markedness, and investigated whether marked onset clusters were more likely to elicit perceptual illusions than lesser marked ones. They also investigated whether the cause for the mistakes that were made are due to phonological repair or phonetic confusion.

In their introduction, Berent et al. first observe that there are unmarked and marked structures, and that the grammar of a language will always prefer the unmarked structure over the marked structure, whether or not both structures are attested in the given language (Blevins, 1995; Greenberg, 1978; Prince & Smolensky, 1993/2004). What if, however, one structure A is unmarked relative to structure B, but both structures are unattested in a language. Will the speakers of this language nevertheless be sensitive to the difference in markedness between these two structures? Berent et al. attempted to find an answer to this question, and investigated the role of markedness within onset clusters. They used the sonority hierarchy to describe four types of clusters, and also argued how each type has a different degree of markedness. Remember that sonority generally rises from the onset to the nucleus, and falls from the nucleus to the coda in syllables. However, when consonant clusters are involved this does not appear to be a universal rule. The sonority hierarchy in (4), taken from Berent et al., and the four types of onset clusters they describe illustrate this point:

(4)

Vowels and glides (5)	**Liquids (4)**	**Nasals (3)**	**Fricatives (2)**	**Stops (1)**
a, e/y, w	l, r	n, m	s, sh, z, f, v, th	p, b, t, d ,k ,g
More sonorous				*Less sonorous*

The least marked onset cluster is the *large rise (1→4)*. The sonority of this type of cluster rises considerably from the first member to the second. An example of this type of cluster is /bl/. A cluster that is a little bit more marked is the *small rise (1→3)*. The sonority of this type of cluster still rises, but to a lesser extend than in a large rise. An example of this type of cluster is /bn/. Yet more marked is a *plateau (1→1)*. The sonority of this type of cluster is stable across its members. The cluster /bd/ provides an example of this type of cluster. Finally, the most marked cluster is the *fall (4→1)*. In this type of cluster, of which /lb/ is an example, the sonority considerably falls from the first member to the second. All of these four types of clusters can appear in languages. As a matter of fact, all of these clusters are legal onset clusters in Russian.

Berent et al. continue to discuss the *implicational universals* regarding sonority profiles in typology by arguing that, if a language allows clusters of a certain degree of markedness, it also allows clusters that are less marked. For example, if a language allows falls, it also allows plateaus and both types of rises. On the other hand, if a language does not allow falls, but does allow plateaus, it will still also allow rises, etc. They verified this claim by the following reconstruction of the data of Greenberg (1978):

a.		**Large rise**	
		+	-
Small rise	+	57	1
	-	18	14

b.		**Rise**	
		+	-
Plateau	+	41	3
	-	35	11

c.		**Plateau**	
		+	-
Fall	+	11	1
	-	33	45

Table 4. The contingency between small sonority rise and larger sonority rise (a); sonority plateau and sonority rise (b); and sonority fall and sonority plateau (c). The presence of a cluster is indicated by +, whereas its absence is indicated by -.

Table 4a shows that Greenberg's survey contained 57 languages that allow both kinds of rises, 18 that only allow large rises, and 14 that allow neither, which are all possible combinations according to the implicational universals of Berent et al. Only one language allows small rises while not allowing large rises. However, the specific language in question, Santee Dakota, lacks liquids altogether, making it impossible for this language to have a large rise (obstruent-liquid cluster), meaning that it is not a true counter example. Tables 4b and c show the same pattern. Most languages either allow both types of clusters, or only the lesser marked, or neither. Only a few languages allow marked clusters while not allowing lesser marked ones. Exceptions apparently do occur, but the majority of the languages follow Berent et al.'s claim. Interestingly, Berent et al. likely claim that languages that allow only lesser marked clusters are more abundant than those with more marked clusters, which is in line with the concept of markedness from Optimality Theory. The data shows that this claim is likewise true. There are more languages that allow both kinds of rises, than there are languages that additionally allow plateaus, etc. The authors provide additional information to this claim by revealing that large rises made up 83%, small rises 64%, plateaus 49%, and falls 13% of the sample.

When speakers of a certain language need to produce a cluster that is illegal in their first language, they tend to split up the cluster and insert an epenthetic schwa in between the two members (Pertz & Bever, 1975 for English; Broselow & Finer, 1991 for Korean and Japanese) For example, speakers produce a word like /lbif/, which has only one syllable and an onset cluster with a falling sonority, as /lə.bif/, a word with two syllables and no onset cluster. Ostapenko (2005) specifically found this repair strategy to be used by an inexperienced English speaker of Russian, whereas two experienced English speakers of Russian favoured deletion of one member of the cluster. In contrast, the English native speakers participating in Davidson's (2000) study provided no systematic effects of sonority profile in production. The few available studies seem to contradict one another.

However, production is different from perception. Altenberg (2005) claimed that researchers should discuss them as two separate issues, and illustrated her point by showing that her subjects performed differently on a perception task than on a production task. Still, the same effect that appeared in some studies in production seems to be present in perception. English native speakers confuse an input like /tla/ with its epenthetic counterpart /tə.la/ (Pitt, 1998). Massaro & Cohen (1983), and Moreton (2002) likewise found that English native speakers' perceptions of illicit syllables are subject to perceptual illusions. Additional studies provided the same evidence for French (Hallé et al., 1998) and Japanese (Dupoux et al., 1999, 2001). In these studies syllables that shared a marked, unattested-unmarked, attested relationship were compared. Berent et al. observe that the perceptual difficulties with the former type of syllables might be simply due to the unfamiliarity with the marked structure. However, linguistic analysis suggests a different possibility. In this view, highly marked illicit inputs that are subject to perceptual illusions are misperceived because they are repaired by the organization of the grammar.

Berent et al. are particularly interested in whether different types of onset clusters, with varying degrees of markedness, and which are all unattested in a given language elicit varying degrees of perceptual illusions. For example, three of the four clusters described before are illegal onset clusters in English, a language that allows only large rises. They conducted four experiments to find an answer. They conducted an additional two experiments in an attempt to provide an explanation for the cause of perceptual illusions.

In experiment 1, native speakers of English were aurally presented with monosyllabic CCVC, and disyllabic CəCVC non-words of all four cluster types (of which the ones that are illegal in English were of interest). The participants had to quickly indicate whether the stimulus had one or two syllables. Experiment 2 was the same experiment, only then conducted among Russian native speakers. The results are provided in figure 1, which is taken from the article:

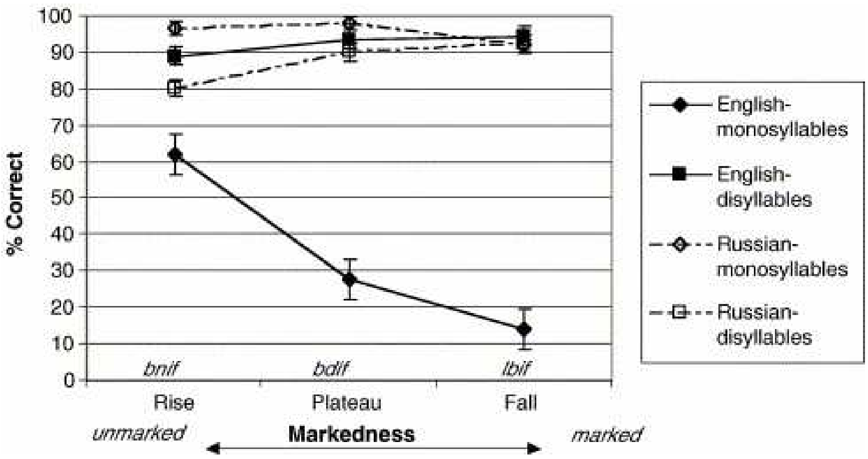


Fig. 1. Mean response accuracy of English and Russian speakers to monosyllabic nonwords and their disyllabic counterparts in Experiments 1-2 as a function of the number of syllables and the markedness of the monosyllabic counterpart. Error bars represent the confidence interval constructed for the difference among the means.

The graphs indicate that both language groups performed very well on the disyllabic stimuli, while the curves of the monosyllabic stimuli show great differences. The Russians were able to perceive them as one syllable, and were thus able to accurately perceive very marked onset clusters. The English however, perceived most monosyllabic stimuli as disyllabic. This effect is stronger when the cluster is more marked.

Since the results of experiments 1 and 2 were so strong, Berent et al. wondered whether native speakers of English were able to discriminate between monosyllabic and disyllabic stimuli at all. Experiments 3 and 4 were discrimination tasks in which the participants were aurally presented with pairs of non-words and had to indicate whether both stimuli were the same or different. Experiment 3 was conducted among native speakers of English, while experiment 4 was the same experiment conducted among Russian native speakers. The results are provided in figure 2, which is likewise taken from the article:

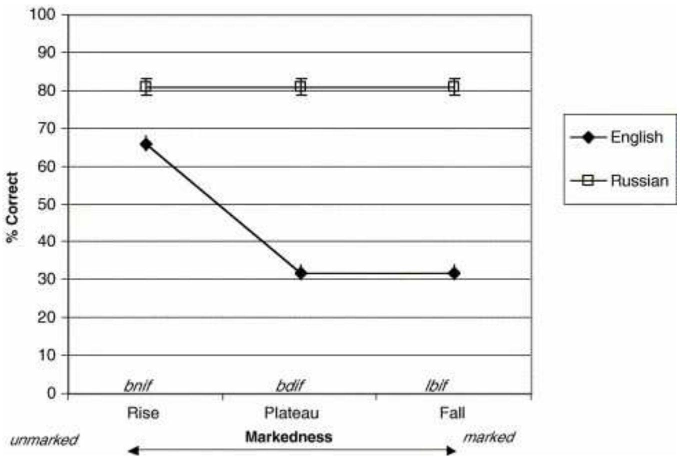


Fig. 2. Mean response accuracy of English and Russian speakers to non-identical trials in Experiments 3-4 as a function of the markedness of the monosyllabic input. Error bars represent the confidence interval constructed for the difference among the means.

Again the Russians performed very well, even showing no difference in the degree of markedness of the clusters. The English on the other hand again had trouble discriminating between the monosyllabic and disyllabic stimuli.

These experiments provide a very interesting picture, but do not explain why the English listeners misperceive monosyllabic non-words with illegal onset clusters as disyllabic. With experiments 5 and 6, which were double lexical decision tasks, Berent et al. tried to find out whether the cause for misperceptions was due to phonological repair, or phonetic confusion. If the participants are phonetically confused then they have trouble discriminating the /l/ of /lbif/ from a vowel, thus perceiving a vowel while it is actually not there. Linguistic experience, such as the Russians are exposed to, can overcome this problem. Phonological repair on the other hand states that if the input does not comply with the grammatical knowledge of the speaker, then the input is phonologically repaired to a form that does comply with the grammar.

The participants were all native speakers of English, but the ones participating in experiment 6 were “tacitly encourage[d] […] to attend to the presence of acoustic cues for a vowel in the auditory stimulus” (619). They argue that if participants are focussed on the marked clusters, then phonological repair can be overcome and errors are decreased, while phonetic confusion cannot be overcome, keeping the results the same. A comparison between the results of these two experiments indicated that the subjects participating in experiment 6 performed better, and that misperceptions are thus due to phonological repair.

