0 Gradience and Categoricality in Phonological Theory*

MIRJAM ERNESTUS

1 Introduction

Within phonological theory, important roles are assigned to the notions of 'gradience' and 'categoricality'. The opposition qualifies sounds and sound patterns, and is crucial both for the definition of the phonological and the phonetic components of generative grammar, and for the development of alternative types of grammatical models. This chapter discusses the assumptions generative phonology and its direct successors (including Optimality Theory) have made about the role of gradience. Moreover, it presents data supporting or contradicting these assumptions, and discusses new models accounting for the conflicting data.

The most important section of this chapter (§2) discusses the opposition between categorical sounds, which are stable and represent clear distinct phonological categories (e.g. sounds showing all characteristics of voiced segments throughout their realizations), and gradient sounds, which may change during their realizations and may simultaneously represent different phonological categories (e.g. sounds that start as voiced and end as voiceless). A shorter section (§3) discusses categorical generalizations over sounds, which are fully productive, and gradient generalizations, which are less productive. The final section (§4) provides a short conclusion.

2 Sounds

2.1 Gradience in generative grammar

In the early days of generative grammar, the opposition between categoricality and gradience was assumed to reflect the fundamental distinction between competence and performance. Competence described speakers' categorical knowledge about their language, abstracted away from performance factors such as vocal tract size, working memory span, articulatory effort, and so on. Performance, in contrast, described speakers' actual linguistic behavior, which could be gradient, and was not in the direct focus of linguistic research (Chomsky & Halle 1968; following Saussure 1916).

The distinction between competence and performance was reflected in the distinction between the phonological and phonetic component. The phonological component contained the speaker's competence and thus represented cognition. It was believed to be language-specific and to include the phonemes of the speaker's language and language-specific phonological processes, such as final devoicing (Chapter 47: Final Devoicing and Final Laryngeal Neutralization) and place assimilation (Chapter 85: Local Assimilation). This knowledge was represented in the form of categorical symbols and rules operating on these symbols. Phonetic mechanisms were responsible for the speaker's performance. These phonetic mechanisms were believed to be universal and the automatic results of speech physiology (Chomsky & Halle 1968: 293; Kenstowicz & Kisseberth 1979). They thus reflected physics and included, for instance, nasalization of vowels preceding nasal consonants, palatalization of consonants preceding high vowels and shortening of vowels preceding voiceless obstruents. Since the phonetic component did not reflect the speaker's competence, it was considered not to be part of the grammar proper.

This view on the phonological and phonetic components changed very quickly, since various studies showed that the exact realization of an abstract symbol (e.g. a phoneme or a phonological feature) may be different in different languages. Moreover, no part of a realization appeared to be the automatic and unavoidable result of speech physiology (e.g. Keating 1985, 1990a; Kingston & Diehl 1994; see also Chapter 15: Distinctive Features). As a consequence, the traditional definition of the phonetic component as containing only universal processes automatically resulting from speech physiology implied that this component was empty. A new distinction had to be developed, which was no longer based on the notions of language-specific *vs.* language-universal and non-automatic *vs.* automatic mechanisms.

The now widely accepted definitions of the phonological and phonetic components are completely based on the opposition between categoricality and gradience (e.g. Keating 1988, 1990b; Pierrehumbert 1990; Cohn 1993: Zsiga 1997). The phonological component is assumed to deal with categorical, abstract, stable, timeless symbols, such as phonemes and phonological features. Phonological processes refer to these symbols and consequently have categorical effects: they change one symbol (e.g. [+voice]) into another one ([–voice]), or they delete or insert symbols. Phonetic processes translate the abstract symbols into articulatory and perceptual targets. This may lead to sounds with acoustic characteristics that do not perfectly represent categorical phonological symbols, but rather have intermediate values, for instance, when obstruents are partly voiced due to coarticulation. These definitions of the phonological and phonetic components have been adopted in several psycholinguistic models of speech production and comprehension (e.g. Levelt 1989; Norris 1994).

Since the distinction between gradience and categoricality is crucial in the definitions of the phonological and phonetic components, it has led to many experimental studies (Chapter 18: The Atoms of Phonological Representation; Chapter

101: Experimental Approaches in Theoretical Phonology). The following subsections discuss their findings and their implications for phonological theory. The first subsections discuss the domains (assimilation and segment deletion) where the evidence for gradience is most convincing but can also be relatively easily reconciled with generative grammar: the relevant processes traditionally characterized as phonological could be reclassified as phonetic. The following subsections (on incomplete neutralization and phonetic detail) discuss evidence for gradience that is less clear but has important theoretical consequences. Generative grammar cannot account for incomplete neutralization without making additional far-reaching assumptions. Further, the evidence for a role for fine phonetic detail in speech processing suggests that words are not lexically represented in the form of abstract phonemes but are stored together with their detailed phonetic properties. These data have stimulated the development of accounts based on assumptions other than those of generative grammar.

2.2 Assimilation: Data

One of the first types of processes traditionally characterized as phonological for which researchers found evidence of gradience is formed by connected speech processes, in particular assimilation (Chapter 85: Local Assimilation). Nearly all instances of assimilation are traditionally described as the categorical spreading of a phonological feature from one segment to another segment in the phonological component. The receiving segment is assumed to be subsequently identical to segments with the same features in their underlying specifications. For instance, [m] would have exactly the same surface phonological representation and phonetic characteristics if it results from an underlying /m/ and if it results from place assimilation, as in the phrase *gree*[m b]*oat* 'green boat'.

Many articulatory studies have investigated the assumed categorical nature of place assimilation using electropalatography (EPG; Hardcastle 1972), which registers contacts between the tongue and the hard palate, or with the help of an electromagnetic midsagittal articulometer (EMMA; e.g. Perkell *et al.* 1992), which allows the tracking of individual fleshpoints by means of small transducer coils attached to various points on the speaker's vocal tract in the midsagittal plane. These studies have provided evidence for the categorical nature of some place assimilation processes. An example is regressive place assimilation in Korean, which is a characteristic of fast colloquial Korean and affects certain consonants preceding certain other consonants. For instance, the phrase /pat¬õp'oda/ 'rather than the field' can be pronounced as [papõp'oda]. Kochetov & Pouplier (2008) showed that this assimilation results in the categorical absence of the gestures for the original articulation place of the assimilated consonant (in this example: for /t¬/) in most tokens. Another example is place assimilation of /n/ to /k/ in Italian, which categorically results in the absence of alveolar gestures (Farnetani & Busa 1994).

Other studies strongly suggest that some place assimilation processes are gradient in nature. For instance, assimilation of alveolar obstruents to the palatality of the following segments (as in American English *hi*/t j/*ou*) often does not lead to completely palatal segments ([c] in the example), but rather to segments that become more palatal during their realizations (within one and the same token) and that consequently differ in their phonetic detail from underlying palatals (e.g. Barry 1992 for Russian; Zsiga 1995 for post-lexical palatalization in American English). The same type of gradience has been reported for place assimilation of coronal obstruents in American English, as in *la*/t k/*alls* (*late calls*) produced as *la*[k…]*alls*. The assimilated obstruents often start with a coronal constriction that gradually assimilates to the articulation place of the following obstruent during their realizations (velar in the above example; Nolan 1992). Other gradient place assimilation processes include assimilation of alveolar nasals in American English (e.g. in *gree*[m b]*oat*; Ellis & Hardcastle 2002) and of /n/ to following post-alveolars in Italian (Farnetani & Busa 1994). Interestingly, some of these assimilation processes show considerable inter-speaker and intra-speaker variation. For instance, Ellis & Hardcastle (2002) found that four of their eight English speakers showed categorical place assimilation of /n/ to following velars in all tokens, two speakers showed either no or categorical assimilation and two speakers showed gradient assimilation. Together, the data show that place assimilation processes, at least those applying across morpheme boundaries, may be gradient in nature. These processes cannot simply be accounted for by the categorical spreading of a phonological feature from one segment to another.

