# The learner of a perception grammar as a source of sound change*

*Silke Hamann*

In this paper, I argue that a regular diachronic sound change is the result of a different interpretation of the same auditory information, as put forward by Ohala (1981 et seq.). Whereas Ohala describes such an account of sound change as purely phonetic, I show that it involves phonological knowledge, namely the language-specific use of auditory cues and their mapping onto language-specific phonological categories.

Two diachronic developments of retroflex segments, namely retroflexion of rhotic plus coronal consonant sequences in Norwegian and retroflexion of labialised coronal obstruents in Minto-Nenana, illustrate these claims. For both, the differences across generations are modelled in Optimality Theory with the help of language-specific cue constraints in a perception grammar (following Boersma 1997 et seq.). This approach is shown to be superior to the descriptive approach of cue re-association proposed by Ohala because it provides a formal account that includes differences in cue weighting (especially the disregard of cues that became unreliable) and differences in emergent phonological categories.

## 1. Introduction

In recent years, a number of phonological studies have turned their attention to the relation between speech perception and phonology and have highlighted the relevance of this topic, see e.g. the articles in Hume and Johnson (2001a) and in the present volume. The conclusions drawn in these studies, however, diverge. The majority of scientists working on the perceptual basis of phono-

* I want to thank Siri Tuttle for drawing my attention to the Athabaskan data and providing me with literature on this topic. Earlier stages of this research were presented at the *11th Manchester Phonology Meeting*, May 2003; at the *16th International Congress of Historical Linguistics in Copenhagen*, August 2003; and at the ZAS Berlin. I thank the audiences on all three occasions for their comments. I have also received helpful suggestions by Paul Boersma and John Ohala. I gratefully acknowledge a VENI postdoctoral fellowship by the Dutch Science Foundation (NWO; GW 016.064.057).

logical processes, for instance Steriade (1995, 2001), Hume and Johnson (2001b), and Wright (2001), assume or explicitly claim that speech perception informs phonology but lies outside the scope of phonological theory, whereas a few, such as Boersma (1998), Broselow (2004, this volume) and Pater (2004), claim that speech perception is part of our phonological knowledge and therefore has to be included in phonological theory.

One argument against the language-specificity of speech perception put forward by its opponents is the fact that speech perception shares cognitive facilities with general perception, as exemplified with the following citation of Hume and Johnson (2001b: 14):

> [T]o the extent that language sound patterns are caused by external factors such as speech perception, these factors are reflected in the formal phonological theory. Yet, to incorporate them directly into phonological theory erroneously implies that they are exclusive to language.

In the present article I argue that there is no inherent conflict between language-specific perception and our general auditory capacity, but that both are present and apply in the perception of speech sounds.

With respect to sound change, Ohala's (1981) groundbreaking article "The listener as a source of sound change" illustrates that an interaction between sound perception and phonology is one cause of diachronic change. Again, this interaction can be interpreted in two ways. Hume and Johnson (2001b) write that "Ohala's (1981) account of the listener as a source of sound change is one of the most explicit accounts of a *point of contact* between speech perception and language sound structure" (p.7, italics mine), see also Holt (1997) and Mielke (2003) for a similar interpretation. Departing from this interpretation, I show in the present article that Ohala's work on sound change (1974 et seq.) provides no evidence for the independence of speech perception and phonology; on the contrary, it actually supports the phonological nature of speech perception.

In the present study I discuss data of two diachronic developments of retroflexes via re-analysis; both processes and a sample language for each are given in (1) (illustrated by one retroflex segment of each language).

(1) *Retroflexion*

a. *in rhotic context:* *rt > ʈ Norwegian (North-Germanic)

b. *via labialisation:* *ʃ$^{w}$ > ʂ$^{ɻ}$ Minto-Nenana (Athabaskan)

Bhat (1973), who provides the first comprehensive typological description of retroflexes and includes a short treatise on their diachronic development, mentions that retroflexes are "introduced into a language mainly through the assimilatory influences of neighbouring sounds such as back vowels, velar consonants, r, or at a later stage by other retroflexed consonants" (p.55).[1] Bhat's "assimilatory influence" is thus restricted to *coarticulatory phenomena*.

In the present study the two processes under (1) are shown to involve *perceptual reinterpretation*, which is often based on coarticulatory variation in the input. Due to this variation in the input, the child acquires a perception grammar that differs from the perception grammar(s) of the previous generation. I formalise the acquisition and workings of a perception grammar within the framework of Optimality Theory (McCarthy and Prince 1993, Prince and Smolensky [1993] 2004; henceforth: OT), following Boersma (1997 et seq.), Boersma, Escudero and Hayes (2003), Escudero and Boersma (2003, 2004), and Hamann (2003a). For this, I use the model of *Bidirectional Phonetics and Phonology* (Boersma 2006, 2007, to appear; henceforth: BiPhon) because it is the only linguistic model at present that includes an explicit formalisation of the phonetics-phonology interface with so-called cue constraints, which map the auditory form onto a phonological form and vice versa.[2]

Within OT, sound change is usually assumed to be characterised by a different ranking of the same constraints between two (or more) diachronic stages of a language, see for instance the work by Jacobs (1995), Gess (1996), Green (1997, 2001), Holt (1997), Ham (1998), Bermúdez-Otero (1999), and Hamann (2005). In contrast to these approaches, I propose in the present article that the language learner does not merely construct a different *ranking* of the same constraints from the input, but can employ different *constraints* than the previous generation, see Boersma (1998, 2003) and Gess (2003).[3] My proposal is based on the assumption that constraints and phonological categories are not innate and therefore not universally available to every learner of every language, but have to be constructed language-specifically on the basis of the input that the learner receives. This emergentist or evolutionarist view of phonology is not new, see e.g. Boersma (1997 et seq.) and Mielke (2004) on the

1. Retroflexes can also emerge from voiced (implosive) stops, and via areal spread (both are mentioned briefly by Bhat 1973: 41 and 50, respectively, for a detailed account of retroflexion via voiced stops, see Hamann and Fuchs to appear).
2. Alternatively, the mapping between auditory and phonological forms could be modelled for instance with neural networks. Since alternatives that account for the phonetics/phonology interface have not been elaborated yet, and an elaboration would go beyond the focus of the present paper, I employ here the existing and working cue constraints.
3. In contrast to the present study and work by Boersma, Gess (2003) focuses on the role of the speaker in initiating sound change.

non-innateness of phonological features and Haspelmath (1999) and Blevins (2004) on the emergence of typologically similar patterns via diachronic adaptation instead of universal constraints.