These results are very interesting from two points of view. First by showing that perceptual illusions are sensitive to the markedness of clusters when all types of clusters are unattested in a given language. When all types of clusters are attested in a given language, hardly any misperceptions occur, and when they do, they are insensitive to the degree of markedness of the cluster. Second, the final two experiments suggest that input is phonologically repaired to conform to the listener’s grammar. These results combined suggest that the hierarchy of structural types (a sonority fall is more marked than a sonority plateau etc.) as presented by Berent et al. are realized in the grammar. This is very interesting for OT, which states that the grammatical knowledge of a speaker is hierarchically organized. The full extend of this claim is discussed in chapter 5, but first a detailed explanation of OT and an analysis of Berent et al.’s clusters is provided in chapter 4.

4. Optimality Theory and markedness in onset clusters

An important feature within typological studies is the notion of *Universal Grammar* (UG). It is assumed that a child is born with the notion of what Language is and how it will look like. For a child UG acts as a *language acquisition device*. In the 1990's McCarthy and Prince (1993) proposed a new way for looking at the grammar of Language. They proposed that the grammar of a particular language is simply a language-specific ranking of universal *constraints*. This theory is now widely used in phonology. This chapter first provides an overview of how their theory, called Optimality Theory, works by providing the basic information and an example to illustrate its workings. Next, I will quickly review and reject the OT analysis of onset clusters by Ostapenko (2005) before moving on to provide my own analysis.

Phonological processes mainly serve one purpose, namely to provide an output that is easier to pronounce than its underlying form. For example, Dutch has a phonological process called *final devoicing*, in which voiced obstruents become voiceless at the end of a word or syllable. English on the other hand does not have a final devoicing as a phonological process. Thus, a word like *hand*, which means the same thing in Dutch and English, is pronounced with a final /t/ in Dutch, but with a final /d/ in English. The underlying form of Dutch's *han[t]* is *han[d]*. This is derived from the fact that the plural is pronounced as *han[d]en* (Zonneveld, 1996). Voiceless obstruents are slightly less sonorous, and hence *unmarked*, compared to their voiced counterparts, and are therefore easier to pronounce. Dutch is said to be an unmarked language, while English is said to be a *marked* language with respect to this phenomenon.

McCarthy and Prince argue that there are two types of universal constraints: *Markedness constraints* (MARK) and *Faithfulness constraints* (FAITH). MARK causes the output to be unmarked, while FAITH causes the output to be the same as its underlying form, whether it is marked or unmarked. For example, again in the final devoicing case, languages that have this phonological process have a highly ranked MARK prohibiting obstruents in coda position to be voiced, causing an unmarked alternative. Languages that do not have this process have a highly ranked FAITH forcing the output to share its voicing features with the underlying form. It is important to note that the output must either be equally marked or unmarked compared to the underlying form. The interaction between MARK, which causes unmarked candidates to prevail over marked candidates, and FAITH, which causes the degree of markedness to be equal, can never result in a more marked output.

The constraints form the core of the theory, but they work in tandem with two mechanisms. Archangeli (1997) provided an overview. She explained that, in OT, UG consist out of three elements, namely CON, GEN, and EVAL. CON is simply a short term for the constraints discussed above. GEN is short for *generator*. This mechanism provides an infinite candidate set for each underlying form. For the underlying form *han[d]* it provides the two logical candidates *han[d]* and *han[t]*, but also *finger, table,* and *pirakoshinerino,* of which the latter is just a non-word. This infinite set of candidates is then evaluated by EVAL, which is short for *evaluation*. EVAL is a mechanism that tests each candidate to the underlying form by use of the specific ranking of CON. Candidates that violate a certain constraint are eliminated from the set until only one candidate, the winner, remains. In English, *finger, table, pirakoshinerino,* and *han[t]* will have to be eliminated in favour of *han[d]*, while in Dutch *han[d]* needs to be eliminated before *han[t]*.

The evaluation of candidates to an underlying form is visualised in tableaux. A tableau is a table that provides the following information. The underlying form is represented in the upper left corner. On the horizontal axis are the constraints in a certain order (ranking). At this point there is no saying how many constraints there are and what their effect is. Also, for the evaluation of one underlying form there are usually only a few constraints that are relevant. A tableau therefore only visualizes the constraints that are relevant for the current evaluation. On the vertical axis are the candidates. It is impossible to visualize the evaluation of an infinite candidate set, so a tableau only provides the most logical ones.

To give an example of how an evaluation is visualized in tableaux I will use the OT analysis of final devoicing by Lombardi (1999). An OT analysis of phonotactic constraints would have served as a more relevant example. However, one of the few available analyses, namely Ostapenko’s (2005), will soon be dismissed because of its flaws, and another analysis, namely Gnanadesikan’s (2004), moves slightly beyond the basics of OT. Lombardi’s final devoicing analysis is relatively simple, and elegant in its simplicity, providing a far better example.

In a former analysis Lombardi argued that the onsets and codas of a syllable can have a *laryngeal node*, which specifies the phone for certain features, namely voicing, aspiration, and glottalization (Lombardi, 1995). If a phone has a laryngeal node which specifies it for [voice], then the phone is voiced. If, however, the phone does not have a laryngeal node, it cannot be specified for [voice], resulting in a lack of voice. Languages that act like English allow laryngeal nodes to appear in both the onset as well as in the coda, allowing voiced obstruents to be in coda position. Languages that act like Dutch however, only allow laryngeal nodes to appear in the onset. Any laryngeal node appearing in de coda, because the underlying form has a voiced obstruent in coda position, will be deleted, resulting in a voiceless obstruent. She proposes an interaction between three constraints, two FAITH and one MARK:

(1) FAITH: IDentOnset(Laryngeal) (abbreviated IDOnsLar)
- ➢ In the output consonants in the onset position should have an identical (faithful) laryngeal specification as the underlying form.

(2) FAITH: IDent(Laryngeal) (abbreviated IDLar)
- ➢ In the output consonants should have identical (faithful) laryngeal specifications as the underlying form.

(3) MARK: *Lar
- ➢ In the output consonants may not have any laryngeal features.

For both Dutch and English we want the onsets to be identical to the underlying form. IDOnsLar should be ranked high in both languages. Next, Dutch needs to rank *Lar high to make sure laryngeal nodes to not appear in the coda, while English needs to rank IDLar high to make sure laryngeal nodes do appear in the coda. Tableaux (1a) and (1b) visualize the evaluation of *band*:

(1a) Dutch			
/band/	IDOnsLar	*Lar	IDLar
1. → bant		*	*
2. band		**!	
3. pant	*!		**
4. pand	*!	*	*

(1b) English			
/band/	IDOnsLar	IDLar	*Lar
1. bant		*!	*
2. → band			**
3. pant	*!	**	
4. pand	*!	*	*

These tableaux show how a slightly different ranking of constraints cause a different candidate to win. For both Dutch and English candidates 3 and 4 violate IDOnsLar, which is indicated by a star. The exclamation point indicates that the violation is fatal, causing the candidates to be eliminated from the evaluation, hence why the subsequent cells are grey. The stars in the grey cells are only meant to provide a complete picture, but have no more say in the remaining evaluation. Next, for Dutch the evaluation of *Lar causes the second candidate to be eliminated. Note that both remaining candidates violate this constraint, but the second candidate violates it twice by having two laryngeal nodes, both in the onset as in the coda. Since the violation of the second candidate is greater than the first, the second is eliminated in favour of the first one. Only one candidate remains after the evaluation of *Lar, which automatically becomes the winner. IDLar has no more effect on the remaining evaluation in Dutch, while it has in English. In English the evaluation of IDLar causes the first candidate to be eliminated in favour of the second. Here *Lar is irrelevant to the outcome. Finally, note that, based on these constraints, candidate 4 can never be the winner, since it violates every constraint. This implicates that there are no languages in this world that have *initial devoicing* as a phonological process. Lombardi claims that this statement is true.

This example provides a good view of how OT explains the differences between languages and how the evaluation can be visualized in tableaux, which is vital to understand the OT analysis of onset clusters in the next section. However, how do Dutch and English children come to rank their constraints differently with regard to this issue? It has been assumed that in the initial state of language acquisition all MARK outrank FAITH (Gnanadesikan, 2004). Constructing the correct constraint hierarchy requires the child to demote MARK below FAITH when it receives marked input. For example, English children are known to go through a stage in which they devoice final obstruents. They eventually demote the MARK *Lar below the FAITH IDLar, enabling them to produce voiced final obstruents.

4.1. Phonotactic OT analysis of onset clusters

This section provides a typological approach to onset clusters within OT. Ostapenko (2005) has attempted to create an OT analysis of onset clusters, which will be quickly reviewed. This analysis has a few flaws and will be rejected for it. I will then move on to my own analysis to describe five types of languages, according to the four types of clusters described by Berent et al.

Ostapenko (2005) provided a short OT analysis of onset clusters by describing the second language acquisition process of Russian syllable structures by English native speakers. She proposes an interaction between four constraints, two FAITH and two MARK. The most important constraint in her analysis is the OSon constraint. This is a MARK that states that "for 2 segments to be parsed in the same onset, a certain distance in the sonority scale must be maintained. This distance would be different for different languages" (147). This constraint carries a number of problems with it. First of all, the constraints that form the grammars of languages are all universal in OT. This is the reason why it is illogical to assume that a language-specific feature can be added to a universal constraint. Furthermore, if such a constraint were to exist, then languages would not so much differ in their specific ranking of the relevant constraints, but in the sonority distance this constraint forces on the members of a cluster. Ostapenko claims that OSon is ranked above both FAITH in English, while it is ranked below both FAITH in Russian. However, it does not become really clear from the tableaux why this would be the specific rankings for these two languages. I will attempt to provide a more complete and accurate analysis.

There are roughly five types of languages concerning the types of onset clusters they allow. Belonging to the first type are languages that disallow any type of clusters. A language that belongs to this type of language is Japanese. The second type of language allows only large rises. As has become evident from the previous chapter, English belongs to this type. The third type of language allows small rises additional to large rises. Berent et al. name Ancient Greek to be such a language. They likewise name Hebrew to be of the fourth type, namely a language that allows plateaus additional to both types of rises. Finally, it has already become evident that Russian allows falls additional to plateaus and rises, making it an example of the fifth type. I will deal with these languages in the following order: (1) Japanese and Russian, (2) Hebrew, (3) Ancient Greek, and (4) English.

4.1.1. Japanese and Russian

Japanese and Russian are each other's complete opposite in the sense that Japanese allows no clusters, while Russian even allows very complex and marked clusters. Their OT analyses should reflect this. Not having any clusters is unmarked compared to having clusters. Likewise inserting a vowel to avoid a cluster is unmarked to keeping the cluster as it is. Remember that the output must either be equally marked or unmarked compared to the underlying form (in this case, the input). This means that the constraints may only cause a cluster to be broken up by a vowel in the output, instead of deleting a vowel to cause the output to have a cluster. The underlying forms must contain a cluster, after which in Japanese a MARK will cause the cluster to be broken up, while a FAITH will keep the cluster as it is in Russian. There are two constraints that conform to this analysis:

(1) MARK: *ComplexOnset (abbreviated *Complex)
- An onset may not consist out of more than one consonant.