The evidence for gradience is clearer for place assimilation than for voice assimilation. The main reason is probably that the difference between [+voice] and [–voice] obstruents is cued by many different acoustic, and hence also articulatory, characteristics, including the duration of the preceding vowel, the duration and intensity of the obstruent, and the duration of glottal vibration during the obstruent. Voice assimilation can thus not be studied only on the basis of electropalatography, and has been mainly investigated on the basis of the acoustic signal instead. For instance, Kuzla *et al.* (2007) studied progressive voice assimilation in German clusters consisting of a voiceless obstruent and a voiced fricative (e.g. the /tv/ cluster in *ha*/t v/*älder* 'had woods' produced as [tf]). They showed that assimilation results in shorter stretches of glottal vibration during the cluster, whereas it hardly affects the duration of the fricative, which is the most important perceptual cue to the [±voice] distinction for German fricatives. Assimilation thus does not affect all perceptual cues of the [±voice] distinction equally, and the phonetic implementation of devoiced fricatives differs from the implementation of underlyingly voiceless fricatives. This is difficult to reconcile with an abstract phonological categorical account of voice assimilation, since in such an account voice assimilation results in phonologically voiceless fricatives, which cannot be distinguished from underlyingly voiceless fricatives during phonetic implementation.

Other studies have investigated regressive voice assimilation in Dutch, that is, the voiced realizations of obstruents before voiced stops (e.g. *we*[t] 'law' is realized as *we*[d] in *wetboek* 'law book'). Ernestus *et al.* (2006) and Jansen (2007) showed that glottal vibration, which is the most important cue to the [±voice] distinction in Dutch obstruent clusters (van den Berg 1988), may be completely absent, partly present or continuously present in clusters subject to regressive voice assimilation, suggesting that regressive voice assimilation in this language is gradient. Ernestus and colleagues (2006) also investigated the effect of a word's frequency of occurrence (i.e. the word's relative number of occurrences in speech, independent of its realization) on voice assimilation (see also Chapter 000: Frequency Effects). They found that higher frequencies correlate with shorter obstruent clusters, a perceptual cue for [+voice], but also with shorter periods of glottal vibration and longer release bursts, which are perceptual cues for [–voice]. These data also suggest that voice assimilation may result in sounds that are neither fully voiced nor fully voiceless.

In conclusion, the data on assimilation suggest that we often perceive assimilation as categorical because we are used to distinguishing between only two values of the relevant phonological feature, but that the actual results from assimilation may be gradient rather than categorical. Before discussing the theoretical implications of these data, I first discuss data showing that also segment deletion may be gradient in nature.

2.3 Segment deletion: Data

In addition to assimilation, many studies have investigated the nature of segment deletion (Chapter 69: Deletion). It is generally assumed that the absence of segments may result from three different sources. First, the lexicon may represent more than one pronunciation variant for at least some words and segment deletion may result from speakers' selection of reduced pronunciations from their lexicons. Examples of lexicalized reduced pronunciations include English *won't* for *will not* and Dutch [tyk] for [natyrl\k] 'of course' (Ernestus 2000). Second, segments may be absent due to phonological deletion processes operating on the lexically represented unreduced pronunciations. These processes result in phonological surface representations without the absent segments. Both mechanisms (i.e. selection of lexically represented pronunciation variants and phonological processes) result in pronunciation variants that do not contain any acoustic cues for the missing segments, and the absence of these segments is categorical in nature. Alternatively, segments may be absent due to gradient phonetic reduction processes, which reduce the durations and articulatory strengths of segments and make different segments overlap in time (Chapter 82: Reduction). Segments that are absent due to such reduction mechanisms typically leave some traces in the acoustic signal or in the word's articulation. In conclusion, the distinction between categoricality and gradience is also relevant for the theory of segment deletion, since it indicates which type of mechanism is responsible for a given type of deletion.

This view has resulted in several studies investigating the categorical *vs.* gradient nature of segment deletion processes. Browman & Goldstein (1990) hypothesized that most highly productive casual speech reduction processes result from reduction in and overlap of articulatory gestures. They showed in an X-ray study that, for instance, the /t/ in a phrase like *perfect memory* may be acoustically absent, but still articulatory present: speakers may close their lips for the production of the /m/ before the closure of the /t/ is released, which makes the release noise of the /t/ (its most important perceptual cue) inaudible (Browman & Goldstein 1992). Several articulatory and acoustic studies of other highly productive reduction processes support this hypothesis. Thus, Manuel (1992) and Davidson (2006) demonstrated that schwa deletion in

American English is gradient (Chapter 23: Schwa). They reported acoustic differences between consonant clusters resulting from schwa deletion (e.g. [sp] from schwa deletion in *support*) and underlying consonant clusters (e.g. [sp] in *sport*). For instance, clusters resulting from deletion may show aspiration, whereas underlying clusters typically do not. Similarly, Russell (2008) showed that the deletion of the first vowel of a sequence of two in Plains Cree is gradient for his two native speakers (vowels may vary in their duration on the full continuum from values typical for accented full vowels to zero, which implies that they may have clear, some, or no traces at all in the acoustic signal).

In contrast, several less productive processes appear categorical in nature. Examples are the possibly morphosyntactically governed coalescence of /a+i/ or /a:+i/ to [e:] in Plains Cree (Russell 2008) and /e/ deletion in the highly frequent French word combination *c'était* 'it was' (Torreira & Ernestus 2009). Furthermore, segments that may also be absent in more careful speech registers are more probable to be (at least partly) categorically absent. An example is word-medial schwa in French (as in *f/\/nêtre* 'window'; see Bürki *et al.* 2010).

2.4 Gradient assimilation and segment deletion: Theoretical implications

Together, these studies suggest that many productive connected speech processes, such as assimilation and segment deletion, are gradient in nature. If the phonological component contains only categorical processes, as is assumed in traditional versions of generative grammar, these gradient processes should be classified as phonetic, which implies a move of a substantial part of the phonological component to the phonetic component. Theoretical research is necessary into the consequences of this move. Furthermore, the experimental data suggest that post-lexical processes, in particular, show gradience. Additional detailed articulatory and acoustic studies have to investigate whether this generalization is correct. Finally, we have to investigate why some processes are categorical and others gradient and why some processes show inter-speaker and intra-speaker variation. For instance, we have to exclude the possibility that differences result from how participants deal with the experimental situation in which they are tested, including the tools that are put in their mouths for the recording of their articulation. Some participants may show normal speech behavior, while others may adapt their speech.