The article has the following structure. In §2, I elaborate Ohala's approach to sound change as initiated by the listener. Section 3 contrasts this with the present approach, where the learner of a perception grammar initiates sound change. The workings of the presented model are elaborated in §4. Section 5 provides the data of the two retroflex processes and a formalisation of these changes in terms of OT perception grammars. Conclusions are given in the last section.

## 2. Ohala's listener as a source of sound change

The first full-fledged perceptual account of sound change has been proposed by Ohala (1974, 1981, 1986, 1989, 1993b, 1993a, 1995) who attributes a large role in diachronic change to the listener. Ohala's account is reduced to sound changes that have been attested independently in many unrelated languages and that are most likely to arise from physiological factors, excluding changes introduced via spelling pronunciation, paradigm uniformity effects, and other language-specific factors. Ohala illustrates that the listener, when reconstructing what the speaker says, is often confronted with contextual assimilation of sounds, and has to undo these assimilatory changes to derive the correct underlying phonological form (Ohala and Feder 1994; Ohala and Shriberg 1990).

An example recurrently used by Ohala is the contextual influence of a coronal stop /t/ on a back vowel /u/. In this context, the low second formant (henceforth: F2) of the back vowel /u/ is realised with rising F2 transitions for the /t/. These rising F2 transitions can be interpreted as belonging to a more fronted vowel such as /y/, which has higher F2 values than /u/. The contextual influence of a /t/ on a back vowel is stronger in the transitions from vowel to consonant (henceforth: VC) than in the transitions from consonant to vowel (henceforth: CV), because the quick tongue tip gestures are anticipated already during preceding non-coronal sounds but do not carry over onto gestures of following non-coronal sounds. Listeners that hear intended sequences like /ut/ have to subtract the influence of the coronal context from the vowel to derive the intended /u/. Ohala (1981: 181, et seq.) represents this performance of the speaker and the listener as in Figure 1.

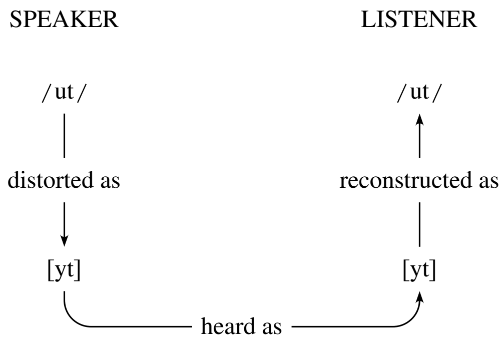


*Figure 1.* Ohala's model of a correct speaker – listener interaction.

We have to keep in mind that the use of the symbol [y] for a back rounded vowel in coronal context is a simplification since it implies that this segment is realised identically to a front rounded /y/, which is not the case: listeners of languages that contrast /u/ and /y/ do not generally misperceive [ut] as /yt/. It is, however, the case that in languages with only one high rounded vowel such as English this sound is usually realised less back than in languages that have a backness contrast for high rounded vowels such as French and German (the so-called *dispersion effect*, see e.g. Liljencrants and Lindblom 1972 and Lindblom 1986). As a result, the /u/ in the non-contrasting languages lies perceptually between the /u/ and /y/ in contrasting languages, see for instance Flege's (1987) study where English native speakers perceived both French /u/ and /y/ as English /u/.

In Ohala's model, sound change can occur if the listener performs one of two possible types of "misperceptions", namely uncorrected distortion and dissimilation. In uncorrected distortion, also referred to as *hypocorrection* or *false association*, the listener is unable to correlate the context-induced distortion with the context, which may be almost inaudible due to factors such as loud background noise. The listener thus fails to correct for the context, as in Figure 2a, where the high F2 transition of the coronal plosive is associated with the vowel and thus the intended /ut/ is perceived as /yt/.

In dissimilation, or *hypercorrection*, the listener subtracts what he/she assumes to be due to contextual influence but was actually intended by the speaker, see Figure 2b. Here the high F2 value of the vowel is interpreted as belonging to the coronal plosive, and therefore /yt/ is perceived as /ut/.

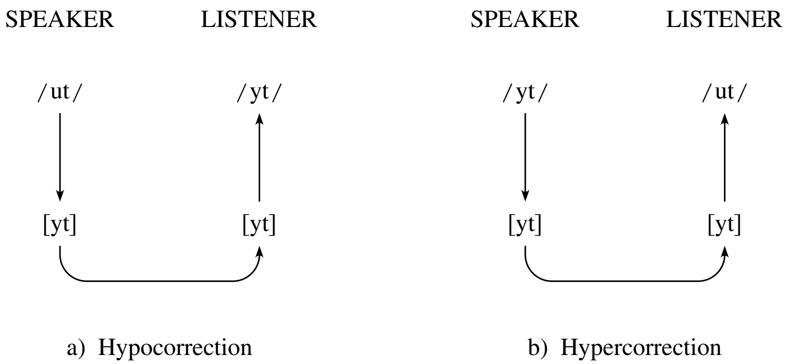


*Figure 2.* The two types of Ohala's sound change if the speaker – listener interaction goes wrong.

According to Ohala, both types of sound changes are some kind of "parsing error[s] by the listener" (1993a: 263), and occur "due to a break-down […] in the system" (1992: 340). They differ in as far as hypocorrection, which can introduce new phonological categories, is only possible if the listener is linguistically inexperienced and does not know enough about contextual influences yet. Hypercorrection, on the other hand, can occur with adult speakers, since the resulting categories are not new in the language.

In Ohala's model, the listener simply reverses the speaker's task of producing [y] from a phonological form /u/, as can be seen in Figure 1: The listener has to retrieve an /u/ from the phonetic form [y]. He or she does this with the help of what Ohala (1981: 183) terms *"reconstructive rules"*. These rules differ from rules in traditional generative phonology in two ways: They operate on a highly variable input, and they derive more abstract representations from less abstract ones (ibid.). In a later account, Ohala (1992) describes the process of reconstruction as "a cognitive act on the part of the listener" (p.326). He goes on to say "I would be willing to call this a *rule of grammar* although of a type not acknowledged by most phonologists" (ibid., italics mine). Ohala explicitly states that speech reconstruction requires the listener to have an "elaborate knowledge base" (1986: 396) that allows him or her to factor out contextual changes. Nevertheless, he refers to this explanation of sound change as purely phonetic and writes: "this account of sound change also locates the mechanism centrally *in the phonetic domain*" (Ohala 1993a: 263, italics mine).