(2) FAITH: Dependency Input-Output (abbreviated Dep-IO)
- Do not insert vowels.

Tableaux (2a) and (2b) visualize the evaluation of Berent et al.'s four types of clusters in these two languages. Note that after the evaluation of *Complex in Japanese there are actually more logical candidates possible. Deleting one consonant is also a possible unmarked repair strategy. For example, the winning candidate for /bnif/ could also be either /bif/ or /nif/. Likewise the evaluation of Dep-IO in Russian does not prevent deletion either; it only causes /bə.nif/ to be eliminated. This can easily be solved by using a FAITH called

Maximality, a constraint that prohibits segments to be deleted. However, to keep matters simple, I will act as if all language types have this constraint ranked highly, eliminating candidates like /bif/ and /nif/ in favour of /bnif/ and /bənif/. This complies with the findings that the input is phonologically repaired, causing a perceptual illusion to occur in which listeners perceive an epenthetic schwa in between the two members of the cluster (Pitt, 1998; Massaro & Cohen,1983; Moreton, 2002; Hallé et al., 1998; Dupoux et al., 1999, 2001). In short, these tableaux visualize the perception of the clusters conform to these previous findings, though they might differ for production.

(2a) Japanese		
/lb, bd, bn, bl/	*Complex	Dep-IO
Lbif	*!	
→ Lebif		*
Bdif	*!	
→ Bedif		*
Bnif	*!	
→ Benif		*
Blif`	*!	
→ Belif		*

(2b) Russian		
/lb, bd, bn, bl/	Dep-IO	*Complex
→ Lbif		*
Lebif	*!	
→ Bdif		*
Bedif	*!	
→ Bnif		*
Benif	*!	
→ Blif		*
Belif	*!	

These tableaux show how the difference in ranking between but two constraints explain two completely opposite languages. *Complex makes sure that any cluster will be eliminated before Dep-IO can prevent it in Japanese, while Dep-IO prevents it in Russian before *Complex can eliminate he clusters.

4.1.2. Hebrew

The evaluation of Russian and Japanese cannot be completely fitted on languages of the other three types. Remember that Hebrew is nearly as marked as Russian, but does not allow clusters of the most marked type, those with sonority falls. Hebrew thus allows a subset of Russian, which is a classical situation for the existence of a MARK. Hebrew needs to be faithful in most cases, so like in Russian Dep-IO should be ranked high. However, Hebrew should not be faithful for sonority falls like in /lbif/. Harms (1973) proposed a universal syllable-wellformedness condition, which stated that voiced obstruents must be closer than voiceless ones to the syllable nucleus. Lombardi (1999) used this generalization in her OT analysis of English plural formation. What is interesting about this generalization is that, since voiced obstruents are slightly more sonorous than voiceless obstruents (remember the sonority hierarchy of Selkirk (1984) in (2)), Harms in essence states that sonority may not fall in the onset, and not rise in the coda, though the actual statement is a bit more limited than this. Inspired by this generalization I propose a new MARK that will set off Hebrew from Russian:

(1) MARK: *SonorityFall [onset] (abbreviated *SonFall [onset])
- ➢ In an onset cluster, sonority may not fall towards the nucleus.

It is important to specify this constraint for the onset, since sonority should actually fall in the coda to be unmarked. In order for this constraint to work for Hebrew, it needs to be ranked before Dep-IO and *Complex, as is shown in tableau (3).

(3) Hebrew			
/lb, bd, bn, bl/	*SonFall [onset]	Dep-IO	*Complex
Lbif	*!		*
→ Lebif		*	
→ Bdif			*
Bedif		*!	
→ Bnif			*
Benif		*!	
→ Blif`			*
Belif		*!	

With the addition of one constraint, subset languages like Hebrew, are set off from superset languages like Russian. Note how it only affects the evaluation of /lbif/, since the other clusters do not contain a falling sonority. In Russian *SonFall [onset] should be ranked below Dep-IO to prevent /lbif/ to be eliminated before /ləbif/. For Japanese it does not really matter how *SonFall [onset] is ranked, since *Complex takes care of all types of clusters. However, remember that it has been assumed that in the initial state of language acquisition all MARK outrank FAITH, and that constructing the correct constraint hierarchy requires the child to demote MARK below FAITH when it receives marked input (Gnanadesikan, 2004). A Japanese child does not receive any marked input concerning onset clusters, and has therefore no reason to demote MARK below FAITH. Although the actual ranking between *SonFall [onset] and *Complex does not make any difference in the outcome, *SonFall [onset] must be ranked above Dep-IO according to this theory. I will return to this issue when discussing the research questions in chapter 5.

4.1.3. Ancient Greek

Languages like Ancient Greek only allow sonority to rise in onset clusters, which makes them a subset of languages like Hebrew. Clusters with a falling sonority will be taken care of by *SonFall [onset], but that still leaves to deal with plateaus. Again an OT analysis of English plural formation provides inspiration. In an overview of OT, Zonneveld (1996) provided an analysis of which one constraint is of interest. In short, English has three plural morphemes, /-z/, /-s/, and /-iz/. The latter morpheme is used when the singular ends in a sibilant. The constraint that makes sure /bus-iz/ wins over /bus-s/ and /bus-z/ is called *SibilantSibilant, a MARK that prohibits two sibilants to be adjacent. In the onset cluster /bd/ a similar situation occurs where two obstruents are adjacent. A constraint like *ObstruentObstruent will make sure /bd/ will be eliminated before its epenthetic counterpart. However, a sonority plateau is simply a cluster in which the sonority is stable across its members, meaning that liquid-liquid, nasal-nasal, etc. are likewise classified as sonority plateaus. *ObstruentObstruent is too limited, giving rise to a more general MARK:

(1) MARK: *Plateau [onset]
   - In an onset cluster, sonority may not be stable across its members.

This constraint may not need to be specified for the onset, since it is logical to assume that, if a language disallows plateaus in the onset, it will also disallow plateaus in the coda. The reason for specifying this constraint for the onset will be discussed in chapter 5. In order for this constraint to work in Ancient Greek, it needs to at least be ranked before Dep-IO. Its relationship to *SonFall [onset] does not matter for the outcome, hence why the border between these two constraints in tableau (4) contains a dotted line.

(4) Ancient Greek				
/lb, bd, bn, bl/	*SonFall [onset]	*Plateau [onset]	Dep-IO	*Complex
Lbif	*!			*
→ Lebif			*	
Bdif		*!		*
→ Bedif			*	
→ Bnif				*
Benif			*!	
→ Blif`				*
Belif			*!	

With this additional constraint subset language like Ancient Greek can be set off from superset languages like Hebrew. Note how this constraint only affects the evaluation of /bdif/, since it is the only cluster that has a plateau. In Russian and Hebrew, *Plateau [onset] needs to be ranked below Dep-IO to prevent /bdif/ to be eliminated before /bədif/. Since this is another MARK, and Japanese children have no reason to demote it, it should outrank Dep-IO, but its ranking to *Complex and *SonFall [onset] does not provide any difference for the end result of the evaluation.

4.1.4. English

Finally, we move on to the last language type. English only allows large rises, so not only /lbif/ and /bdif/ need to be eliminated, but /bnif/ as well, while /blif/ should win. The sonority rises in both clusters, preventing a constraint that demands sonority to rise to have any effect. A possible constraint could demand the sonority difference between its members to be optimal. This constraint carries a number of issues along with it. First of all, it would not only eliminate /bn/, but /bl/ as well, since the optimal sonority difference between two members of a cluster is a voiceless obstruent followed by a liquid. In the evaluation of /blif/, /plif/ would be the optimal candidate. Second, since it wants the sonority difference of all cluster types to be optimal it would likewise select /plif/ as an optimal candidate for /bnif/. Even if it does select /pnif/ as the optimal candidate an illegal English cluster remains. Maybe Ostapenko's OSon constraint can be of help, by specifying it for a certain sonority distance that is equal to all languages. However, as can be seen by comparing Selkirk's hierarchy in (2) and Berent et al.'s hierarchy in (4) there are differences in how different sounds are classified. Ostapenko likewise used a sonority hierarchy that is even more basic than the one from Berent et al., by assigning all obstruents, whether they are fricatives or plosives, voiced or voiceless, the same index. It is difficult to assign OSon a clear sonority difference. The final MARK that I will propose keeps matters much simpler, even though the nature of the constraint cannot be made valid by other theories:

(1) MARK: *SecondConsonant [nasal, onset] (abbreviated C2 Nas [onset])
- In the onset, the second member of the cluster may not have the feature [+nasal].

This constraint has to be specified for the onset, since liquid-nasal clusters in coda's are perfectly legal in English. If this constraint is ranked above Dep-IO in English, it will make sure /bnif/ is eliminated before /bənif/. Its specific ranking concerning the other two MARK, *SonFall [onset] and *Plateau [obs], does once again not matter for the outcome. Tableau (5) shows the evaluation of the clusters for English:

(5) English					
/lb, bd, bn, bl/	*SonFall [onset]	*Plateau [onset]	*C2 Nas [onset]	Dep-IO	*Complex
Lbif	*!				*
→ Lebif				*	
Bdif		*!			*
→ Bedif				*	
Bnif			*!		*
→ Benif				*	
→ Blif`					*
Belif				*!	

Once again the addition of one constraint can set off subset languages like English from the other superset languages by affecting only one type of cluster. *C2 Nas [onset] needs to be ranked below Dep-IO for languages that allow obstruent-nasal clusters to make sure /bnif/ is not eliminated before /bənif/. Since *C2 Nas [onset] is a MARK, it is ranked above Dep-IO in languages like Japanese.

4.2. Conclusion

The interaction between five constraints, four MARK and one FAITH, explains how languages deal with different types of onset clusters with different degrees of markedness. Essential is the ranking of Dep-IO with respect to the MARK. This typological analysis shows how illicit input is segmented by the listener and supports Berent et al.'s findings. The previous chapters served to provide enough information to understand the line of thought that will lead to the formulation of two research questions, which will be discussed in the next chapter.

Before moving on to the reflection of both McQueen and Berent et al., several final issues need to be addressed concerning this OT analysis of onset clusters. It provides a bit of a black and white picture. One example is that the analysis implies that if a language allows large rises, it allows all large rises, while this is not true. There are two English clusters that have a large rise in sonority but are illegal. They are /tl/ and /dl/. Interestingly, the members of these clusters share the place of articulation. There is a constraint that deals with these kinds of situations. The *Obligatory Contour Principle* (OCP) (McCarthy, 1986) prohibits representations with two identical elements to be adjacent. The specific element can be specified. By specifying the element for [place] the MARK OCP [place] prohibits representations, in this case the members of onset clusters, with the identical element [place] to be adjacent. If this constraint is ranked above Dep-IO, it will make sure /tl/ and /dl/ are eliminated in favour of their epenthetic counterparts.

Another issue that needs to be addressed are several /sC/ clusters that are legal in English. Consider the following legal English onset clusters: /sl/ (e.g. *sleep*), /sn/ (e.g. *snake*), /sf/ (e.g. *sphere*), and /st/ (e.g. *stone*). These clusters are of different markedness values. The above analysis of English causes /sn/, /sf/, and /st/ to be eliminated by *C2 Nas [onset], *Plateau [onset], and *SonFall [onset] respectively. OCP [place] even judges three of these clusters, namely /sl/, /sn/, and /st/, to be illegal. Remember that the /s/ is known to cause exceptions, and that /sC/ might be perceived as a single unit (Ostapenko, 2005). The /sC/ clusters are avoided in the upcoming experiment, hence a further analysis of these types of clusters is irrelevant and beyond the scope of this thesis.