The evidence for the gradient nature of many connected speech processes has stimulated the development of new theoretical accounts, which do not make a fundamental distinction between the phonological and phonetic components. One of the most influential theories is Articulatory Phonology, developed by Browman & Goldstein (1986 1992) (see also Chapter 18: Atoms of Phonological Representation). This theory assumes that lexical phonological representations consist of strings of articulatory gestures (articulatory scores), which are specified for time and space, and that languages differ in how these gestures may reduce in size and overlap in time. Gradient reduction in gestural size and gradient increase in gestural overlap naturally explain the gradient natures of assimilation and segment deletion processes. For instance, nasal place assimilation in English *gree*[m b]*oat* may result from the early onset of the bilabial closure, during the realization of the preceding nasal, which makes this nasal partly bilabial. In addition, Articulatory Phonology can account for categorical connected speech processes, either by incorporating the processes in the lexical representations of the words (e.g. the French word *c'était* 'it was' may have two lexical representations: one with, and one without, the gestures for the vowel /e/), or by processes that reduce gestural sizes to zero and make gestures completely overlap in time. Note that these different types of mechanisms make Articulatory Phonology a very powerful theory, which can basically explain any reduction pattern. More research is necessary to investigate how this theory can account only for those sound patterns that are actually attested. Furthermore, detailed research is necessary to explain how listeners translate the acoustic signal into gestural scores, which are the basic units of the phonological lexical representations in Articulatory Phonology.

While Browman & Goldstein (1986 1992) proposed Articulatory Phonology as an alternative for theories making a sharp distinction between the phonological and the phonetic components, many researchers (e.g. Byrd & Choi, to appear) do not consider this theory as a competitor for these theories. Rather, they incorporate the ideas of Articulatory Phonology (especially the idea of reduction in and overlap of articulatory gestures) into the phonetic component of generative grammar. Obviously, theoretical research is necessary to investigate the consequences of this incorporation.

2.5 Incomplete neutralization: Data

Final devoicing is another phonological process whose possible gradient nature has received a great deal of attention in the literature (see also Chapter 47: Final Devoicing and Final Laryngeal Neutralization). It has always been assumed to imply a categorical change of voiced obstruents into voiceless ones, and thus a complete neutralization of the distinction between underlyingly voiced and voiceless obstruents in their phonological surface representations and articulatory and acoustic characteristics (e.g. Booij, 1995). Within traditional generative phonology, the output of final devoicing (i.e. final voiceless obstruents) forms the input to other categorical phonological processes (see below). Hence, if final devoicing turns out to lead to incomplete neutralization (i.e. to slightly voiced obstruents) and thus, according to the definitions of generative grammar, to be phonetic in nature, this has consequences for the theoretical accounts of these other phonological processes as well. That is, a gradient nature of final devoicing would have more important theoretical consequences than the gradient nature of the connected speech processes discussed above. Consequently, the possibility of incomplete neutralization has attracted attention from many researchers.

Most experimental studies have investigated the nature of final devoicing by comparing the acoustic characteristics of words differing only in the underlying voice specifications of their final obstruents. The acoustic characteristics that are typically investigated are known to correlate with perceived voicing. They include the duration of the vowel preceding the final obstruent, the duration of the final stop's closure, the duration of this stop's burst, the complete duration of the final fricative, and the duration of glottal vibration during the final obstruent. For instance, Port & O'Dell (1985) investigated ten minimal word pairs in German (e.g. *Rat* 'counsel' *vs. Rad* 'wheel'), read aloud by ten speakers, and showed that all acoustic measures mentioned above provided cues to the underlying voice specification of the final obstruent. In line with this, cluster analysis could correctly classify the underlying voice specifications of the obstruents on the basis of these acoustic measurements for 63% of the tokens. Similar studies have provided evidence for incomplete neutralization in Polish (e.g. Slowiaczek & Dinnsen 1985) and Dutch (e.g. Warner *et al.* 2004). They report acoustic differences between underlyingly voiced and voiceless obstruents in word-final position but also that that these differences may be very small (e.g. Warner and colleagues observed a difference in vowel duration of only 2.5 msec).

Other studies have cast doubt on these findings. For instance, Port & Crawford (1989) recorded five native speakers of German reading three minimal word pairs in four different contexts. The underlyingly voiced final obstruents differed in their realization slightly from the underlying voiceless final obstruents in all four contexts, which is in line with the incomplete neutralization hypothesis. However, speakers differed in which acoustic cues were relevant for the distinction, and, more importantly, whether acoustic characteristics typically cueing voiced obstruents (e.g. longer preceding vowels) were combined with underlyingly voiced or voiceless obstruents. One possible explanation for (part of) these mixed results may be the nature of one of the minimal pairs (*seid,* a verb form of 'to be' *vs. seit* 'since'), since the final obstruent of the member *seid* never occurs in onset position in Modern German and there is consequently no synchronic evidence that this obstruent is underlyingly voiced. Another study showing mixed results was conducted by Charles-Luce (1985), who investigated eight German minimal word pairs. Each of the words appeared in four different sentences, in which it was in sentence-final or medial position. Vowel duration appeared the only reliable cue to underlying voicing, distinguishing /t/ and /d/ in both sentence positions, but /s/ and /z/ only in sentence-final position.

Several studies have raised the question whether the reported evidence for incomplete neutralization may result from the experimental tasks speakers had to perform. Participants typically read sentences aloud, and their pronunciation may therefore show spelling effects. Fourakis & Iverson (1984) investigated this possibility by asking their German participants to conjugate strong verbs after having heard the infinitives (e.g. they heard *reiten* and had to form *ritt* and *geritten*). In this task, participants' attention was not drawn to the spelling of the words to be pronounced. Only 10% of the statistical analyses showed a significant difference between the words ending in underlyingly voiced and underlyingly voiceless obstruents. Importantly, the differences were much smaller than those obtained for the same words in a word-reading task performed by the same speakers. Dinnsen & Charles-Luce (1984) addressed the role of spelling by studying five Catalan minimal word pairs whose members differed from each other in the underlying voice specification of the final obstruent, but not in spelling (e.g. /fat/ *fat* 'fate' *vs.* /fad/ *fat* 'silly'). The words were embedded in carrier sentences, and five speakers read the sentences five times. Two speakers showed incomplete neutralization, one in the expected direction (vowels were 10% longer before underlyingly voiced obstruents in one context condition), and one in the unexpected direction (15% longer closures for underlyingly voiced obstruents). Finally, Warner *et al.* (2006) addressed the role of spelling by comparing two types of Dutch word pairs consisting of morphologically related homophones that differed underlyingly only in the presence of the singleton /t/ *vs.* the geminate /tt/. Importantly, only one of these two types of word pairs reflects the underlying difference in spelling. For instance, /het+\n/ [het\n] *heten* 'are called' *vs.* /het+t\n/ [het\n] *heetten* 'were called' reflects the underlying difference, whereas /het/ [het] *heet* 'am called' *vs.* /het+t/ [het] *heet* 'is called' does not. The results suggest that only those underlying differences that are reflected in orthography lead to pronunciation differences, and that these pronunciation differences are comparable in size to the pronunciation differences induced by incomplete neutralization resulting from final devoicing. Together, these results suggest that incomplete neutralization may be completely driven by orthography.