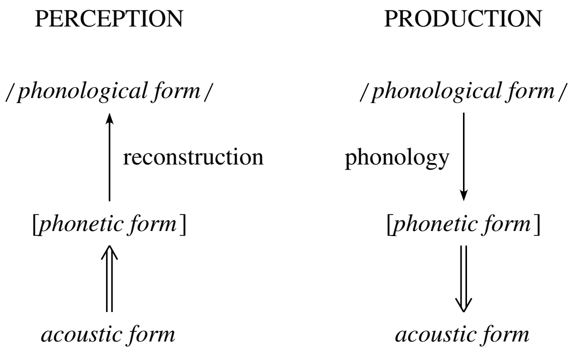


*Figure 3.* A depiction of Ohala's grammar modules (double arrows indicating automatic processes).

In line with the notation employed in this book, Ohala's grammar model can be depicted as in Figure 3. Perception is assumed not to involve phonological knowledge in this model. The double arrows from phonetic to acoustic form and vice versa indicate that this is an automatic process, described by Ohala as "is heard as". In this model, phonology maps phonological forms directly onto phonetic forms (without any intermediate phonological representation) and is restricted to the process of speech production, both in line with Chomsky and Halle (1968), as elaborated in Boersma and Hamann (this volume).

## 3. The present model: phonological speech perception

Based on its rule-like nature and its acquisition on the basis of elaborate knowledge of the language, Ohala's reconstruction module can be interpreted as phonological: It maps auditory information language-specifically to a phonological form. This mapping is only possible when we know what auditory information is of importance in the language under investigation and what phonological categories this language has.

Several studies on cross-language speech perception (e.g. Werker and Logan 1985 and Strange 1995) provide evidence for the language-specificity of the mapping from perceptual information onto phonological form. In psycholinguistics, this mapping is referred to as prelexical speech perception, and complemented by word recognition, where the phonological form is mapped onto a form in the lexicon, see Cutler and Norris (1979), Cutler, Mehler, Norris and Segui (1987), and McQueen and Cutler (1997). Boersma (2006 et seq.) employed such a two-staged speech recognition in the BiPhon model (but see already Boersma 1998), where the mapping between forms is formalised with OT constraints and occurs in parallel. In contrast to Ohala's model in Figure 3 and to earlier phonological models, BiPhon employs two phonological forms, namely a surface and an underlying (or: lexical) form. The surface form contains predictable information like foot structure and stress, whereas the underlying form is stripped from this, and includes only the information that has to be stored in the lexicon. The surface form is connected to the non-discrete phonetic form.

The linguistic modelling of speech recognition in BiPhon is complemented by a model of speech production, which employs the same forms and the same mappings, that is constraints, as the ones used in speech perception. The full BiPhon model is illustrated in Figure 4. Since the present article is concerned with the perception of speech sounds, we will focus on the perception grammar (boldface in Figure 4).

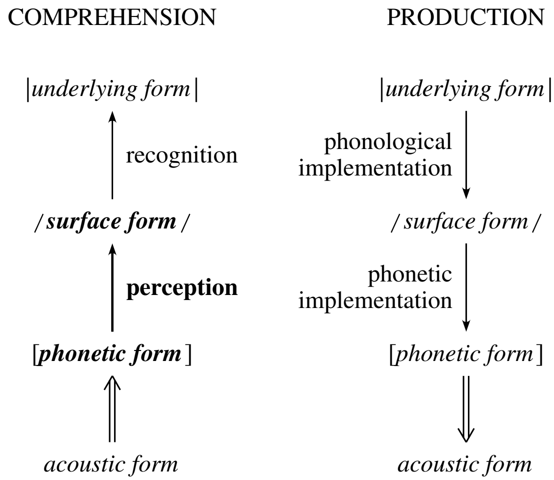


*Figure 4.* Boersma's BiPhon model, with the additional automatic processing of acoustic forms as auditory forms (double arrows at the bottom).

Before we move on to show how a perception grammar works and can account for sound changes, we have to deal with the claim made by opponents of a phonological view of perception that speech perception cannot be phonological because it is part of our general auditory capacity. Stating that speech perception is phonological does not imply that *all* perception is phonological. Instead, the present article proposes that speech perception *shares* perceptual abilities with general auditory (and visual) perception (see Fowler 1986), namely the ability to turn incoming acoustic data into processable auditory representations. This process is depicted by the double arrow at the bottom left of figure 4. Auditory and speech perception differ in that speech perception has as output phonological categories, i.e. it employs phonology, whereas auditory perception has a non-linguistic output.

Support for this difference between general auditory and language-specific perception comes from neurolinguistics. Best and Avery (1999) tested the perception of Zulu clicks by English and Zulu speakers in a dichotic listening task. Though the two groups had similar overall performance levels, only the Zulu listeners showed a left hemisphere advantage, which is typical for language-specific tasks (Kimura 1961; Studdert-Kennedy and Shankweiler 1970; among others). This result indicates that the Zulu listeners treated the clicks as phonological information, i.e. mapped the incoming data onto phonological categories, whereas the English listeners treated the clicks as general auditory information only. Further evidence comes from Dehaene-Lambertz *et al.* (2005), who found large differences in brain activity between auditory and speech perception. The fact that the perception of speech is processed with general cognitive as well as with language-specific means rebuts the argument by Hume and Johnson (2001b: 14–15) that speech perception cannot be inherent to language because it is part of a general cognitive module.

## 4. Sound change as the acquisition of a different perception grammar

The present section is concerned with perceptual cues and how they are mapped onto phonological categories in a perception grammar: §4.1 shows that sound changes do not simply involve a re-association of cues, §4.2 elaborates how the mapping from cues onto categories can be acquired, and §4.3 provides a formalisation of this mapping and its change over two generations of speakers with the example of Ohala's hypocorrection.

### 4.1 Sound change is more than a change in cue association

In "the perceptual basis of some sound patterns" (1995), Ohala describes the processes of hyper- and hypocorrection by referring to cues, or "phonetic events", as he also calls them (see also Ohala and Busà 1995). Hypocorrection can then be schematically represented as in Figure 5a (leaving aside the dotted lines for the movement): as a process whereby the listener parses a perceptual cue $b_1$ of the context (segment B) as belonging to segment A. In hypercorrection, a perceptual cue $a_2$ of a segment A is interpreted as belonging to the context (segment B), see Figure 5b. In this representation of hyper- and hypocorrection we can see that both processes involve perceptual cues that change their *association* from one segment to another.

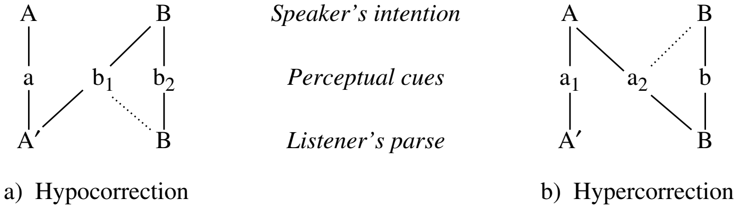


*Figure 5.* Hypo- and hypercorrection exemplified with cues, based on Ohala (1995: 90). A = original segment, A′ = changed segment, B = context.