Finally Dutch is a language that shows how grey the in-between area is. At first glance Dutch seems to be like Ancient Greek, because it allows both large and small rises. However, the only small rises Dutch allows are those in which the first member is a voiceless obstruent and the second member an /n/, but, with the exception of /kn/, most clusters are low frequent. A highly ranked OCP [place] will take care of clusters like /pm/, /dn/, etc., but still leaves us to deal with clusters like /bn/, /tm/ etc. Is Dutch like Ancient Greek, but with restrictions on what type of small rises it allows, or is like English, but with a few exceptional small rises? Ranking a FAITH like –Voice + /n/ before *C2 Nas [onset] and Dep-IO in Dutch analyses it as English with a few exceptional small rises, but I must admit this constraint might be a little bit too specific. Again, a more detailed analysis is irrelevant and beyond the scope of this thesis. The above analysis provides a basis on which specific languages can build upon.

5. Reflection McQueen and Berent et al., and research questions

McQueen's (1998) finding that listeners use their language-specific phonotactic constraints in the segmentation of speech is very interesting. However, McQueen did not take the markedness of the different types of clusters he used into account. Since listeners use phonotactics in speech segmentation it might be interesting to see whether the degree of markedness of a medial cluster makes any difference. If, in a hypothetical situation, cluster A is marked with respect to cluster B, and both clusters are illegal as onset or coda clusters in the given language, will the occurrence of cluster A provide a stronger syllable boundary than cluster B? For example, suppose for the moment that the cluster /mn/ is the marked cluster A, while /tn/ is the unmarked cluster B, and both clusters force a listener of this particular language to place a syllable boundary in between its two members. Will the occurrence of /mn/ in [dem.not] (assuming that both *[demn.ot] and *[de.mnot] are illicit segmentations) give a stronger syllable boundary than the occurrence of /tn/ in [di:t.not] (assuming that both *[di:tn.ot] and *[di:.tnot] are illicit segmentations)?

It is important to note that the markedness of a particular cluster depends on its syllabification. As soon as listeners receive the input *deemnoot* their phonotactic constraints will cause it to be syllabified as CVC.CVC, because the CVCC.VC and CV.CCV syllabifications will cause a cluster to appear that is illegal in the listeners' language. The CVC.CVC syllabification can never be more marked than the other two syllabifications, because not having a cluster is always unmarked opposed to having one. The segmentation of *deemnoot* into *deem+noot* is not marked opposed to the segmentation of *dietnoot* into *diet+noot*, because both segmentations contain no cluster. Markedness starts playing a role the moment the grammar of the language considers the other possible syllabifications. If both sequences are syllabified as CV.CCVC two different clusters with different degrees of markedness are the result. Since /mn/ is the more marked onset cluster, it is less preferred than /tn/, providing a stronger cue for a syllable boundary; the CVC.CVC syllabification.

Interestingly, the opposite is true when both sequences are syllabified as CVCC.VC. As a coda cluster the /tn/ cluster is actually more marked than /mn/. However, the CVCC.VC syllabification is less likely to occur than a CV.CCVC syllabification, because having an onset is unmarked opposed to not having one (Gnanadesikan, 2004). The CVC.CVC and the CV.CCVC syllabifications are the only relevant ones.

If it is the case that more marked clusters provide a stronger cue for a syllable boundary then the question remains how this manifests itself. What effect does a stronger syllable boundary have on the segmentation process? I propose that a strong syllable boundary caused by a relative marked cluster is segmented faster than a weak boundary caused by a relative unmarked cluster. In this view, the input *deemnoot* of the above hypothetical situation is segmented faster into *deem+noot* than *dietnoot* is in *diet+noot.* It is now possible to assume that the target word *noot* can be spotted faster in a word-spotting experiment when it occurs in a context with a marked cluster than in a context with an unmarked cluster. This may result in increasingly faster reaction times when the cluster becomes more marked. This is in essence the first research question, which is put in more general terms in (1):

(1) Is a target word easier to find in context A than in context B, in which A and B share a marked-unmarked relationship respectively, resulting in faster reaction times for A than for B in a word-spotting experiment?

Before we can move on to actually discussing this type of experiment, we first need to reflect back on Berent et al. (2007), specifically in combination with my OT analysis of chapter 4. At this point it may be relatively unclear how my OT analysis of markedness in onset clusters supports Berent et al.'s findings that perceptual illusions are due to phonological repair of the input and are sensitive to the degree of markedness of the cluster. Remember that the English participants misperceived illegal clusters as their epenthetic counterparts, and that this effect was stronger when the cluster was more marked. For the outcome of the analyses, the ranking between the MARK does not matter, but what if their ranking influences how input that does not comply with the speaker's grammatical knowledge is perceived? For example, instead of the ranking *SonFall [onset], *Plateau [onset], *C2 Nas [onset], *Complex >> Dep-IO, in which the order of the MARK is free, I propose the ranking is actually *SonFall [onset] >> *Plateau [onset] >> *C2 Nas [onset] >> *Complex >> Dep-IO, in which the order of MARK is set. In this setup, inputs with sonority falls are eliminated sooner than inputs with plateaus, giving a stronger rejection pattern. The Russians did not show a preference for unmarked over marked clusters. This could be because all epenthetic counterparts are eliminated at the same stage of evaluation by Dep-IO. A set ranking of constraints will provide an explanation as to why marked clusters provide a stronger syllable boundary, speeding up the segmentation process. The question whether this is the case is in essence the second research question, which is stated in (2):

(2) Within Optimality Theory, when multiple rankings between several constraints cause the same candidates to win, is their order really free, or are they set, providing an explanation for the phonological repair strategy found by Berent et al.?

The MARK *SonFall [onset] restricts clusters of a more marked type than *Plateau [onset] does. The more marked a structure is, the less preferred it is by Language, making it conceivable that *SonFall [onset] outranks *Plateau [onset]. For the same reason *Plateau [onset] might outrank C2 Nas [onset]. However, this argument does not hold for *Complex, since this constraint restricts any type of cluster, regardless of how marked they are. Nevertheless, there are plenty of languages that allow consonant clusters, making it conceivable that, although having clusters is marked opposed to not having them, it is not so great a violation. In any case, a set ranking of constraints can account for the Berent et al.'s results that the English native speakers made more mistakes when the provided structure was more marked.

Note that I have deliberately specified all the MARK for [onset]. Reason being that it strengthens the claim that these MARK share a hierarchical relationship in which they are universally ranked. Gnanadesikan (2004) discussed a similar situation by incorporating the sonority hierarchy as multiple 'sub-constraints' within the FAITH Onset, a constraint that demands all syllables to have an onset. She is not the first to translate *harmony scales* within OT constraints (Prince& Smolensky, 1993; Jäger, 2003). She made her analysis after observing the speech of her young daughter. Young children go through a stage in which they reduce consonant clusters by deleting one of its members. Gnanadesikan argues that at this stage of language acquisition the MARK *Complex still has to be demoted. Most interestingly is the fact that children are very consistent in which consonant they delete. Gnanadesikan's daughter consistently said [ki:n] when she meant *clean*, and never said [li:n]. By hierarchically ranking the harmony scale of the universal sonority hierarchy Gnanadesikan was able to explain this consistency. In short, by ranking highly sonorous sounds like the /l/ before lesser sonorous sounds like the /k/, the /l/ and thus /li:n/ is eliminated before /k/ and thus /ki:n/. Perhaps phonotactic constraints are governed by one main constraint, which is subdivided into several sub-constraints that are universally ranked.

Another important part of the claim that the ranking between these MARK is not free is that is offers much for the implicational universals discussed by Berent et al., who stated that if a language allows clusters of a certain degree of markedness, it has to allow clusters that are less marked, which has been proven to be correct despite a few exceptions. If a child encounters clusters with sonority falls, it will eventually demote *SonFall [onset] below Dep-IO. This child's UG will tell it that, because its language has sonority falls, it should likewise expect to encounter plateaus and rises. This expectation might cause the child to demote

*Plateau [onset], C2 Nas [onset], and *Complex on the basis of just one kind of input. This is in line with a claim made by Dinnsen (2007), who argued that providing a specific type of input can solve multiple problems in phonological delayed children, by demoting multiple MARK below FAITH at once. In short, all a child has to do is determine where to put the FAITH Dep-IO in relation to the MARK.

Berent et al.'s reconstruction of Greenberg's (1978) data shows that there are a few exceptional languages, and that there are thus more constraint combinations are possible. Consider a language with the following constraint hierarchy: *SonFall [onset], *C2 Nas [onset] >> Dep-IO >> *Plateau [onset], *Complex. This hypothetical language does not allow falls and small rises, while it does allow plateaus and large rises. This language clashes with the implicational universals stating that languages must allow clusters that are less marked if this language allows clusters that are more marked. The clash occurs between the small rise and the plateau. Since this language allows plateaus, it should also allow small rises, while this is not true. How does my theory deal with this kind of situation? A child growing up in this situation is exposed to plateaus, and will therefore expect both kinds of rises, causing it to demote all MARK but *SonFall [onset] below Dep-IO. For this child it causes no problems to have *C2 Nas [onset] set below Dep-IO. In essence, this language can be described as a language that actually does allow small rises, but does not have any positive evidence of its existence; these types of cluster have a frequency of 0. This assumption needs to be tested among people living in this kind of situation. Although their language does not allow any small rises, are they able to accurately perceive and produce such clusters? This investigation is beyond the scope of this thesis, but is an interesting one nonetheless and had to be addressed.

6. Experiments

Two experiments were conducted to find an answer to both research questions. By putting a target word in different contexts, of which the markedness of the medial cluster as an onset cluster differs, the influence of the degree of markedness of the cluster on the segmentation process can be investigated. If markedness plays a role in the segmentation process, and marked clusters speed up the segmentation process by providing a stronger syllable boundary than unmarked clusters, then this might result in faster reaction times when the cluster is more marked. This likewise means that the ranking of the MARK is set, causing more marked clusters to be eliminated before lesser marked clusters. If such an effect is not found it does not mean that the degree of markedness is not used, it merely indicates that its role might be too small to be found in this experiment. It likewise does not prove the ranking of the MARK is free.

Section 6.1 discusses the materials, participants, and results of experiment 1 in more detail. Experiment 2 is discussed in section 6.2, which was conducted as a control of experiment 1 in order to see whether any effect are due to phonotactic cues, or other acoustic cues.

6.1 Experiment 1

This experiment is a word-spotting experiment. This type of experiment was chosen because it is closely related to segmentation, and McQueen has already shown that this type of experiment provides a good way to investigate the role of phonotactic constraints on the segmentation process. My segmentation task differs from that of McQueen. In McQueen's task the context in which the target word appeared demanded different syllabifications. The syllabification of my items will always be the same. The effect of markedness is a more indirect effect, which may be more difficult to find.