The nature of final devoicing has also been investigated in several perception studies, addressing the question whether listeners are sensitive to the minimal acoustic differences assumed to be present between underlyingly voiced and voiceless obstruents. If they are, this supports the hypothesis of incomplete neutralization. Participants typically listened to words in isolation and indicated which word they heard by selecting the corresponding orthographic representation (e.g. German listeners heard [rat] and indicated whether they had heard *Rat* 'counsel' or *Rad* 'wheel'). All studies showed that participants tend to choose the intended orthographic representation at just above chance level (e.g. 59% in Port & O'Dell 1985; 62% in Warner *et al.* 2006). In another type of study (Ernestus & Baayen 2007), Dutch participants rated rhymes (i.e. monosyllabic words without their onsets) as 0.7 more voiced on a scale of one to five if the final obstruent was underlyingly voiced compared to voiceless. These studies thus suggest that listeners are sensitive to the minimal cues of incomplete neutralization.

It is legitimate to wonder to what extent the results from the perception experiments are simple task effects, reflecting unnatural linguistic behavior. All studies reported above asked participants to choose between orthographic forms, and hence drew participants' attention to spelling. Moreover, participants could not perform their tasks without taking the acoustic cues to incomplete neutralization into account. Ernestus & Baayen (2006) circumvented this problem by presenting Dutch participants auditorily with non-existing verb stems and asking them to produce the corresponding past tense forms. According to Dutch regular morphology, the appropriate past tense allomorph is *-te* if the final obstruent of the verbal stem is underlyingly voiceless; otherwise it is *-de.* Earlier research had shown that participants interpret the final obstruents of nonce words on the basis of the phonologically similar existing words (Ernestus & Baayen 2003). Ernestus & Baayen (2006) showed that if the final obstruents slightly differ in their voicing, participants interpret these acoustic differences as resulting from incomplete neutralization, and use these differences as a cue for their interpretations of the final obstruents as well. They do so even if their interpretations have no consequences for the spelling of these final obstruents. These findings suggest that listeners are sensitive to incomplete neutralization also if this is not necessary for the experimental task and has no consequences for spelling.

In conclusion, several experimental studies have shown that final devoicing may be incomplete, and that listeners are sensitive to the resulting minimal acoustic differences between underlyingly voiced and voiceless obstruents. Other studies, however, have cast doubt on these findings. Further research into this issue is indispensable.

2.6 Incomplete neutralization: Theoretical implications

The possibility that final devoicing may be gradient is unexpected within generative grammar, since it has always been classified as a phonological process. If final devoicing is phonetic in nature (see e.g. Port & O'Dell 1985, who suggested that final devoicing and incomplete neutralization together form one phonetic implementation process), its output cannot form the input of purely phonological processes. This complicates the theoretical account of several other processes.

One example is the devoicing of voiced fricatives following syllable-final obstruents in Dutch (e.g. while *maa*/n+v/*is* 'angelfish' is pronounced as *maa*[nv]*is,* *gou*/d+v/*is* 'goldfish' is pronounced as *gou*[tf]*is* (see e.g. Booij 1995). In the traditional generative account, this fricative devoicing results from phonological progressive voice assimilation, which is fed by phonological final devoicing (i.e. in the example *gou*[tf]*is*, final devoicing turns word-final /d/ into [t], which triggers devoicing of the following /v/). If final devoicing is phonetic, we have to assume that the devoicing of fricatives results from a phonological process that precedes and is independent of final devoicing. Another possibility is that progressive voice assimilation is phonetic as well, an assumption for which we do not have any acoustic or articulatory support.

A second phonological process that appears to follow final devoicing is resyllabification. In Dutch, word-final obstruents form syllables with following vowel-initial clitics (Chapter 88: Clitics), and the word-final obstruents then occupy onset positions (e.g. *weet ie* /vet i/ 'knows he' is pronounced as /ve-ti/). Importantly, these word-final obstruents are typically voiceless, independently of their underlying voice specification. If final devoicing precedes resyllabification, this is as expected. Hence, if final devoicing is part of the phonetic component, we have to assume that resyllabification is phonetic as well, or we have to assume a phonological process, independent of phonetic final devoicing, which devoices resyllabified obstruents. In summary, if final devoicing is phonetic in nature, we have to assume that other phonological processes are also phonetic or that there are several phonological processes doing partly the same work as final devoicing.

Since both options appear unattractive, Dinnsen & Charles-Luce (1984), as well as Slowiaczek & Dinnsen (1985), suggest that phonetic implementation rules (including final devoicing) may apply before phonological rules. Note that this solution implies that phonetic processes may be of very different types. Traditional phonetic implementation processes translate segments or phonological features into phonetic scores (for articulation) that correspond well with these symbols. Final devoicing, in contrast, would change [+voice] into (almost completely) [–voice].

Given the problems facing a phonetic account of final devoicing, some researchers have proposed that the process is phonological in nature, and that incomplete neutralization results from phonetic implementation processes. These accounts have to solve the question of how the phonetic component can distinguish between obstruents that should be realized as completely voiceless and those that should be slightly voiced. Van Oostendorp (2008) proposes that obstruents may be phonologically specified as voiced ([voice]), as voiceless (no specification for voice) or as devoiced (the feature [voice] is not in a pronunciation relation), and argues that this possibility directly results from assumptions about the phonological component that are necessary for the explanation of unrelated phenomena.

A completely different account for incomplete neutralization is proposed by Ernestus & Baayen (2007). Their account is based on the assumption that the mental lexicon contains representations for all words of the language, including morphologically complex words. Thus, the Dutch lexicon contains both the singular *man*[t] 'basket' and the plural *man*[d]*en* 'baskets'. This assumption is supported by the finding that all words of high frequencies of occurrence, including morphologically inflected and derived words, are recognized and produced more quickly and with fewer errors than words of low frequencies (e.g. Baayen *et al.* 1997; Alegre & Gordon 1999; Chapter 000: Frequency Effects). If the lexicon contains all words of a language, all word-final obstruents can be lexically represented as voiceless. The information that obstruents are voiced in morphologically related words is present in the lexical representations of these related words themselves. Thus, the Dutch word for 'basket' can be lexically represented as *man*/t/, since the plural *man*/d/*en* is stored as well. In this account, incomplete neutralization may be explained in two ways. First, lexical representations may be gradient and contain detailed information about the exact pronunciations of the segments (see also §2.7). Word-final obstruents may thus be represented as slightly voiced. Second, the realization of a word may be affected by the pronunciations of phonologically and morphologically related words. If a stem-final obstruent is voiced in most words, these voiced specifications may affect the pronunciation of the stem-final obstruent in word-final position, which is consequently produced as slightly voiced. This type of lexical analogy would also explain why, in the absence of an abstract mechanism of final devoicing, the final obstruents of new words are always produced as voiceless: this results from the influence of all final voiceless obstruents in the lexicon.

In conclusion, incomplete neutralization has attracted much attention in the theoretical literature, framed both within and outside generative grammar. This may be surprising, since we saw above that the phenomenon is not yet well-established. Note, however, that if future research will show that incomplete neutralization is just an artefact of our experimental paradigms, we still need to explain how these experimental effects can arise in speech production and comprehension. Incomplete neutralization will therefore remain an important theoretical topic.

2.7 Fine phonetic detail in speech processing

Within generative grammar, lexical representations are categorical in nature, as they consist of strings of phonemes or phonological features, abstracting away from phonetic detail which is not necessary to distinguish between these units. In contrast, several researchers are now considering the hypotheses that lexical representations are gradient in nature and reflect fine phonetic detail (see e.g. the account of incomplete neutralization of Ernestus & Baayen 2007, mentioned above) and that one and the same word may have many lexical representations reflecting slightly different pronunciations. These hypotheses are based on the findings that phonetic detail may play an important role in speech comprehension (Chapter 104: Perceptual Effects).