Re-association of cues is a simplified view of sound change for several reasons. First of all, cues are often associated with more than one segment, that is, they are *shared*. The auditory event of a high F2 transition, for instance, can cue a coronal consonant but also an adjacent front vowel. A hypercorrection of an intended /yt/ as /ut/ is only possible because both /y/ and /t/ share this cue and learners might fail to associate the cue with /y/ since they can already associate it with the coronal. The sharing of a cue is illustrated in Figure 5 with the additional dotted lines: In Figure 5a, the cue $b_1$ is shared by the younger generation between the segment under investigation and the context, whereas in Figure 5b the cue $a_2$ was shared by the older generation between segment under investigation and context.

Secondly, sound change can only occur if some perceptual cues are given less importance by the listener, i.e. the younger generation, than by the speaker, i.e. the older generation, or are even totally *ignored*. An intended /yt/ can only be perceived as underlying /ut/ and vice versa if the listeners pay no attention to the F2 values at the beginning or middle of the vowel, where no coarticulatory transitions are observable, and therefore a clear differentiation between /y/ and /u/ is possible. After all, [yt] and [ut] are not confused by listeners of languages that contrast the two high rounded vowels, as mentioned in §2.

Thirdly, sound change is more gradual than a re-association of some cues (and ignorance of others) from one generation to the next suggests. It seems to develop from a change in the *weighting* of cues over generations. Learners acquire a different weighting because one cue is less reliable than it used to be for former generations, whereas others are more or similarly reliable. Such a situation emerges if categories start exhibiting variation along one perceptual dimension, as elaborated in §4.3 below.

Support for a language- and generation-specific difference in cue weighting comes from the fact that listeners do not give equal importance to all cues available to them (as shown e.g. by Dorman, Studdert-Kennedy and Raphael 1977 and by Whalen 1981). As Beddor and Krakow (1998) noted, diachronic change is therefore "not so much a breakdown or failure of normal perceptual processes, but rather a shift in the relative weighting of factors contributing to the linguistic percept" (p. 332).

## 4.2 The two-staged acquisition of a perception grammar

To acquire the mapping between perceptual information and phonological categories, infants can be assumed to pass through two consecutive stages. In the first stage, at the age of 6–8 months, the infants focus on a few, presumably the most salient, auditory dimensions (such as formant transitions, e.g., Nittrouer 1992). They keep track of the statistical distribution of items along these cue dimensions. Based on the statistical distributions, the infants construct language-specific phonetic categories. Psycholinguistic support for this stage is given for instance by Maye and Gerken (2000), Maye, Werker and Gerken (2002), and Pierrehumbert (2003).

In the second stage, the infants acquire labels for the learnt phonetic categories, and are guided in this by the lexicon. The lexicon informs the learners about the abstract categories necessary to distinguish words in the language and therefore has to be present at that stage. The abstract level created in the second stage also allows the learner to focus on more cues and to map several perceptual dimensions onto the same category (i.e., to integrate cues).

Psycholinguistic studies provide evidence for the two-stage development of perceptual capabilities in infants (e.g. Best 1993; Maye 2000; Stager and Werker 1997; Werker and Pegg 1992). Best, McRoberts and Goddell (2001: 791) describe that infants begin recognizing language-specific phonetic patterns and only later detect phonological classes, "perhaps in relation to increases in size of their early lexicon" (ibid.). The creation of phonetic categories, that is, the end point of the first stage in the acquisition of a perception grammar, explains why infants from nine months of age are no longer able to discriminate all phonetic contrasts but only those that occur in their native language(s) (Werker, Humphrey and Tees 1981; Werker and Tees 1984). Furthermore, it has been observed that infants and children often employ different cues than adults and seem to change their weighting of acoustic cues with more linguistic experience (e.g. Nittrouer Manning and Meyer 1993; Ohde and Haley 1997; Walley and Carrell 1983).[4]

Let us illustrate the two stages with a hypothetical language prone to undergo hypocorrection of /ut/ to /yt/. In this language, the /u/ tokens in coronal context differ from the remaining /u/ tokens with respect to the F2 transitions. We assume that the distribution of both types of tokens along this dimension is bimodal, as shown in figure 6. Infants learning this language hear tokens from this bimodal distribution and keep track of the probabilities of the tokens. On the basis of this input they will construct two separate phonetic categories, which we can term [ut] and [u]. In the second stage of learning, the infants deduct from form alternations with the same meaning that the items with F2 values between 1000 and 1200 Hz (i.e., [ut]) and those with F2 values between 600 and 1000 Hz belong to the same category /u/, and that the former are actually only contextual variants of the latter.

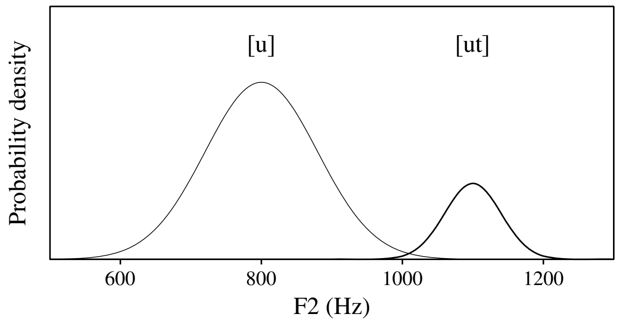


*Figure 6.* Assumed distributions for [u] and [ut] allophones along the dimension of F2 transitions.

A bimodal distribution of tokens alone does not trigger a sound change, as we know from the study by Harrington, Kleber and Reubold (2007), where the older generation of English RP speakers showed a similar distribution of /u/ and /ut/ as in Figure 6. This, however, only triggered a gradual shift of the /u/ tokens towards the /ut/ tokens for the younger generation of RP speakers, that is, the category merged again along this perceptual dimension, but no new phoneme /y/ emerged. A sound change in the shape of a split into the two catego-

4. Cues are often in a trading relation (e.g., Repp 1979), that is, a cue can receive more weight if certain other cues are not available or less prominent in specific contexts. This fact is not further incorporated in the present account.

ries /u/ and /y/ can only happen if the learners lack phonological or lexical information to associate both phonetic categories with one phonological category.

## 4.3 The formalisation of a perception grammar

The task of a perception grammar is to map each discrete auditory input onto an abstract phonological category. Within an OT framework, this mapping can be formalised with *cue constraints* (see Boersma 1998 et seq.; and Escudero and Boersma 2003 for the introduction of the term).