6.1.1 Materials

It would have been interesting to use Berent et al.'s four main clusters /bl/, /bn/, bd/, and /lb/, but this was not possible. First of all, remember that targets in initial or final position resulted in different response latencies in McQueen's experiments. This means that target words in initial and final position may not be compared, and have to be in the same position across its different contexts, making it impossible to include the most marked cluster /lb/. Moreover, since Dutch has final devoicing no words exist that phonetically end in /b/. Furthermore, in order to provide a clear picture on the influence of markedness of the cluster in a word-spotting experiment, more different types of clusters have to be examined. These issues cause me to divert from Berent et al.'s four types of clusters. Instead, each target word was put in three different contexts, in which the target always appeared in final position. The final consonant of the preceding context (C1) and the first consonant of the target (C2) provided the medial cluster of interest. By keeping the C2 constant across three different C1's, three different types of clusters were created with different markedness values. This was done for three categories of target words, each category with a different C2, creating a total of nine different types of clusters. The exact specification of types and examples are given below in (5). The sonority differences between the members of the cluster are according to Berent et al.'s sonority hierarchy, and the examples are written in standard Dutch orthography. For a complete overview of the materials used, see the appendix. Per category the first cluster is the least marked, the last cluster the most marked:

(5)

P-words: C1 is either a /t/, /m/, or /r/, while C2 is always a /p/			
Plosive-plosive	/tp/	Plateau (1→1)	Kaatpoot
Nasal-plosive	/mp/	Small fall (3→1)	Toempoot
Liquid-plosive	/rp/	Large fall (4→1)	Marpoot

N-words: C1 is either a /t/, /m/, or /r/, while C2 is always a /n/			
Plosive-nasal	/tn/	Small rise (1→3)	Dietnoot
Nasal-nasal	/mn/	Plateau (3→3)	Deemnoot
Liquid-nasal	/rn/	Very small fall (4→3)	Larnoot

L-words: C1 is either a /t/, /m/, or /r/, while C2 is always a /l/			
Plosive-liquid	/tl/	Large rise (1→4)	Kotlood
Nasal-liquid	/ml/	Very small rise (3→4)	Doomlood
Liquid-liquid	/rl/	Plateau (4→4)	Werlood

If the markedness of a cluster plays a role in the speed of its segmentation, then *poot* in *toempoot* should be found faster than *poot* in *kaatpoot*, and *poot* in *marpoot* should be found faster still. Each category holds five Dutch words, creating a total of 45 targets. An additional 45 items were created in which the preceding context and the medial cluster were the same, but the final syllable was a non-word (e.g. *kaatpoer*). The 60 filler items (30 with a Dutch word and 30 with a non-word) held no medial cluster to distract the participants' attention from clusters (e.g. *kuuhond* and *kuuhors*). Lastly, the 10 practice items (5 with a Dutch word and 5 with a non-word) either held a medial cluster or not. The ones with a medial cluster did not contain the same cluster as the targets did, and neither practice item appeared in the actual experiment. This brings the total number of items to 160.

All of the clusters in (5) are illegal onset clusters, while the clusters /mp/, /rp/, and /rn/ are legal coda clusters in Dutch. However, remember that CVCC.VC syllabifications are unlikely to occur, because languages prefer their syllables to have an onset. This caused no target to be misaligned with a phonotactic boundary, allowing the listeners to spot the target words on the basis of their phonotactic knowledge of Dutch. This knowledge will tell the listeners that they have to put a syllable boundary in between the C1 and C2, because putting a syllable boundary before C1 and C2 will cause the second syllable to contain an illicit onset cluster. The different degrees of markedness of this illegal onset cluster may cause the difference between the different contexts.

The construction of the contexts was done very carefully. First of all, naturally no context forms a Dutch word. Additionally the frequency of the CV sequence of the context was taken into account. Each possible Dutch CV sequence was put in a database containing 114992 Dutch words. The database next provided information on how many of the 114992 words start with each particular CV sequence. For example, of the 114992 words, 284 start with *taa*, while 647 start with *tee* (Dutch orthography). For each target word the frequency of the contexts was kept relatively the same. For example, the frequencies of the contexts of *kaatpoot*, *toempoot*, and *marpoot* are 702, 756, and 751 respectively. This was done because several studies have indicated that the frequency of the contexts influences the reaction times of the listeners (Norris et al., 1995; Vroomen & de Gelder, 1995). In order to be able to compare the reaction times between these three items the lexical frequency needs to be controlled. The highest possible frequencies were used in the construction of the contexts of the targets. The contexts of the filler items were all different from the contexts used for the targets. Since they only served as fillers, the frequency of the CV sequence was not taken into account for these items. A complete overview of the frequencies of the contexts of the targets is provided in the appendix.

Each item was recorded by a female native Dutch speaker in a sound attenuated booth, onto DAT tape, sampling at 48,000 Hz, using the Audacity software. The recordings were digitally stored as a wav-file. The speaker was instructed to put every item in a carrier sentence of her choice, and to try to keep the intonation the same across all 160 recordings. She always put main stress on the initial syllable, which is natural in Dutch. She was likewise instructed to carefully articulate the medial cluster. Whenever the C1 was an /r/ she was asked to try to pronounce it with a trill, in which she was very successful. Most Dutch speakers have a trilling /r/ in onsets, but an approximant in codas, though some speakers also have a trilling /r/ in codas. By making the speaker pronounce a coda /r/ with a trill it created a clearer articulation of the cluster, and made it more likely to be a part of an onset. Next the items were excised from their carrier sentence using the Praat software (Boersma & Weenink, 2001). It was attempted to keep the duration of the excised items approximately the same, especially those containing the same target word. This caused some slight modifications to be necessary for some of the items, for example by either cutting away or adding several periods in the vowels. Whenever an item was modified it had to remain sounding natural. Sounding natural was more important than a roughly equal duration.

Lastly, the items were programmed using the FEP experiment program (Veenker, 1998). This program allows the user to program two lists. One practice list, of which the user has to determine the order of appearance of the items, and one experimental list, which the program randomizes each time the application is started.

6.1.2 Participants and procedure

The 20 participants were all volunteers from the participant database of the University of Utrecht. At the end of the experiment they were paid a small amount for their participation. All of the participants were native speakers of Dutch and reported to have no hearing disabilities. There were 3 male and 17 female participants, with age ranging from 18 to 47, and with a mean age of 22.85. The participants were tested individually.

After the participants had arrived they were directed to a sound attenuated booth for testing. The participants were told that they were about to listen to disyllabic words of which one syllable was sometimes a Dutch word. Whenever they heard a Dutch word they had to press a button with their dominant hand as fast as they could and say out loud which word they had heard. If they had not heard a Dutch word they had to do nothing. The instructions were provided on paper and the participants had the chance to ask any questions after reading them. They were also told that they had the chance to ask additional questions after the practice session.

The experimenter was seated next to the participant with a checklist to note down any mistakes that were made. In order for the experimenter to know when a mistake was made, she had to be able to hear the items. The items were presented over speakers for this reason. The items were played at a clear, but comfortable volume. No participant reported that the sound was either too soft or too loud. The participants could keep track of their progress on the computer screen, which counted down from 150 to 0 during the course of the experiment. They received no additional feedback during the experiment.

Six of the participants failed to read the instructions well and asked no questions, assuming they understood what was required of them. Still, they failed to press the button even once during the practice list. After the list had ended the experimenter explained the procedure again and told the participant that 5 Dutch words were imbedded in the practice list and provided one example. Afterwards, the practice list was done again, during which these participant performed as was required of them. After doing the practice list twice, they moved on to the experimental list.

6.1.3 Results and discussion

The reaction times (RTs) were measured from target onset; from the beginning of the second syllable. The duration of the second syllable was subtracted from the measured RT to provide the pure RT. Missed responses and responses accompanied either by no oral response or by a word other than the intended word were treated as errors (7.6% of the targets). Additionally, RTs faster or slower than 2.5 SD of the average of that particular target were likewise treated as errors and removed from any further analysis (a further 2.3% of the data). If the missed responses and the removed responses yielded at least 7 (35%) results of a particular item to be excluded from the analysis the entire item was removed. Since the analysis will involve the comparison of the RTs within one particular target word across its three different contexts, the other two items had to be likewise excluded from any further analysis if one item was removed. For example, 8 participants failed to spot *park* in *saarpark*, which means that *saarpark* is excluded from the analysis. This resulted in excluding *neetpark* and *kumpark* as well, even though the response accuracy for these two items was very good. Asides the *park* items, the *lood* items were removed as well. All remaining items yielded accurate response accuracy, and were included in the analysis. The logarithm of each RT of the remaining items was calculated to create a more normally distributed set.

Remember that all the /tC/ targets were the least marked, the /mC/ targets slightly more marked, and the /rC/ targets the most marked. If the markedness of a cluster plays a role in the speed of its segmentation, there should be a general effect in which the RTs of /tC/ > /mC/ > /rC/. The graph in figure 3 shows that this effect is not present:

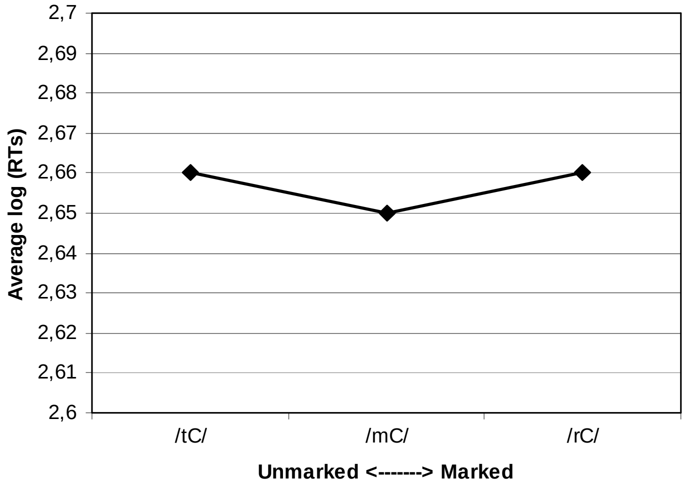


Fig 3. Overall response speed across different types of contexts.

As expected the average RT of /tC/ > /mC/, but the RT of /rC/ is slower than /mC/ instead of faster. Moreover, the differences in RTs are not large. A one-way Analysis of Variance (ANOVA) likewise revealed no significant effect ($F(2, 17) = 0.82$, $p = 0.45$).

A problem with merely analysing the overall RTs across different types of contexts is that within one context the markedness of its clusters is different. Within /tC/, the /tl/ cluster is a large rise, which is lesser marked than the /tn/ cluster, which is a small rise. The different categories need to be pulled apart to provide a clearer picture of what is going on. This is provided in figure 4:

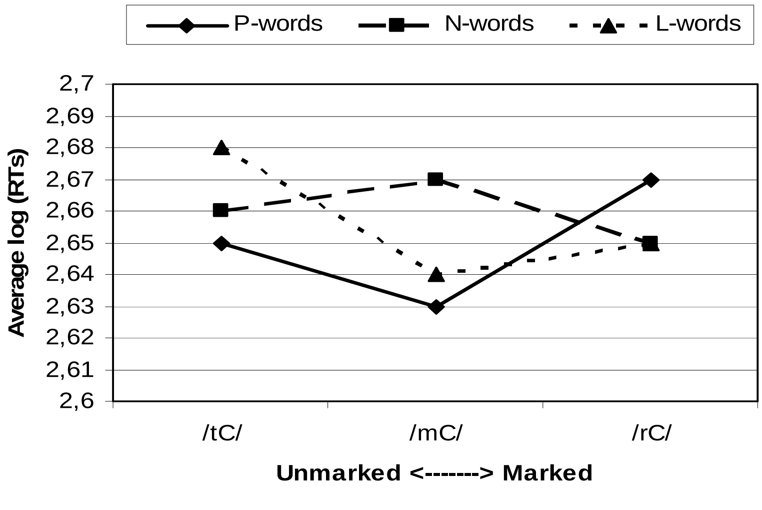


Fig 4. Overall response speed across different types of contexts for each separate category.