First, experimental data show that listeners are sensitive to phonetic detail providing information about upcoming segments. For instance, in several languages, the relative duration of a vowel is a cue to the presence of additional syllables within the same word, as vowels are typically shorter if they are followed by more syllables. Listeners use these durational cues and predict that syllables like *ham* and *dive* produced with relatively short vowels are part of longer words (i.e. *hamster* and *diver*; e.g. Davis *et al.* 2002; Kemps *et al.* 2005). Similarly, listeners use fine phonetic cues in syllable onsets to predict the presence of /r/ or /s/ in syllable codas (e.g. Heinrich & Hawkins 2009).

Second, several experiments have shown that listeners remember voice characteristics and that these memory traces may affect speech processing. For instance, participants are faster in determining whether two words in a sequence are identical if these two words are presented in the same voice than if they are presented in different voices (Cole *et al.* 1974). Participants tend to complete morphological stems with those suffixes that result in words they have just heard before, especially if these complex words were produced by the same voice as the stems (Schacter & Church 1992). Furthermore, participants tend to mimic previously heard pronunciations in their phonetic detail (Goldinger 1998).

These phonetic detail effects can be accounted for within generative grammar by means of the phonetic component and performance factors. The phonetic component may translate long stretches of phonological segments, rather than single segments, into acoustic signals. Likewise, listeners may analyze acoustic signals to extract not only their segments but also information on following segments (see e.g. Norris & McQueen 2008). This would explain the existence and perceptual relevance of acoustic cues distributed over longer stretches of speech. The effects of voice characteristics may result from the storage of acoustic signals in short-term memory.

In addition, these data may be accounted for by assuming that the detailed phonetic properties of a word are stored in the mental lexicon together with all other information about that word. Thus, the lexical representation *diver* may contain the information that the first vowel is relatively short. Episodic models (e.g. Goldinger 1998) assume that the mental lexicon contains such detailed representations for all tokens of all words that a speaker has ever encountered (such representations are called exemplars). These models can easily explain the processing effects of voice characteristics: If a lexicon contains a word token with the characteristics of a given speaker, the mapping of a new token of that word produced by that same speaker with the exemplars in the mental lexicon is easier than if the mental lexicon does not already contain a token by that speaker.

Episodic models are especially popular in psycholinguistics. So far, two purely episodic models have been developed and computationally implemented for speech processing: Johnson (1997)'s XMOD and Goldinger (1998)'s Minerva. The XMOD model is based on the Lexical Access from Spectra (LAFS) model developed by Klatt (1979), and assumes that the incoming speech signal is transformed into a sequence of spectra. MINERVA was originally developed by Hintzman (1986) and applied to speech by Goldinger. Both XMOD and MINERVA assume that during the recognition process, exemplars respond to an acoustic input in proportion to their similarities to this input, and that their activations spread to the abstract word nodes (XMOD) or to the working memory (MINERVA), which enables recognition.

In addition to these purely episodic models, several hybrid models have been formulated, which assume both abstract lexical representations (strings of phonemes or features) and exemplars. These models can account for all experimental evidence supporting abstract lexical representations (including categorical perception, e.g. Liberman *et al.* 1957) and for the role of fine phonetic detail in speech processing. In addition, they can account for the recent finding that speaker characteristics affect speech processing only if for some reason processing is slow. McLennan & Luce (2005) and Mattys & Liss (2008) showed that tokens produced by the same voice are recognized more quickly than tokens produced by different voices only if the experimental task produces delayed responses (e.g. a shadowing task with a long set response time or a lexical decision experiment that is difficult because of the many word-like pseudowords).

An important hybrid model for speech production is proposed by Pierrehumbert (2002). She assumes that speech production involves the activation of abstract representations, the application of abstract phonological rules (e.g. Prosodic Final Lengthening) and the activation of exemplar clouds of phonological units (e.g. phonemes and phoneme sequences). Two hybrid models for word recognition are Goldinger (2007)'s Complementary Learning System and the model that McLennan *et al.* (2003) developed on the basis of the Adaptive Resonance Theory (Grossberg & Stone 1986). Both models assume that the incoming signal is first analyzed into abstract phonological units, which are matched with the abstract representations in the lexicon, and only then the signal is matched with the stored exemplars. Another hybrid model for word recognition is PolySP (Polysystemic Speech Perception), developed by Hawkins and Smith (Hawkins & Smith 2001; Hawkins 2003). This model assumes that a memory trace does not only contain acoustic information, but also multi-medial context, for instance, visual information about the speaker's gestures. In addition, the model assumes that the analysis of an acoustic input into its linguistic units (phonemes, etc.) may precede (and contribute to) or coincide with or follow word recognition or not take place at all, depending on the circumstances.

In conclusion, experimental evidence suggests that gradient acoustic characteristics play a role in speech processing. More research is necessary showing which types of acoustic characteristics are relevant, how this gradient information is accessed under which conditions and how the role of this type of information should be accounted for in speech production and comprehension models.

2.8 Conclusions

Gradience appears to be a much more important characteristic of speech sounds than is traditionally assumed. Place and voice assimilation, segment deletion and final devoicing often result in sounds showing incomplete neutralization, i.e. they result in sounds that contain characteristics of more than one phoneme or that are only partly absent. Since generative grammar assumes that gradience is a characteristic of the phonetic component, these data suggest that within this theory many processes that have always been classified as phonological actually belong to the phonetic component. Particularly in the case of final devoicing, this reclassification has consequences for the classification of other speech processes as well. Alternative theories have been developed, which assume that the phonological primitives are articulatory gestures or that lexical representations reflect the gradient nature of speech sounds. These theories are supported by data showing that fine phonetic detail affects speech processing.

3 Productive sound patterns

3.1 Introduction

Gradience does not only play a role in the discussion of the phonological and phonetic components and of the nature of lexical representations, but also in the theoretical discussion of the nature of productive (morpho)phonological processes. Within traditional generative phonology, a productive process applies always and to all input that satisfies its structural description. Productive processes are thus categorical in nature. Recent research suggests, however, that some productive processes show gradience. The following two subsections discuss evidence for gradient phonological processes, their implications for generative phonology and alternative theories that account for the gradient data.

3.2 Phonotactic constraints

The first type of phonological processes whose categorical nature has been seriously questioned are phonotactic generalizations. Within traditional generative phonology, all illegal sequences are considered equally illegal, all legal sequences as equally legal, and there is no gradience in legality.

If this assumption is correct, differences in the frequencies of occurrence of phonemes and phoneme sequences are based on coincidence. Pierrehumbert (1994) studied the frequencies of consonants and consonant clusters at the beginning of words, at the end of words (excluding the phonological appendix), and in syllable onset and coda positions within morpheme-internal consonant clusters (e.g. the frequency of [n] in words like *vanquish*, where it is in syllable coda position within a consonant cluster, and the frequency of [st] in words like *lobster*, where it is in syllable onset position in a consonant cluster) in an American English dictionary. If the consonants are randomly distributed over the positions, the frequencies of a given consonant (cluster) in the different positions should be unrelated. This appeared not to be the case. The frequency of a morpheme-internal cluster appears highly correlated to the frequency of its first part (i.e. the consonants in coda position) in word-final position and the frequency of its second part (i.e. the consonants in onset position) in word-initial position. Phonemes and phoneme sequences structurally differ in their frequencies in a language.