Let us take again Ohala's example of hypocorrection to illustrate the workings of cue constraints in a perception grammar. We recall that /ut/ changes to /yt/ because the younger generation uses the high F2 value in the VC transitions as a cue for both /t/ and a front rounded /y/ (the cue is shared), whereas the older generation employed it only to cue a coronal stop. To properly formalize this, we have to distinguish the cue of F2 values at the VC transition, such as [high F2]$_{VC}$, from the cue of F2 values in the middle of the vowel, such as [high F2]$_{V}$. In the language spoken by our older generation, a [high F2]$_{VC}$ is only a cue for a coronal consonant, expressed as the first cue constraint in (2a). The second cue constraint in (2a) is necessary for the correct perception of a back vowel.[5] It refers to the feature /±back/ but we could have used the features /±front/ or /±round/, instead.[6]

(2) *Cue constraints for Ohala's example of hypocorrection*

a. [high F2]$_{VC}$ /cor/: high F2 values in VC transitions cue coronals
[low F2]$_{V}$ /+back/: low F2 values in the vowel cue back vowels

b. [high F2]$_{VC}$ /+front/: high F2 values in VC transitions cue front vowels
[low F2]$_{V}$ /−front/: low F2 values in the vowel cue non-front vowels

To correctly formalise the perception of the younger generation, the additional constraints in (2b) are necessary. The younger generation has three high vo-

5. The vowel cues in this example are reduced to F2, as we are only interested in the difference in vowel backness and rounding. Further cues necessary for a complete description are F1 and duration. Other consonantal cues than the F2 transition are also not included, since the perception of /t/ as coronal voiceless stop is of no relevance for our example.
6. Backness correlates with rounding for the older generation in the cross-linguistically attested way with front unrounded and back rounded vowels. We have, however, no evidence that speakers of this language employ an additional feature /±round/ to specify their vowels.

wels instead of two. We thus need two features to distinguish these vowels. We will employ /±back/ and /±front/, but in contrast to traditional feature definitions (e.g. Chomsky and Halle 1968) we define /y/ as /–back, –front/, in order to keep the feature system small and to stay close to the acoustic correlate of F2 values. The first cue constraint in (2b) connects the auditory form [high F2]$_{VC}$ with the phonological class of front vowels, and together with the first cue constraint in (2a) it formalises the fact that one cue is shared by two segmental classes. The second cue constraint in (2b) mirrors the fact that the [low F2]$_{V}$ could also cue the additional high vowel category /y/ that this younger generation created. The four constraints in (2) are sufficient to describe the perception of [ut] by the younger listeners. A full formalisation of the perception grammar of both generations would of course require far more cue constraints than the ones in (2).[7]

As is evident from (2), cue constraints are not universal but depend both on the perceptual cues and the phonological categories that are employed in the respective languages. Thus constraints that refer to /±back/ and /±front/ are only used in a language with more than two distinctions in vowel backness; a language with a two-way distinction only employs one of the two features and the corresponding constraint(s).

In (3) we see perception tableaus for both generations of our assumed language. The input to these tableaus is the auditory form, and output candidates are abstract phonological forms that exist for each generation.

The perception tableaus, just like traditional OT tableaus, allow *Richness of the Base* (Prince and Smolensky [1993] 2004), as there are no language-particular restrictions on the input. In a perception tableau this means that a listener is able to perceive all possible auditory inputs (even though he/she is not always able to assign a linguistic form to them).

7. The cue constraints in (2) depart from the ones employed by Boersma (1998 et seq.) in three ways. Firstly, they involve relative cue values instead of concrete ones. This reduces the number of constraints considerably and implies that some kind of speaker normalisation has taken place. Secondly, the cue constraints are positively formulated, based on the assumption that a connection between an auditory form and a phonological form is only ever created if the learner has positive evidence for it. This again reduces the number of cue constraints enormously. Thirdly, the cue constraints refer to (mainly binary) features instead of segments. If the positive cue constraints were formalised in terms of segments, the highest ranked constraint would simply determine the output, even if lower-ranked cue constraints for e.g. less salient cues favour the non-winning candidate (see Boersma and Escudero 2004: 20–21 on the necessity of negatively formalised cue constraints for segmental categories). For instance, [low F2]$_{V}$ /u/ ≫ [low F2]$_{V}$ /y/ would never allow the candidate /y/ to win, unless [high F2]$_{VC}$ /y/ was highest ranked (compare to the current formalisation in tableau 3b).

(3) a. *Perception grammar of an older listener of Ohalaish*[8]

| [ut] [low F2]$_V$ [high F2]$_{VC}$ | [low F2]$_V$ /+back/ | [high F2]$_{VC}$ /cor/ |
|---|---|---|
| ☞ i. /ut/ | | |
| ii. /it/ | *! | |
| iii. /uk/ | | *! |

b. *Perception grammar of a younger listener of Ohalaish*

| [ut] [low F2]$_V$ [high F2]$_{VC}$ | [low F2]$_V$ /–front/ | [high F2]$_{VC}$ /–back/ | [high F2]$_{VC}$ /cor/ | [low F2]$_V$ /+back/ |
|---|---|---|---|---|
| i. /ut/ | | *! | | |
| ii. /it/ | *! | | | * |
| iii. /uk/ | | *(!) | * (!) | |
| ☞ iv. /yt/ | | | | * |

The perception grammars for the two generations differ in several points. They have different sets of output candidates: the younger listener in (3b) has constructed an additional phonological category /y/ (recall the mechanism described in §4.1 that leads to this), which results in the output candidate /yt/. This new category wins because the younger speaker also has a different, high-ranked cue constraint [high F2]$_{VC}$ /–back/, which formalizes that the listener uses the F2 transition as an important cue for a non-back vowel. This constraint rules out candidate (i) with /u/, which did win in the perception tableau of the older listener. The F2 transition is used by both generations to cue the coronal place of the consonant, and the respective constraint [high F2]$_{VC}$ /cor/ is violated by the candidates with non-coronal consonants (iii) in both tableaus.

Though not mentioned by Ohala, the low F2 in the vowel must be considered a very unreliable cue for back or non-front vowels by the young listener

8. If the older generation used the feature /±round/ in their phonology (recall footnote 6), we could add a candidate /y/ with the new feature combination /+round/ and /–back/ to tableau (3a), though this vowel itself does not occur in the language. Evidence for the emergence of new combinations of existing features comes for instance from loanword adaptation (Paul Boersma, p.c.). The candidate /y/ would violate a structural constraint */+round, –back/, because this featural combination does not surface in the language of the older generation. The interaction of such structural constraints (militating against non-occurring phonological structures) and cue constraints in the perception tableau illustrates again the phonological nature of speech perception.