The P-words and the L-words show the same trend as the average in figure 3 by showing that the average RT of /tC/ > /mC/, but the RT of /rC/ is slower than /mC/ instead of faster. For the P-words the RT of the /rC/ targets are even slower than of the /tC/ targets. The N-words show the opposite trend. The RT of /mC/ > /rC/, but the RT of /tC/ is faster instead of slower than

/mC/. A one-way ANOVA again revealed no significant effect for either category (P-words: $F(2, 1) = 0.39$, $p = 0.69$, N-words: $F(2, 2) = 0.08$, $p = 0.92$, L-words: $F(2, 1) = 0.57$, $p = 0.59$).

Finally, the data was pulled apart even further to purely look at how the participants responded to different degrees in markedness. Although per category the /tC/ cluster was lesser marked than the /mC/ cluster, the overall markedness of the clusters was not distributed this way. The /tp/ cluster is a sonority plateau, which is more marked than the /ml/ cluster, which holds a very small rise in sonority. Figure 5 displays the general trend of markedness on the RTs:

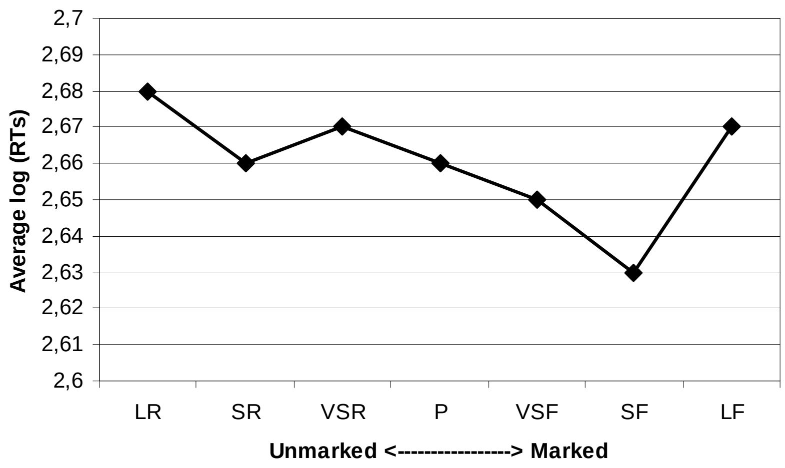


Fig 5. Overall response speed across different degrees of markedness. LR = large rise, SR = small rise, VSR = very small rise, P = plateau, VSF = very small fall, SF = small fall, LF = large fall.

Remember that there was only one type of cluster for each degree of markedness, except for the plateau, of which there were three clusters. The average provided in figure 5 for the plateau is the average RT of the three separate clusters. It was possible to do this, because a one-way ANOVA revealed no significant differences between the three kinds of plateaus ($F(2, 10) = 0.11$, $p = 0.90$). Interestingly, figure 5 suggests that there is at least a trend in which the markedness of the cluster affects the speed of its segmentation.

The average RT of the large fall poses an intriguing exception. This was the /rp/ cluster, which, as stated earlier, is a legal coda cluster in Dutch. However, the very small fall /rn/ and the small fall /mp/ are also legal coda clusters in Dutch, and they follow the general downward curve of figure 5. It is possible that /rp/ is a very bad onset but therefore also a very good coda cluster, causing this particular cluster to be segmented with a syllable boundary following the cluster (e.g. *marp.oot*). If this was the case, than the phonotactic misalignment of the targets containing the /rp/ cluster should result in higher error rates, analogous to McQueen’s findings. However, of the total number of missed targets only 1.6% is accounted for by the /rp/ clusters, including *saarpark*, which was removed from the analysis. Moreover, note how this is approximately one-seventh of the 7.6% missed targets, and there are seven degrees of markedness in this experiment. The /rp/ cluster did not yield more erroneous responses. A final possibility is that the consideration of syllabifying this string as CVCC.VC was present, but not to a large enough extent to fail spotting the intended target. The consideration might only have caused the participants to doubt, resulting in a slower RT, possibly because /rp/ is a well-formed and frequent medial cluster in Dutch. Still, it is strange that this consideration did not occur with the /rn/ and /mp/ clusters.

6.2 Experiment 2

A word-spotting experiment is traditionally accompanied by a lexical decision task. This type of experiment serves as an excellent control to see whether the trends found in experiment 1 are due to the acoustic properties of the targets, or due to the phonotactic markedness of the clusters.

6.2.1 Materials

The same 160 items from experiment 1 were used for this experiment. The second syllables were all excised from their preceding contexts using the Praat software, but remained otherwise unadjusted. The excising proceeded by taking a careful look at the oscillograms and my own auditory judgements. The target words were relatively easy to excise from the /tC/ contexts, because they followed a plosive. The P-words were likewise easy to excise for the same reason. Separating two sonorants proved to be much harder. The duration of the excised parts was later used for experiment 1, by being able to determine the target onset and to adjust the measured RTs to the pure RTs.

These items were likewise programmed using the FEP experiment program, though in a different script. This script also contained the possibility to create two lists. One practice list, of which the user has to determine the order of appearance of the items, and one experimental list, which the program randomizes each time the application is started.

6.2.2 Participants and procedure

The 20 participants were all volunteers from the participant database of the University of Utrecht, and did not participate in experiment 1. At the end of the experiment they were paid a small amount for their participation. All of the participants were native speakers of Dutch and reported to have no hearing disabilities. There were 4 male and 16 female participants, with age ranging from 19 to 56, and with a mean age of 26.35. The participants were tested individually.

After the participants had arrived they were directed to a sound attenuated booth for testing. The participants were told that they were about to listen to words which would sometimes be a Dutch word and otherwise a non-word. They had to press the YES button with their dominant hand whenever they heard a Dutch word, and the NO button with their non-dominant hand when they heard a non-word as fast as they could. The instructions were provided on paper and the participants had the chance to ask any questions after reading them. They were also told that they had the chance to ask additional questions after the practice session.

In order to mimic the sound quality of experiment 1, the items were likewise presented over speakers. The items were played at a clear, but comfortable volume. No participant reported that the sound was either too soft or too loud. The participants could keep track of their progress on the computer screen, which provided the participants with immediate feedback. If they had pressed the correct button the word ‘correct’ would flash on the screen. Pressing the wrong button caused the message ‘fout’ (which means ‘wrong’) to appear. A late or no response elicited the message ‘te laat’ (which means ‘too late’). Outside of the booth, the experimenter could keep track of the participant’s progress by seeing the same messages flash across the screen.

Sadly, the results of one of the participants were unusable. She failed to press the buttons correctly, resulting in many unregistered responses. Her results have been excluded from the analysis and replaced by the results of a family member, who was unaware of the procedure and goal of the experiment.

6.2.3 Results and discussion

The reaction times (RTs) were measured from item onset. The duration of the item was subtracted from the measured RT to provide the pure RT. Missed (prompting the ‘te laat’ message) and incorrect (prompting the ‘fout’ message) responses were treated as errors (8.9% of the targets). Additionally, RTs faster or slower than 2.5 SD of the average of that particular target were likewise treated as errors and removed from any further analysis (a further 1.4% of the data). Since the *park* and the *lood* items were removed from the word-spotting experiment, they were likewise removed from any further analysis from this experiment. All remaining items yielded accurate response accuracy, and were included in the analysis. The logarithm of each RT of the remaining items was calculated to create a more normally distributed set.

The graphs of experiment 1 revealed that, though there seems to be a trend (figure 5), the markedness of the cluster may not play a large role in phonotactic speech segmentation. If this trend can be attributed to the markedness of the particular cluster, and not to any other acoustic cues, then the excised items should reveal no significant differences between the different types. This means that the RTs of /tC/ = /mC/ = /rC/. If significant differences do occur in experiment 2, then there might be a chance that markedness actually had a larger role in experiment 1, but that its effect was extinguished by other acoustic cues. The graph in figure 6 shows the overall response speed:

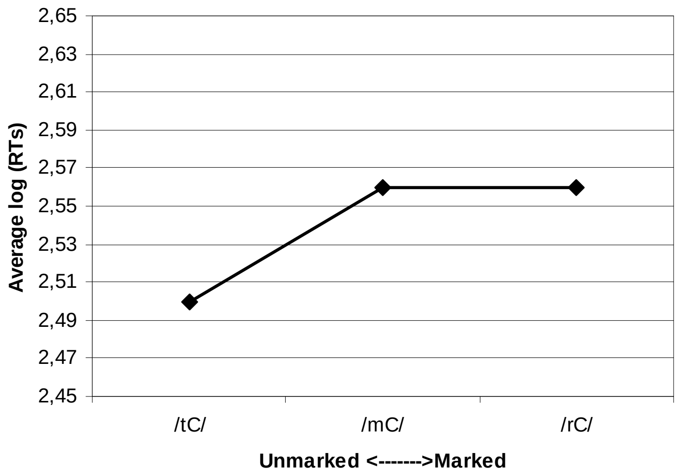


Fig 6. Overall response speed across different types of contexts.

The /mC/ and /rC/ types yielded the same average RT, which indicates that the trends of experiment 1 are not caused by acoustic cues other than the phonotactic markedness of the cluster. However, the items excised from /tC/ contexts were responded to much faster. A one-way ANOVA even revealed a significant effect ($F(2, 17) = 15.27$, p = <.0001).

It is important to pull this graph apart again into the three different categories to see whether this significant effect is category bounded. This is provided in figure 7:

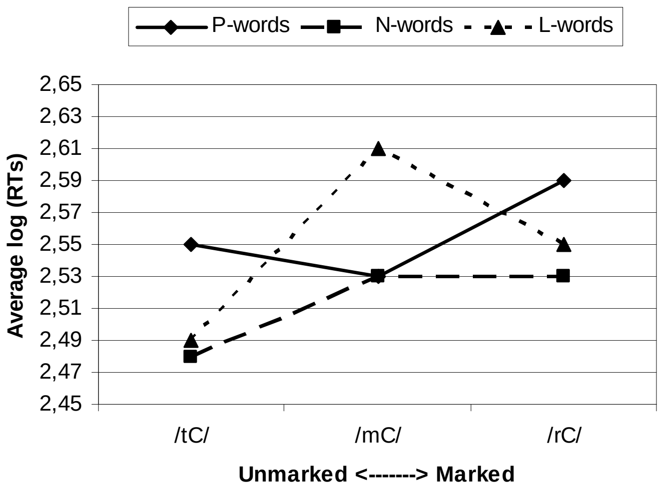


Fig 7. Overall response speed across different types of contexts for each separate category.