Crucially, language users reflect these frequencies in their well-formedness judgments of nonce words and parts of words (Chapter 90: Morpheme Structure Constraints). Speakers typically judge high-frequency rhymes as ‘phonologically’ better than low-frequency rhymes (Treiman *et al.* 2000), phonotactically legal nonce words as better if they contain phoneme sequences of high frequencies (e.g. Vitevitch *et al.* 1997; Frisch *et al.* 2000) and nasal–obstruent clusters as better if these clusters are more frequent (Hay *et al.* 2004). Thus, *blick* is rated as a good English word, *bnick* as an impossible word and *bwick* is rated in between. Importantly, these gradient well-formedness judgments are obtained both if participants are allowed to provide gradient responses and if they have to provide categorical judgments, with the judgments being averaged over participants (Frisch *et al.* 2000). This strongly suggests that phonotactic constraints are gradient, rather than categorical.

Language users’ judgments of a nonce word are also affected by the phonological distance of this word from existing words (Chapter 91: Neighborhood Effects). Thus, participants rate a nonce word as more well-formed if it differs in fewer phonemes from an existing word (Greenberg & Jenskins 1964; Ohala & Ohala 1986). In addition, their well-formedness judgments are related to the size of a word’s phonological neighborhood (Bailey & Hahn 2001; Hammond 2004), which is typically defined as the number of existing words that can be changed into that word by the substitution, addition or deletion of one single phoneme. Importantly, the effect of the word’s phonological neighborhood is independent of the effects of the frequencies of the word’s constituents (i.e. the effect is also present if words with small and larger neighborhoods are matched in the frequencies of their constituents). This shows again that well-formedness judgments are not categorical (i.e. it is not the case that a word is either completely well-formed or completely ill-formed). Rather, these judgments are gradient between completely well-formed and completely ill-formed.

Importantly, the measures affecting well-formedness judgments also play a role in other (psycho-)linguistic tasks. The frequencies of the phonemes and phoneme sequences in a word have been shown to affect speech production, recognition, and learning. For instance, participants are better at repeating nonce words made up of high-frequency rather than low-frequency phoneme sequences (Vitevitch *et al.* 1997) and at transcribing such words orthographically (Hay *et al.* 2004). Participants tend to interpret ambiguous fricatives as the most probable ones given the preceding and following segments (Pitt & McQueen 1998). Nine-month-old infants prefer to listen to words consisting of high-frequency rather than low-frequency phoneme sequences (Jusczyk *et al.* 1994). Furthermore, when both eight-month-old infants and adults are presented with continuous speech from a non-existing (artificial) language, they extract the words of this language on the assumption that frequent phoneme sequences form (parts of) words, while the less frequent ones span word boundaries (Saffran *et al.* 1996a; Saffran *et al.* 1996b). Similarly, speech production and comprehension are affected by a word's phonological neighborhood. Thus, participants recognize words with large neighborhoods more slowly in auditory lexical decision (e.g. Luce & Pisoni 1998) and produce them with more expanded vowel spaces (Munson & Solomon 2004), while pre-school-aged children produce such words more quickly and with fewer errors in picture-naming tasks (Arnold *et al.* 2005).

Several generative linguists have assumed that the gradience of well-formedness judgments may be merely a task effect, resulting from performance factors (for a discussion, see Schütze 2005). This account is in line with the finding that the variables affecting well-formedness ratings also play roles in speech production, perception and learning, which are certainly modulated by performance factors.

In addition, there is a continuum of accounts which differ in their assumptions about the contributions of the phonological component and the mental lexicon. The models at one end of the continuum assume that the gradience of well-formedness judgments results from the gradient nature of the phonological component itself. This component would be gradient due to the probabilistic nature of its constraints or rules. For instance, Hammond (2004) frames his account of gradient well-formedness judgments within Probabilistic Optimality Theory, which is based on Stochastic Optimality Theory, developed by Boersma (1998). The idea is that the ranking of constraints is variable, and that a given (markedness or faithfulness) constraint outranks some other constraint with a certain probability. If this probability is smaller than 1, the phonological component shows variation, sometimes favoring one form and sometimes another, which results in gradient well-formedness rankings. The probability of a given ranking (and consequently the judgment of a given form) may be co-determined by the frequencies of phoneme sequences and the exact contents of the mental lexicon. Models at the other end of the continuum assume that well-formedness judgments for a given word result only from the comparison of that word with all words in the mental lexicon and their constituents. The visual or auditory presentation of a word leads to the activation of all (phonologically) similar words in the lexicon and their constituents, and a higher total lexical activation leads to a higher well-formedness rating. In these analogical models, there is thus no role for an abstract phonological component with hardwired phonological constraints or rules (e.g. Bailey & Hahn 2001). Models positioned between the two ends typically assume that the effects of constituent frequencies result from phonotactic knowledge and the effects of phonological neighborhood from lexical knowledge. Phonotactic knowledge is permanently stored in the phonological component, while lexical knowledge is deduced from the mental lexicon if necessary (e.g. Bailey & Hahn 2001; Albright 2009).

In summary, the evidence for gradience in well-formedness judgments is undisputed. Detailed research is necessary, in different domains of phonology, to establish the best theoretical account.

3.3 Allomorphy

A second type of productive phonological processes that appears gradient are those involved in morphological processing. These morphophonological processes select affixes on the basis of the phonological properties of the words' stems (Chapter 103: Phonologically Conditioned Allomorphy). For instance, Dutch regular past tense forms consist of a verbal stem and the suffix -*te* or -*de*. According to the traditional literature (which follows Dutch orthography), the correct allomorph is -*te* if the verbal stem ends in an underlyingly voiceless obstruent (e.g. *sta*/p+t/*e* 'stepped'), otherwise it is -*de* (e.g. *kra*/b+d/*e* 'scratched'). It has been shown recently that at least some of these apparently perfectly categorical generalizations do not do justice to the full data.

One example is the above-mentioned regular past tense formation in Dutch. Ernestus & Baayen (2004) show that the description of the selection of the past tense allomorph given in the literature is too simplistic. Speakers tend to choose the non-standard allomorph for verbal stems that are special, in that the underlying voice specification of their final obstruent is unexpected given the other stems ending in the same type of final rhyme in the lexicon. For instance, speakers often choose the non-standard allomorph for *kra*/b/ (creating *kra*/b+t/*e*), which is one of the few Dutch verb stems ending in a short vowel and a voiced (instead of voiceless) bilabial stop. The pattern 'short vowel–underlyingly voiceless bilabial stop' is much more common (e.g. *sto*/p/ 'stop', *kla*/p/ 'clap', *me*/p/ 'slap', *ni*/p/ 'sip') than the pattern 'short vowel–underlyingly voiced bilabial stop', and speakers tend to add the allomorph that is correct for the majority of verbs ending in a short vowel and a bilabial stop to the minority of verbs for which it is incorrect (i.e. verbs ending in a short vowel and an underlyingly voiced bilabial stop). These findings can easily be incorporated in all types of theoretical accounts, since the only adaptation necessary is that the broad generalizations are replaced or supplemented by generalizations that are more specific for the precise phonological properties of the words. Apparently, Dutch requires a generalization stating that stems ending in short vowels and bilabial stops tend to select -*te*.