(as indicated by the low ranking of [low F2]$_V$ /+back/ in 3b) to allow for the observed change in vowel quality from [ut] to [yt]. As discussed in §4.1 above, such an abrupt change in weighting is not likely to have occurred from one generation to the next. The F2 values of high vowels probably became a more and more unreliable cue over generations. This development might have started with coarticulation of the vowels according to consonantal context, which showed in slightly spread distribution of vowel tokens along the F2 dimension. Learners mirrored this distribution with their cue constraints (recall §4.2) by ranking the constraint [low F2]$_V$ /+back/ lower than the previous generation. The reuse of the same cue constraint ranking in their production grammar (this is one of the essential ideas in BiPhon) caused the learners to produce the same spread token distribution as they had observed in their input. In the production grammar, the cue constraints interact with articulatory constraints. If the language of these learners had high-ranked articulatory constraints that caused more coarticulation, the output tokens of this generation and input to the next would have had even more spread values along the F2 dimension. This gradual coarticulatory change continued until a generation of learners did not profit anymore from using F2 as a main cue to distinguish high vowels, and focussed on other cues instead.[9]

To summarise, an OT perception grammar as employed here expresses the fact that auditory information is not used homogeneously by all speakers across all languages, that no matter what speech sounds we hear we try to assign them abstract categories based on our language-specific knowledge, and that in sound change such language-specific knowledge changes across generations. In contrast to Ohala's account, sound change is not considered a misperception or a break-down in the communicative system. Instead, sound change is the learner's effort to construct the most efficient phonological system and a mapping from auditory form to this phonological system (i.e., a perception grammar) on the basis of the available input.

9. The interaction of cue and articulatory constraints in an OT grammar over several generations can account for the aforementioned phenomenon of segmental dispersion, see Boersma and Hamann (2008).

## 5. Two diachronic developments of retroflexes

In this section I illustrate that the reinterpretation of auditory information and its formalisation as perception grammars of two generations can account for real instances of hypocorrection. Two diachronic developments illustrating hypocorrection are discussed here, namely those in (4) (repeated from 1).

(4) *Retroflexion*

a. *in rhotic context:* *rt > ʈ Norwegian (North-Germanic)

b. *via labialisation:* *ʃʷ > ʂ˞ Minto-Nenana (Athabaskan)

Both processes involve changes from non-retroflex to retroflex coronals. As the accounts for these sound changes are based on the differences in the employment of acoustic information as perceptual cues between adult and language learner, §5.1 gives a short description of the acoustic properties of retroflexes and mentions other segments sharing these properties. Section 5.2 deals with the case of retroflexion in Norwegian (4a), and §5.3 with retroflexion in Minto-Nenana (4b). In these language-specific descriptions, the acoustic properties of the relevant segments are given in relative rather than absolute terms to ease the comparison between the two languages and between the different stages of one language. A further reason for relative descriptions is the lack of acoustic information on the former stages of the two languages.

### 5.1 Acoustic properties of retroflexes

Retroflex segments are generally distinguished from other coronals by lowered third formant (F3) transitions (both VC and CV; see Stevens and Blumstein 1975 and the summary of language-specific acoustic studies of retroflexes by Hamann 2003b: 59–60). Retroflex sonorants and vowels are furthermore characterised by an inherent low F3. Retroflex fricatives and affricates show a low and narrow frequency spectrum (see e.g. the centre of gravity measurements by Żygis and Hamann 2003 and Hamann and Avelino 2007). The location of F2 formants and formant transitions is the same for retroflexes and non-retroflexes according to Stevens and Blumstein (1975: 219), and seems to depend heavily on the quality of the adjacent vowels.

Retroflexes share the lowered F3 transitions (and for retroflex sonorants additionally the inherent low F3) with other segmental classes. Rounding in vowels like /u, y/ or labialisation in consonants has the effect of lowering F3 (Stevens 1998). A low F3 can also be found in velarised consonants or back vowels (Brosnahan and Malmberg 1970). Coronal trills show a low F3 (Lindau 1985), too, which might be due to the fact that they are always apical (Recasens 1991) and articulated with a low tongue middle and a retracted tongue back to enable the rapid movement of the tongue tip (see Recasens and Pallarès 1999 and Solé 1999 on the trill in Catalan).[10] This retracted tongue back, i.e. velarisation, causes lowering of the F3.

As we will see below, the similarities in low F3 between coronal rhotics, labialised sounds and retroflexes are responsible for the introduction of retroflex consonants into a number of languages.

## 5.2 Retroflexion of rhotic plus consonant sequences: Norwegian

In Urban East Norwegian, orthographic forms of <r> plus <t, d, s, n, l> in monomorphemic words are realised by the respective retroflex segments [ʈ, ɖ, ʂ, ɳ, ɭ], see the examples in (5a) (from Kristoffersen 2000). Historically, these forms were pronounced as sequences of rhotic trill plus apical consonant, see Haugen (1982a: 62). In this article, the present-day retroflex segments are assumed to be underlying, for a discussion of alternative views the reader is referred to Jahr and Lorentz (1981) and Hamann (2003b: 84–85).

(5) *Norwegian*

| | | | | | |
|---|---|---|---|---|---|
| a. | kart | [kɑʈ] | 'map' | | |
| | kors | [kɔʂ] | 'cross' | | |
| | barn | [bɑːɳ] | 'child' | | |
| b. | bror+s | [bɾuːʂ] | 'brother'possessive | cf. [bɾuːɾ] | bror |
| | vår+dag | [ˈʋoː.ˌɖɑːg] | 'spring day' | cf. [ʋoːɾ] | vår |
| | Per ser | [ˌpeː.ˈʂeːɾ] | 'Per sees' | cf. [peːɾ] | Per |

The examples in (5b) illustrate that Norwegian also has a synchronic process of retroflexion across morpheme and word boundaries. This synchronic process will not be dealt with in the following; for an OT account and a description of exceptions see e.g. Bradley (2007).

10. A low F3 does not hold for all coronal rhotics. Ladefoged and Maddieson (1996: 244) describe the retroflex approximants in Hausa and in Arrernte as both having high F3.