The N-words show the same trend as the average in figure 6 by showing that the average RT of /mC/ = /rC/, but the RT of /tC/ is faster. The L-words show a similar trend, though with much more variation. The graph of the P-words show a response pattern opposite to that of the N-words and L-words. Interestingly, a one-way ANOVA revealed a significant effect for the N-words and the L-words, but not for the P-words (P-words: $F(2, 1) = 0.76$, $p = 0.51$, N-words: $F(2, 2) = 7.44$, $p = 0.01$, L-words: $F(2, 1) = 8.55$, $p = 0.02$).

Finally, the data was pulled apart even further to purely look at how the participants responded to different degrees in markedness. If this final graph still reveals a downward curve it becomes very likely that the trend of figure 5 is due to acoustic cues of the different targets and not due to the markedness of the cluster. Figure 8 displays the general trend of markedness on the RTs:

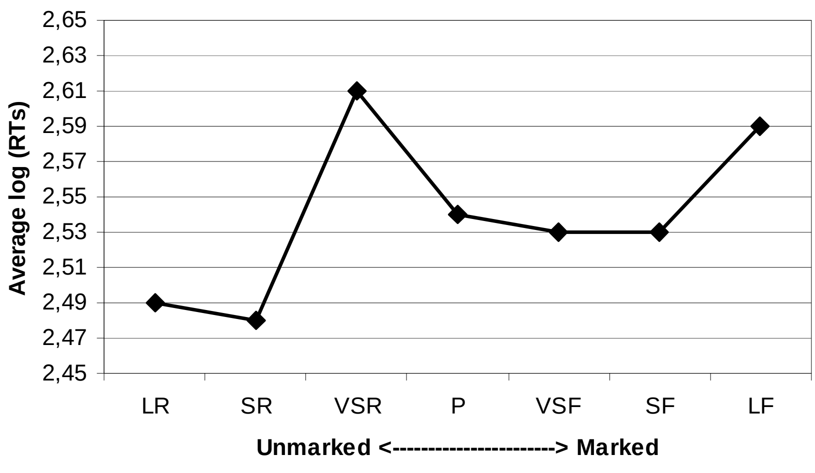


Fig 8. Overall response speed across different degrees of markedness. LR = large rise, SR = small rise, VSR = very small rise, P = plateau, VSF = very small fall, SF = small fall, LF = large fall.

In this graph the average for the plateau is likewise the average RT of the three separate clusters. It was again possible to do this, because a one-way ANOVA revealed no significant differences between the three kinds of plateaus ($F(2, 10) = 0.19$, $p = 0.83$). As becomes evident from this figure, there was hardly a trend of markedness, at least not in the direction expected. As a matter of fact, the clusters that should have yielded the slowest RTs in experiment 1 yielded the fastest RTs in this experiment. The very small rise cluster /ml/, and the large fall cluster /rp/ are the main cause for the significant effects.

What caused significant effects to appear where they were not present in experiment 1? Are there other acoustic cues in play? A possibility is the sound quality of the excised parts, specifically of the onset. The reason the /tC/ contexts yielded faster RTs might be because the target words were relatively easy to excise after a plosive. Maybe no significant interaction occurred within the P-words, because they were likewise easy to excise. Both situations provided target words with decent sound qualities. The quality of the onset of the targets separated from two sonorants might have been affected to such an extent that the participants experienced a moment of doubt. Several even reported this to be the case; that they needed a slight moment to realize what they had just heard was actually a word.

Another possibility has already been briefly discussed at the beginning of this section, namely that the markedness of a cluster does have an effect on the speed of its segmentation, but that the targets carried other acoustic cues, which might have cast a shadow over the effect of markedness. In any case, the overall results of this experiment differ enough to conclude that the trend shown in figure 5 is more likely due to the markedness of the clusters than due to other acoustic cues.

7. General discussion and conclusion

This thesis attempted to combine two separate findings. The first finding is that listeners use their language-specific phonotactic constraints as cues in speech segmentation, and that this information is compatible with the lexical competition model Shortlist (McQueen, 1998). The second finding is that listeners are sensitive to the degree of markedness in onset clusters when these clusters are illegal in their native language (Berent et al., 2007). These two findings led to the first research question, namely whether existing words are easier to find in a highly marked context than in a lesser marked context.

In addition to reviewing Berent et al.'s article, this thesis provided a detailed typological analysis of their onset clusters using Optimality Theory. This theory states that the grammar of a language is a language-specific ranking of universal constraints. If the underlying form or the input does not comply with the grammar of the listener, because markedness constraints are ranked high, an unmarked alternative that does comply with the grammar of the listener will become the winner, causing a perceptual illusion. This is in accordance with the phonological repair strategy found by Berent et al. However, the only way to explain the sensitivity to the markedness of the clusters that the English native speakers of Berent et al.'s study experienced, the ranking of the MARK of my OT analysis may not be free. This gave rise to the second research question, namely whether the order of the constraints is set or free when the specific ranking between constraints does not matter for the outcome of the evaluations.

A word-spotting experiment was conducted to find an answer to both questions. If the markedness of the cluster acts as an additional cue to the segmentation process, then highly marked clusters should signal a stronger syllable boundary and be segmented faster than lesser marked clusters, resulting in faster reaction times. If this effect was found it provides evidence that the order of the MARK is set and not free.

The initial results showed no consistent sensitivity to the different degrees of markedness the different contexts provided for the target words. However, when ordering the clusters according to their different markedness values, there seems to be a gradual decline in RTs when the clusters become more marked. Even though comparing different target onsets is technically impossible, this is an interesting bit of information nonetheless. There are four logical explanations for the trends found in this experiment. The first one is that markedness plays no role at all in the segmentation process. The second one is that markedness of the cluster plays a role in segmentation, but not on the speed. The gradual decline in RTs shown in figure 5 is bases on coincidence for these two options. The third one is that markedness does play a role and affect segmentation speed, but other acoustic cues of the targets made its role unable to be found. The last one is that markedness plays a role, but its role is too small to be able to be found.

There is a reason to believe that the strength of a phonotactic syllable boundary indeed has only a subordinate role in the segmentation process. Remember that, if segmentation proceeds by lexical competition, there is one central mechanism (Norris et al., 1997). This competition process can be modulated by additional information, such as metrical or phonotactic information, which signal where syllable and possible word-boundaries may occur. This type of information is subordinate to the central mechanism. The information the markedness of a consonant cluster provides is in turn subordinate to the overall phonotactic information. Whether a Dutch listener encounters the large rise /tl/ or the plateau /tp/, which are differently marked, this listener is either way forced to place a syllable boundary in between these two members. In essence, the supposed strength of the syllable boundary is unnecessary for the result of the segmentation process in the same way that the ranking of the MARK is irrelevant for the outcome of the evaluations as long as Dep-IO is ranked correctly. The only effect markedness can have is on the speed of the segmentation process. However, this process is already very fast to begin with. How much difference can a subordinate cue like markedness make? This line of thought supports the fourth option as to why no effects were found in experiment 1, only a possible trend.

A lexical decision experiment was conducted as a control. Interestingly enough, significant effects were found in this experiment. This may indicate that other acoustic properties may have shadowed the effect of markedness in the word-spotting experiment, supporting the third reason as to why no effects were found in experiment 1. A closer examination of the data reveals another possible reason as to why effects were found in experiment 2, namely that these effects might be due to the quality of the onset of the excised targets, rather than to other acoustic properties. Their graphs show a very different trend than the ones from experiment 1, which leads me to conclude that the trends found in experiment 1 are due to the different degrees of markedness of the clusters, and not due to other acoustic properties of the targets, again supporting the fourth reason as to why no effects were found in experiment 1.

My experiments have at least shown that something is going on, but since the results are inconsistent more research needs to be done on this topic to form any solid conclusions. It would be very interesting to perform this experiment among Japanese native speakers. Since

Japanese allows no consonant clusters to appear in neither the onset nor the coda all the clusters used above are fully illegal in Japanese (Dupoux et al., 1999), while the sonority falls were legal coda clusters in Dutch. Japanese listeners might therefore produce stronger trends, possibly resulting in a significant effect.

Additionally, remember that Berent et al. found that their Russian native speakers were not sensitive to the different degrees of markedness, presumably because all epenthetic counterparts are eliminated at the same stage of evaluation by Dep-IO. Repeating my experiment among Russian native speakers should yield the same trends and no significant effects, meaning that they should show a relative flat-line when the different clusters are ordered according to their different markedness values. It might be difficult to create illegal medial clusters in Russian, since it allows many types of clusters. To illustrate, of the nine clusters used in experiment 1, six clusters (/tp/, /mp/, /rp/, /tn/, /rn/, and /rl/) are illegal onset clusters in Russian, but only two are illegal coda clusters (/tp/ and /tn/); that is, these clusters are zero-frequent (Trapman & Kager, 2009).

Comparing the results of multiple similar experiments might shed more light on the role of markedness in speech segmentation. For now, its role is shown to be there, but not with a strong, solid and consistent piece of evidence.

References

Altenberg, E.P. (2005). “The judgement, perception, and production of consonant clusters in a second language.” *International Review of Applied Linguistics in Language Teaching* **43:** 53-80.

Anderson, S. R. (2005). *How many languages are there in the world?* The Linguistic Society of America, Washington, D.C. 28 Apr 2011 <http://www.lsadc.org/pdf_files/howmany.pdf>

Archangeli, D. (1997). “Optimality Theory: an Introduction to Linguistics in the 1990s.” In D. Archangeli and D.T. Langendoen (eds.), *Optimality Theory, an Overview*. Blackwell Publ., Oxford, 1-31.

Beckman, M.E. and J. Edwards. (1990). “Lengthenings and shortenings and the nature of prosodic constituency.” In J. Kingston and M.E. Beckman (eds.), *Papers in laboratory phonology I: Between grammar and physics of speech*. Cambridge: Cambridge University Press, 152-78.

Berent, I. et al. (2007). “What we know about what we have never heard: Evidence from perceptual illusions.” *Cognition* **104**: 591-630.

Blevins, J. (1995). “The syllable in phonological theory.” In J. Goldsmith (Ed.), *The handbook of phonological theory*. Cambridge: Blackwel, 206-44.

Boersma, P. and D. Weenink. (2001). “Praat, a system for doing phonetics by computer.” *Glot International.* **5 (9/10):** 341-5.

Broselow, E. and D. Finer. (1991). “Parameter setting in second language phonology and syntax.” *Second Language Research* **7:** 35-59.

Cutler, A. and D. Norris. (1988). “The role of strong syllables in segmentation for lexical access.” *Journal of Experimental Psychology: Human Perception and Performance* **24:** 381-410.

Dinnsen, D.A. (2007). “Fundamentals of Optimality Theory.” Chapter 1 of D.A. Dinssen and J.A. Gierut (eds.). *Optimality Theory, Phonological Acquisition and Disorders*. London: Equinox.

Dupoux, E. et al. (1999). “Epenthetic vowels in Japanese: A perceptual illusion?” *Journal of Experimental Psychology: Human Perception and Performance* **25:** 1568-78.

Dupoux, E. et al. (2001). “New evidence for prelexical phonological processing in word recognition.” *Language and Cognitive Processes* **5:** 491-505.