Importantly, however, the facts are more complex. First, Ernestus & Baayen (2004) observe that if participants select the standard allomorph, they do so more quickly for verbs following the majority patterns than for exceptional verbs (i.e. they produce forms of the type *stapte* more quickly than forms of the type *krabde*). Second, Ernestus & Baayen (2003, 2004; see also Ernestus 2006) find that speakers show stochastic behavior: they often do not agree with each other, and the same speaker may choose -*te* for some verbs and -*de* for other verbs of the same type. Similar results have been found, among others, for past tense formation in English (Albright & Hayes 2003), the choice of the English indefinite article (*a vs. an*; Skousen 1989) and vowel harmony in

Hungarian (Hayes & Londe 2006). Apparently, the morphophonological processes that have to replace or supplement the traditional broad generalizations are not simple categorical rules that apply whenever their structural description is met. The processes are gradient in nature.

Speakers' probabilistic behavior has been accounted for in the two types of approaches (forming a continuum) that also explain the gradience of well-formedness ratings (see above). The first approach holds that constraints or rules are probabilistic in nature. Thus, in Stochastic Optimality Theory (Boersma 1998), constraint rankings are stochastic, and in the rule-based account proposed by Albright & Hayes (2003), rules differ in their confidence intervals. Both accounts assume that the probability of a constraint ranking or rule (and thus of a given form) is determined by the exact contents of the mental lexicon. While this approach can account well for the observed probabilistic effects, additional assumptions are necessary to explain why speakers are slower in selecting the standard allomorph if it receives less lexical support than the other allomorph (for a discussion see Ernestus 2006). The second approach to speakers' stochastic behavior assumes that when speakers select an allomorph for a word, they check all words in their lexicons online. The probability that they select a given allomorph is proportional to its support from the words in the lexicon, with words that are more similar to the target word being more influential. If the target word itself is in the lexicon as well and supports a different allomorph than the one receiving the greatest lexical support from the other words, this may result in severe competition between the two allomorphs, which may lead to the selection of the non-standard allomorph and longer response latencies (Ernestus & Baayen 2004).

In conclusion, phonologically driven allomorphy also strongly suggests that gradience is an important characteristic of phonology. The generalizations formulated in the generative literature appear too crude, given that speakers show probabilistic behavior. Several models can account for the obtained observations so far. More data is necessary to tease the different accounts apart.

4 Conclusions

In the early days of generative grammar, phonology was assumed to be completely categorical in nature. The present chapter has provided a summary of different types of corpus-based and experimental studies which strongly suggests that many processes traditionally classified as phonological are in fact gradient in nature. Sounds may contain characteristics of different categories and speakers may show probabilistic behavior. These data have given rise to modifications of traditional generative phonology and to the development of new theories, including theories assuming different types of phonological primitives and phonological representations, and theories challenging the role of abstract generalizations. Further research is necessary to obtain a more detailed view of the role of gradience in phonology and to test different theoretical accounts. Until then, we have to conclude that gradience is an important challenge for phonology.

References

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26. 9–41.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90. 119–161.

Allegre, Maria & Peter Gordon. 1999. Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40. 41–61.

Arnold, Hayley S., Edward G. Conture & Ralph N. Ohde. 2005. Phonological neighborhood density in the picture naming of young children who stutter: Preliminary study. *Journal of Fluency Disorders* 30. 125–148.

Baayen, R. Harald, Ton Dijkstra & Robert Schreuder. 1997. Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language* 36. 94–117.

Bailey, Todd M. & UIrike Hahn. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44. 568–591.

Barry, Martin C. 1992. Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication* 11. 393–400.

Berg, Rob J. H. van den. 1988. *The perception of voicing in Dutch two-obstruent sequences*. Enschede: Sneldruk Enschede.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives.* The Hague: Holland Academic Graphics.

Booij, Geert. 1995. *The phonology of Dutch*. Oxford: Clarendon Press.

Browman, Catherine P. & Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3. 219–252.

Browman, Catherine P. & Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech*, 341–376. Cambridge: Cambridge University Press.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49. 155–180.

Bürki, Audrey, Mirjam Ernestus & Ulrich Frauenfelder. 2010. Is there only one fenêtre in the production lexicon? On-line evidence about the nature of phonological representations. *Journal of Memory and Language* 62. 421–437.

Byrd, Dani & Susie Choi. To appear. At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio & Nathalie Vallée (eds.) *Laboratory Phonology 10.* Berlin: Mouton de Gruyter.

Charles-Luce, Jan. 1985. Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics* 13. 309–324.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Cohn, Abigail. 1993. Nasalisation in English: Phonology or phonetics. *Phonology* 10. 43–81.

Cole, Ronald A., Max Coltheart & Fran Allard. 1974. Memory of a speaker's voice: Reaction time to same- or different-voiced letters. *Quarterly Journal of Experimental Psychology* 26. 1–7.

Davidson, Lisa. 2006b. Schwa elision in fast speech: Segmental deletion or gestural overlap. *Phonetica* 63. 79–112.

Davis, Matthew H., William D. Marslen-Wilson & M. Gareth Gaskell. 2002. Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 28. 218–244.

Dinnsen, Daniel A. & Jan Charles-Luce. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12. 49–60.

Ellis, Lucy & William J. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30. 373–396.

Ernestus, Mirjam. 2000. *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology-phonetics interface.* Utrecht: LOT.

Ernestus, Mirjam. 2006. Statistically gradient generalizations in phonology. *The Linguistic Review* 23. 217–234.

Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79. 5–38.

Ernestus, Mirjam & R. Harald Baayen. 2004. Analogical effects in regular past tense production in Dutch. *Linguistics* 42. 873–903.

Ernestus, Mirjam & R. Harald Baayen. 2006. The functionality of incomplete neutralization in Dutch: the case of past-tense formation. In Louis M. Goldstein, Douglas Whalen & Catherine T. Best (eds.) *Laboratory phonology* 8, 27–49. Berlin & New York: Mouton de Gruyter.

Ernestus, Mirjam & R. Harald Baayen. 2007. Intraparadigmatic effects on the perception of voice. In Jeroen van de Weijer & Erik Jan van der Torre (eds.) *Voicing in Dutch: (De)voicing - phonology, phonetics, and psycholinguistics,*. 153–174. Amsterdam: John Benjamins.

Ernestus, Mirjam, Mybeth Lahey, Femke Verhees & R. Harald Baayen. 2006. Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America* 120. 1040–1051.

Farnetani, Edda & M. Grazia Busà. 1994. Italian clusters in continuous speech. *Proceedings of the 3rd International Conference on Spoken Language Processing*, 359–362. Yokohama: Acoustical Society of Japan.

Fourakis, Marios & Gregory K. Iverson. 1984. On the 'incomplete neutralization' of German final obstruents. *Phonetica* 41. 140–149.

Frisch, Stefan A., Nathan R. Large & David B. Pisoni. 2000. Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42. 481–496.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105. 251–279.

Goldinger, Stephen D. 2007. A complementary-systems approach to abstract and episodic speech perception. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*, 49–54. Saarbrücken: Saarland University.

Greenberg, Joseph H. & James J. Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20. 157–177.

Grossberg, Stephen & Gregory Stone. 1986. Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review* 93. 46–74.

Hammond, Michael. 2004. Gradience, phonotactics, and the lexicon in English phonology. *International Journal of English Studies* 4. 1–24.

Hardcastle, William J. 1972. The use of electropalatography in phonetic research. *Phonetica* 25. 197–215.