The diachronic process of retroflexion in the context of an alveolar rhotic emerged in the transition from Old to Modern Scandinavian (Haugen 1982a); the exact date of this event is difficult to determine (Eliasson 2005: 1124). The following two stages can be assumed:

(6) *Old Scandinavian* *Modern Scandinavian/Norwegian*
*rt, rd, rl, rs, rn → ʈ, ɖ, ɭ, ʂ, ɳ

The process in (6) is traditionally referred to as ‘postalveolarisation’, since the resulting segments are apical postalveolars and do not show a bending-backwards of the tongue tip usually found in the retroflex segments of the Indian subcontinent (see Hamann 2003b on the definition and occurrence of this type of retroflexion, and Simonsen, Moen and Cowen 2008 on the articulation of present-day retroflexes in Norwegian).

Norwegian also has a retroflex flap /ɽ/, the so-called ‘thick l’, which is widely agreed to stem from the sequence /rð/ (e.g. Larsen 1907: 70–72; Seip 1955: 177; Haugen 1976: 275). The emergence of this sound is not the topic of the present study, but the interested reader is referred to Molde (2005) for an articulatory explanation and an OT formalisation of this development. Besides postalveolarisation, central Scandinavian also developed a process of ‘cacuminalisation’, whereby heteromorphemic sequences of the so-called ‘thick l’ /ɽ/ plus following /t, d, s, n/ were realised as apical palatals with a bend-backwards tongue tip (Eliasson 2005: 1124). The older Oppdal dialect of Sør-Trøndelag still has both phonetic series of retroflexion, according to Haugen (1982a: 39–41). In most dialects, however, the results of ‘postalveolarisation’ and ‘cacuminalisation’ merged to apical postalveolars.[11]

Let us look at the perceptual cues that were probably involved in the Norwegian process. As elaborated in §5.1 above, coronal rhotics have low F3 values during their articulation but also at the transitions to neighbouring segments, fairly similar to the cues of a retroflex consonant. We can therefore assume that whereas Old Scandinavian listeners associated the low F3 values with the rhotic it occurred with, younger generations associated it with the following coronal segment. In Figure 7, the cue re-association is given in Ohala’s notation, with a voiceless stop as representative of the coronal class. Here, the low F3 values of the rhotic are reduced to the transition to the following coronal, transcribed as [low F3]$_{VC}$.

11. It has been suggested that the retroflexion of non-rhotic consonants in Middle Scandinavian was first introduced by the retroflex flap, and then this process was transferred to the alveolar trill, see e.g. Torp (1983: 73) and Torp and Vikør (1996: 72). Since there is no evidence for this assumption, the present study treats the process in (6) as an independent development.

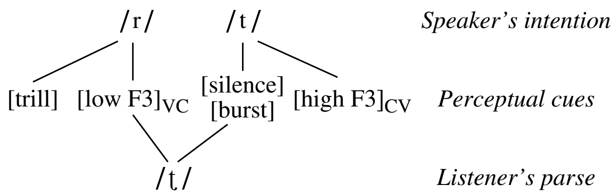


*Figure 7.* Development of retroflexion in Norwegian as re-association of low F3.

In Figure 7 we can see that the [low F3]$_{VC}$ which cued the rhotic is used by the younger generation to cue a non-rhotic coronal. Furthermore, the new learner ignores or gives less weight to a number of cues that were of significance to the speakers of the preceding generation, namely the rhotic manner cue (assumedly trill) and the place cue of the anterior coronal (high F3 transition from the consonant to the following vowel, i.e. [high F3]$_{CV}$). In (7) are the cue constraints necessary to formalise the change that took place across the two generations. To reduce the number of cue constraints, we employ only the silence as plosive cue (ignoring the burst), and the low F3 transition as rhotic cue (ignoring the trilling).

(7) *Cue constraints for Norwegian retroflexion of a plosive*

a. [low F3]$_{VC}$ /r/: low F3 values in VC transitions cue a rhotic
[high F3]$_{CV}$ /coronal/: high F3 values in CV transitions cue an (anterior) coronal (probably the only coronal series, and thus not further specified)
[silence] /plosive/: a silence cues a plosive

b. [low F3]$_{VC}$ /–anterior/: low F3 values in VC transitions cue a retroflex
[high F3]$_{CV}$ /+anterior/: high F3 values in CV transitions cue a non-retroflex

The cues in (7a) are the ones necessary to model the older generation, the ones in (7b) replaces the first two in (7a) for the younger generation.

The development of retroflexion in Norwegian is modelled with the perception grammars in (8) for two generations of Norwegian listeners. I assume that the older generation in (8a) had only one coronal series, and thus no feature such as /±anterior/ available to create a non-anterior, i.e. retroflex, coronal output candidate.

(8) a. *Perception grammar of an Old Scandinavian listener*

| [rt] — [low F3]$_{VC}$ [silence] [high F3]$_{CV}$ | [low F3]$_{VC}$ /r/ | [silence] /plosive/ | [high F3]$_{CV}$ /coronal/ |
|---|---|---|---|
| ☞ i. /rt/ | | | |
| ii. /t/ | *! | | |
| iii. /r/ | | *! | |
| iv. /rk/ | | | *! |

b. *Perception grammar of a Modern Norwegian listener*

| [rt] — [low F3]$_{VC}$ [silence] [high F3]$_{CV}$ | [low F3]$_{VC}$ /–anterior/ | [silence] /plosive/ | */r+coronal/ | [high F3]$_{CV}$ /+anterior/ |
|---|---|---|---|---|
| i. /rt/ | *(!) | | *(!) | |
| ii. /t/ | *! | | | |
| iii. /r/ | *(!) | *(!) | | |
| iv. /rk/ | *! | | | * |
| ☞ v. /ʈ/ | | | | * |

The learners of the younger generation in (8b) do not interpret any perceived input as a sequence of rhotic plus coronal. This is expressed in the structural constraint */r+coronal/. The same generation gives little weight to the cue of a high F3 in the CV transitions, expressed by the low ranking of the respective constraint. This cue might have been even ignored by this generation, in which case there would be no corresponding constraint at all. A cross-splicing perception experiment conducted by Hamann (2003a) provides psycholinguistic evidence for the low-weighting of [high F3]$_{CV}$. In the experiment, Norwegian listeners did not make use of the VC transitions but categorised coronals almost exclusively by their VC transitions; the latter is reflected in the perception grammar in (8b) by the high-ranked constraint on VC cues.

We can assume, like we did for the example of Ohalaish, that the Norwegian sound change developed more gradual and that there were more than two generations involved. The initial rhotic presumably caused a slightly more postalveolar articulation of the following coronal, leading to a more spread distribution in the coronal input to the following generation. This distribution was replicated with cue constraints by the younger generation, and due to the use of the same constraints and a high-ranked constraint requiring coarticulation in the production grammar, the subsequent generation received even more spread input. Over several generations, this lead to a bimodal distribution of post-rhotic versus post-vocalic rhotics, and since the two are distinguished by meaning in Norwegian, a new generation of learners assigned different abstract categories to them.