Escudero, P. and P. Boersma. (2004). “Bridging the gap between L2 speech perception research and phonological theory.” *Studies in Second Language Acquisition* **26**: 551-85.

Gnanadesikan, A. (2004). “Markedness and faithfulness constraints in child phonology.” In R. Kager, J. Pater and W. Zonneveld (eds.). *Constraints in Phonological Acquisition.* Cambridge University Press.

Gow, D.W. and P.C. Gordon. (1995). “Lexical and prelexical influences on word segmentation: Evidence from priming.” *Journal of Experimental Psychology: Human Perception and Performance* **21:** 344-59.

Greenberg, J. H. (1978). “Some generalizations concerning initial and final consonant clusters.” In E. A. Moravcsik (Ed.). *Universals of human language*. Stanford, CA: Stanford University Press, Vol. 2, 243–79.

Guasti, M.T. (2002). *Language Acquisition: The Growth of Grammar.* Cambridge: MIT Press.

Hallé, P. et al. (1998). “The processing of illegal consonant clusters: A case of perceptual assimilation?” *Journal of Experimental Phonology: Human Perception and Performance* **24:** 592-608.

Harms, R.T. (1973). “Some Nonrules of English”, paper distributed by Indiana University Linguistics Club, Bloomington.

Jäger, G. (2003). *Learning constraint sub-hierarchies: The Bidirectional Gradual Learning Algorithm.* Manuscript, Potsdam University.
Jusczyk, P.W. et al. (1993). "Infants' sensitivity to the sound patterns of native language words." *Journal of Memory and Language* **32:** 402-20.
Kaye, J. (1989). *Phonology, a Cognitive View*. Hillsdale, New Jersey: Lawrence Erlbaum Associates Inc.
Klatt, D.H. (1975). "Vowel lengthening is syntactically determined in a connected discourse." *Journal of the Acoustical Society of America* **59:** 1208-21.
Lehiste, I. (1960). "An acoustic-phonetic study of internal open juncture." *Phonetica* **Supplement 5:** 1-54.
Lombardi, L. (1995). "Laryngeal Neutralization and Syllable Wellformedness." *Natural Language and Linguistic Theory* **13:** 39-74.
Lombardi, L. (1999). "Positional Faithfulness and Voicing Assimilation in Optimality Theory." *Natural Language and Linguistic Theory* **17:** 267-302.
Massaro, D.W. and M.M. Cohen. (1983). "Phonological constraints in speech perception." *Perception & Psychophysics* **34:** 338-48.
McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry* **17**: 207-263.
McCarthy, J.J. and A.S. Prince (1993). "Generalized Alignment." *Morphology Yearbook* **3**: 79-153.
McQueen, J.M. et al. (1994). "Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20:** 621-38.
McQueen, J.M. (1998). "Segmentation of Continuous Speech Using Phonotactics." *Journal of Memory and Language* **39:** 21-46.
Moreton, E. (2002). "Structural constraints in the perception of English stop-sonorant clusters." *Cognition* **84:** 55-71.
Norris, D.G. (1994). "Shortlist: A connectionist model of continuous speech recognition." *Cognition* **52:** 189-234.
Norris, D.G. et al. (1995). "Competition and segmentation in spoken word recognition." *Journal of Experimental Psychology: Learning, Memory, and Cognition* **21:** 1209-28.
Norris, D.G. et al. (1997). "The possible-word constraint in the segmentation of continuous speech." *Cognitive Psychology* **34:** 191-243.
Ostapenko, O. (2005). "The Optimal L2 Russian SyllableOnset." *LSO Working Papers in Linguistics* **5:** 140-50.
Pertz, D.L. and T.G. Bever. (1975). "Sensitivity to phonological universals in children and adolescents." *Language* **51:** 149-62.
Pitt, M.A. and A.G. Samuel. (1995). "Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology* **29:** 149-88.
Pitt, M.A. (1998). "Phonological processes and the perception of phonotactically illegal consonant clusters." *Perception & Psychophysics* **60:** 941-51.
Prince, A.S. and P. Smolensky. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell pub.
Rietveld, A.C.M. and V.J. van Heuven. (1997/2001). *Algemene fonetiek*. Bussum: Coutinho.
Selkirk, E. (1984). "On the Major Class Features and Syllable Theory." In M. Aronoff & R.T. Oehrle (eds.). *Language Sound Structure.* 107-36.
Slowiaczek, L.M. and M. Hamburger. (1992). "Prelexical facilitation and lexical interference in auditory word recognition." *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18:** 1239-50.

Strange, W. (1995). “Cross-language studies of speech perception: a historical review.” In C.A. Ferguson, L Menn & C. Stoel-Gammon (eds.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research.* Timonium, MD: New York Press Inc. 3-45.

Trapman, M. and R. Kager. (2009). “The Acquisition of Subset and Superset Phonotactic Knowledge in a second language.” *Language Acquisition* **16:** 178-221.

Trommelen, M. (1983). *The Syllable in Dutch, with Special Reference to Diminutive Formation*. PhD dissertation, Utrecht University.

Veenker, T.J.G. (1998). *Flexible Experiment Programme: A tool for designing and running computerized experiments.* Utrecht Institute of Linguistics – OTS.

Vroomen, J. and B. de Gelder. (1995). “Metrical segmentation and lexical inhibition in spoken word recognition.” *Journal of Experimental Psychology: Human Perception and Performance* **21:** 98-108.

Wallace, W.P., M.T. Stewart and C.P. Malone. (1995). “Recognition memory errors produced by implicit activation of word candidates during the processing of spoken words.” *Journal of Memory and Language* **34:** 417-39.

Wallace, W.P. et al. (1995). “False positives in recognition memory produced by cohort activation.” *Cognition* **55:** 85-113.

Zonneveld, W. (1996). “Optimality Theory, roots and aims.” *South African Journal of Linguistics* **Supplement 31:** 23-54.

Appendix

Targets (45)					
Item		**Gloss**	**Frequency of CV sequence of the context**	**Total duration (ms)**	**Duration target (ms)**
Context	**Target**				
Kaat	Poot	Paw	702	527	240
Toem			756	514	295
Mar			751	507	255
Paat	Pols	Wrist	500	586	303
Maam			524	578	295
Ror			494	567	308
Neet	Park	Park	410	547	272
Kum			393	522	303
Saar			412	568	260
Toot	Poes	Puss	307	527	280
Poom			317	527	289
Ser			315	519	243
Lut	Pil	Pill	227	501	272
Laam			235	539	266
Rar			268	523	250
Diet	Noot	Nut	537	520	327
Deem			565	540	303
Lar			558	535	313
Ket	Nier	Kidney	476	482	269
Mim			481	516	328
Rer			474	514	267
Dot	Naam	Name	353	568	332
Pom			386	554	332
Mer			391	578	334
Taat	Nest	Nest	284	564	337
Beem			290	567	330
Der			305	585	344
But	Nacht	Night	211	559	336
Piem			220	524	326
Gar			244	551	311
Kot	Lood	Lead	1748	502	311
Doom			962	515	290
Wer			1281	513	327
Tet	Leuk	Nice/fun	411	539	326
Moom			408	563	271
Dir			421	558	318
Miet	Lied	Song	312	514	294
Siem			331	533	274
Mor			317	527	288
Geet	Lamp	Lamp	245	562	292
Noom			268	570	298
Ber			275	562	332
Sut	Lens	Lens	205	525	287
Biem			207	513	286
Tar			242	526	325

Non-words (45)			
Item		Total duration (ms)	Duration target (ms)
Context	Target		
Kaat	Poer	507	240
Toem		503	240
Mar		516	240
Paat	Pelk	531	250
Maam		550	250
Ror		545	250
Neet	Pulg	569	256
Kum		551	301
Saar		574	250
Toot	Pig	544	271
Poom		529	278
Ser		552	270
Lut	Peup	538	273
Laam		558	287
Rar		563	289
Diet	Noes	505	292
Deem		520	306
Lar		532	312
Ket	Nems	554	356
Mim		557	345
Rer		576	331
Dot	Nolk	536	341
Pom		524	341
Mer		547	332
Taat	Nig	535	286
Beem		526	307
Der		519	308
But	Neuf	570	390
Piem		573	385
Gar		579	361
Kot	Loeg	503	285
Doom		517	267
Wer		525	298
Tet	Lerk	535	348
Moom		545	295
Dir		530	339
Miet	Lup	501	282
Siem		516	258
Mor		513	268
Geet	Lif	554	253
Noom		558	281
Ber		553	319
Sut	Leup	524	281
Biem		514	310
Tar		516	319

Practice (10)		
Item		Gloss
Context	Target	
Fiet	Kant	Side
Wam	Reus	Giant
Ler	Merk	Brand
See	Pot	Pot
Muu	Lip	Lip
Fuup	Kert	
Won	Rost	
Loel	Musk	
Soo	Purk	
Muu	lups	

Practice (10)			
Item		Total duration (ms)	Duration target (ms)
Context	Target		
Fiet	Kant	566	322
Wam	Reus	558	325
Ler	Merk	549	309
See	Pot	536	272
Muu	Lip	504	296
Fuup	Kert	552	290
Won	Rost	535	322
Loel	Musk	538	289
Soo	Purk	539	257
Muu	lups	523	327

Fillers (60)									
Item		Gloss	Total duration (ms)	Duration target (ms)	Item		Gloss	Total duration (ms)	Duration target (ms)
Context	Target				Context	Target			
Nie	Band	Bond	504	318	Roo	Wiel	Wheel	517	297
	Burk		511	325		Wef		537	325
Seu	Boef	Crook	572	304	Raa	Kast	Cupboard	552	322
	Berf		558	291		Kuig		560	321
Voe	Mast	Mast	523	333	Fe	Raaf	Raven	522	358
	Mef		530	345		Rarf		517	379
Gij	Feest	Party	583	316	Heu	Tas	Bag	541	317
	Fum		527	283		Torp		559	316
Baa	Feit	Fact	560	315	Kaa	Raam	Window	505	302
	Felk		556	318		Rolg		521	298
Gee	Bank	Bank/couch	560	321	Duu	Maag	Stomach	503	344
	Bis		546	317		Mijp		497	347
Mui	Duif	Pigeon	561	309	Veu	Boek	Book	507	293
	Duk		558	315		Bop		512	285
Foe	Riet	Reed	539	337	Puu	Muis	Mouse	508	346
	Rops		520	348		Muip		501	324
Zaa	Deur	Door	511	254	Pui	Doek	Cloth	506	306
	Darf		529	285		Dens		512	323
Kui	Kat	Cat	509	285	Pie	Teen	Toe	511	338
	Koes		508	288		Tern		513	342
Peu	Film	Film/movie	532	330	Fij	Geur	Smell	532	307
	Fon		521	334		Gep		517	279
Zuu	Buik	Belly	533	315	Fie	Rok	Skirt	548	317
	Beg		532	312		Rolp		546	336
Soe	Mond	Mouth	533	324	Kee	Ramp	Disaster	509	298
	Meuk		558	328		Rulp		527	315
Ne	Kerk	Church	529	306	Fui	Ton	Tub	527	231
	Kuin		517	296		Tier		527	267
Kuu	Hond	Dog	501	317	Hui	Hout	Wood	554	329
	Hors		507	322		Hogt		571	330