Hawkins, Sarah. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31. 373–405.

Hawkins, Sarah & Rachel Smith. 2001. Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Rivista di Linguistica* 13. 99–188.
Hay, Jennifer, Janet B. Pierrehumbert & Mary E. Beckman. 2004. Speech perception, well-formedness and the statistics of the lexicon. In John Local, Richard Ogden & Rosalind Temple (eds.) *Phonetic interpretation: Papers in laboratory phonology VI*. Cambridge: Cambridge University Press. 58–74.
Hayes, Bruce & Zsuzsa Cziráky Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23. 59–104.
Heinrich, Antje & Sarah Hawkins. 2009. The effect of r-resonance information on intelligibility. In Maria Uther, Roger Moore & Stephen Cox (eds.) *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech 2009),* 804–807. Brighton: Causal Productions.
Hintzman, Douglas L. 1986. 'Schema abstraction' in a multiple-trace memory model. *Psychological Review* 93. 411–428.
Jansen, Wouter. 2007. Dutch regressive voicing assimilation as a 'low level phonetic process': Acoustic evidence. In van de Weijer & van der Torre (2007). 125–151.
Johnson, Keith 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.) *Talker variability in speech processing,* 145–165. San Diego: Academic Press.
Jusczyk, Peter W., Paul A. Luce & Jan Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33. 630–645.
Keating, Patricia. 1985. Universal phonetics and the organization of grammars. In Victoria A. Fromkin (ed.) *Phonetic linguistics: Essays in honor of Peter Ladefoged*, 115–132. Orlando: Academic Press.
Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5. 275–292.
Keating, Patricia. 1990a. Phonetic representations in a generative grammar. *Journal of Phonetics* 18. 321–334.
Keating, Patricia. 1990b. The window model of coarticulation: Articulatory evidence. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech*, 451–470. Cambridge: Cambridge University Press.
Kemps, Rachèl, Lee H. Wurm, Mirjam Ernestus, Robert Schreuder & R. Harald Baayen. 2005. Prosodic cues for morphological complexity: Comparatives and agent nouns in Dutch and English. *Language and Cognitive Processes* 20. 43–73.
Kenstowicz, Michael & Charles W. Kisseberth. 1979. *Generative phonology: Description and theory*. New York: Academic Press.
Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70. 419–454.
Klatt, Dennis H. 1979. Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics* 7. 279–312.
Kochetov, Alexei & Marianne Pouplier. 2008. Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. *Phonology* 25. 399–431.
Kuzla, Claudia, Taehong Cho & Mirjam Ernestus. 2007. Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics* 35. 301–320.
Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Liberman, Alvin M., Katherine S. Harris, Howard S. Hoffman & Belver C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54. 358–368.
Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 19. 1–36.

Manuel, Sharon Y. 1992. Recovery of ‘deleted’ schwa. In Olle Engstrand & Catharina Kylander (eds.) *Perilus XIV: Papers from the Symposium on Current Phonetic Research Paradigms for Speech Motor Control,* 115–118. Stockholm: University of Stockholm.

Mattys, Sven L. & Julie M. Liss. 2008. On building models of spoken-word recognition: When there is as much to learn from natural ‘oddities’ as artificial normality. *Perception and Psychophysics* 70. 1235–1242.

McLennan, Conor T. & Paul A. Luce. 2005. Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 31. 306–321.

McLennan, Conor T., Paul A. Luce & Jan Charles-Luce. 2003. Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 29. 539–553.

Munson, Benjamin & Nancy P. Solomon. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language and Hearing Research* 47. 1048–1058.

Nolan, Francis. 1992. The descriptive role of segments: Evidence from assimilation. In Gerard Docherty & D. Robert Ladd (eds.) *Papers in laboratory phonology II: Gesture, segment, prosody*, 261–280. Cambridge: Cambridge University Press.

Norris, Dennis. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52. 189–234

Norris, Dennis & James M. McQueen. 2008. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* 115. 357–395.

Ohala, John J. & Manjari Ohala. 1986. Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In John J. Ohala & Jeri Jaeger (eds.) *Experimental phonology,* 239-252. Orlando: Academic Press.

Oostendorp, Marc van. 2008. Incomplete devoicing in formal phonology. *Lingua* 118. 1362–1374.

Perkell, Joseph S., Marc H. Cohen, Mario A. Svirsky, Melanie L. Matthies, Iñaki Garabieta & Michael T. T. Jackson. 1992. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America* 92. 3078–3096.

Pierrehumbert, Janet B. 1990. Phonological and phonetic representation. *Journal of Phonetics* 18. 375–394.

Pierrehumbert, Janet B. 1994. Syllable structure and word structure: a study of triconsonantal clusters in English. In Patricia A. Keating (ed.) *Phonological structure and phonetic form: Papers in laboratory phonology III*, 168–188. Cambridge: Cambridge University Press.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory phonology* 7, 101–139. Berlin & New York: Mouton de Gruyter.

Pitt, Mark A. & James M. McQueen. 1998. Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39. 347–370.

Port, Robert F. & Penny Crawford. 1989. Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17. 257–282.

Port, Robert F. & Michael O’Dell. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13. 455–471.

Russell, Kevin. 2008. Sandhi in Plains Cree. *Journal of Phonetics* 36. 450–464.

Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport. 1996. Statistical learning by 8-month old infants. *Science* 274. 1926–1928.

Saffran, Jenny R., Elissa L. Newport & Richard N. Aslin. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35. 606–621.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale.* Lausanne & Paris: Payot.

Schacter, Daniel L. & Barbara A. Church. 1992. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18. 915–930.

Schütze, Carson T. 2005. Thinking about what we are asking speakers to do. In Stephan Kepser & Marga Reis (eds.) *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 457–484. Berlin & New York: Mouton de Gruyter.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer.

Slowiaczek, Louisa M. & Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13. 325–341.

Torreira, Francisco & Mirjam Ernestus. 2009. Vowel elision in connected French: The case of vowel /e/ in the word *c'était*. In Maria Uther, Roger Moore & Stephen Cox (eds.) *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, 448–451. Brighton: Causal Productions.

Treiman, Rebecca, Brett Kessler, Stephanie Knewasser, Ruth Tincoff & Margo Bowman. 2000. English speakers' sensitivity to phonotactic patterns. In Michael B. Broe & Janet B. Pierrehumbert (eds.) *Papers in laboratory phonology V: Acquisition and the lexicon,* 269–282. Cambridge: Cambridge University Press.

Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce & David Kemmerer. 1997. Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40. 47–62.

Warner, Natasha, Erin Good, Allard Jongman & Joan A. Sereno. 2006. Orthographic *vs.* morphological incomplete neutralization effects. *Journal of Phonetics* 34. 285–293.

Warner, Natasha, Allard Jongman, Joan A. Sereno & Rachèl Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32. 251–276.

Weijer, Jeroen van de & Erik Jan van der Torre (eds.) 2007. *Voicing in Dutch: (De)voicing – phonology, phonetics, and psycholinguistics.* Amsterdam & Philadelphia: John Benjamins.

Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In Bruce Connell & Amalia Arvanti (eds.) *Papers in laboratory phonology IV: Phonology and phonetic evidence*, 282–302. Cambridge: Cambridge University Press.

Zsiga, Elizabeth C. 1997. Features, gestures, and Igbo vowels: An approach to the phonology–phonetics interface. *Language* 73. 227–274.