The development of retroflexion in rhotic context is cross-linguistically very common. We find it for instance also in Swedish (Eliasson 1986) and Faroese (Sandøy 2005)[12], in the Australian languages Watjarri (Douglas 1981) and Ndjébbana (McKay 2000), in Sanskrit (Whitney 1889), Sardinian (Bhat 1973) and the Iranian languages Yidgha and Munji (Skjærvø 1989). Tibetan languages show that besides apical alveolars also velars and labials can develop into retroflex in rhotic context, cf. the correspondence between orthography and pronunciation in the following words from Ladakhi, spoken in Jammu and Kashmir, India: *drug* 'six' [ʈuk], *phrugu* 'child' [ʈugu], *grodpa* 'belly' [ʈotpa] (Koshal 1982). A similar development can be found in the Tibetan languages Spiti and Lhoke (Grierson 1908).

## 5.3 Retroflexion of labialised sounds: Minto-Nenana

In Minto-Nenana, a Northern Athabaskan language spoken in Alaska at the Tanana River, retroflex fricatives and affricates were introduced via labialised segments (Krauss 1962; Tuttle 1998). The resulting segmental classes have rhotic releases, see (9a) for examples with fricatives (which show a voiced-voiceless distinction) and (9b) for examples with affricates (which are voiced, voiceless, or voiceless ejective).[13]

12. In contrast to Norwegian and Swedish, the retroflexion process in Faroese usually preserves the adjacent rhotic, with the exception of [ʂ], where it is always deleted (Barnes 2005: 1576).
13. Howe and Fullop (2005) discuss the diachronic development of sibilants in Athabaskan, which covers the processes investigated here. They provide an analysis in terms of the features [grave] and [flat], which they claim to be acoustic, not phonological as originally proposed by Jakobson, Fant and Halle (1952). Their proposal suggests that acoustic information is universally mapped onto acoustic features, which in turn are universally mapped onto phonological structures. This goes counter the argumentation and evidence given here that speech perception is language-dependent, i.e. that perceptual cues are weighted language-specifically and that the mapping between these cues and the phonological categories (segments or features) is language-specific.

(9) a. ʂʵa ‘sun/moon’
nəʐʵunh ‘it is good’

b. ʈʵ’exe ‘woman’
ʈʵəx ‘he cries’
ɖʵen ‘day’

The retroflex fricatives with rhotic releases in (9a) stem from the Proto-Athabaskan fricatives *ʃʷ and *ʒʷ, respectively, as illustrated in (10).

| (10) | *Proto-Athabaskan* | | *Minto-Nenana* |
|---|---|---|---|
| | *ʃʷ, *ʒʷ | → | ʂʵ, ʐʵ |

The origin of the retroflex affricate series (9b) is less clear, since the status of the relevant Proto-Athabaskan segments is under dispute. They are assumed to have been either labialised velar stops (Krauss 1973, 1979), labialised alveolar stops (Tharp 1972) (derived from the Proto-Athabaskan-Eyak *kʷ), or labialised postalveolar affricates (Cook and Rice 1989, 1989b; Krauss and Leer 1981), as depicted in the developments in (11a) – (11c), respectively.[14]

| (11) | *Proto-Athabaskan* | | *Minto-Nenana* |
|---|---|---|---|
| | a. *kʷ, *kʷ’, *gʷ | → | ʈʵ, ʈʵ’, ɖʵ |
| | b. *tʷ, *tʷ’, *dʷ | → | ʈʵ, ʈʵ’, ɖʵ |
| | c. *t͡ʃʷ, * t͡ʃʷ’, * d͡ʒʷ | → | ʈʵ, ʈʵ’, ɖʵ |

The present study follows Tharp (1972) in the assumption that (11b) is correct. This decision is based on the fact that (11b) involves a change from (front) coronal to back coronal and is most similar to the change occurring in the coronal fricatives in (10), which allows us to explain the two processes in parallel without any additional assumptions for the affricates.

In the development of retroflexes in Minto-Nenana, the younger generation re-interpreted the perceptual information in a way that departed in a number of points from the representations that the older generation had constructed. The labial release in [ʃʷ, ʒʷ] and [tʷ, tʷ’, dʷ] was taken by the younger listeners as a rhotic release because both share long transitions (henceforth: [trans]) and a

14. Some scholars (e.g. Leer 1996, 2005: 284 and Rice 2004: 324) assume that the development from labialised to retroflex affricates already took place in the development from Pre-Proto-Athabaskan to Proto-Athabaskan. Leer (1996: 212) includes the stem-final retroflex fricatives in this assumption.

low amplitude, both typical for glides (see Lehiste 1964), and also a lowered F3. The lowered F3 was furthermore interpreted as cue for the retroflexivity of the obstruent segment. For the fricatives, the low frequency friction noise caused by the labialisation was taken by the younger generation as additional cues for the retroflex nature of the fricative. These re-interpretations are graphically represented in Figure 8, with an alveolar labialised voiceless stop as representative of the class of segments that underwent the process (hence fricative-specific cues such as the friction noise are not included).

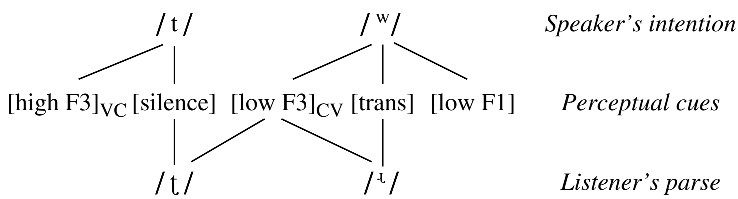


*Figure 8.* Development of retroflexion in Minto-Nenana as re-association and interpretation of a low F3.

In the diachronic development of retroflexion in Minto-Nenana, as in Norwegian, the low third formant of a segment (in this case: the approximant release) is associated with the adjacent front coronal, which is interpreted as retroflex. In contrast to Norwegian, the process in Minto-Nenana is not a re-association but a sharing of cues, since the low F3 is also used to cue the retroflex nature of the approximant release. Again, the change can only happen because the learner ignores or gives less weight to other perceptual cues: the high F3 transition into the alveolar stop and the low F1 typical for labialised approximants.

The following cue constraints are minimally necessary to model the change in Minto-Nenana, again restricted to the voiceless stop. The list in (12) does not include the low F1 values as cue for the labialised approximant release and the long transitions for any kind of approximant release. In (12a) are the constraints necessary for the older generation of Minto-Nenana, in (12b) those necessary for the younger generation.

(12) *Cue constraints for Minto-Nenana retroflexion*

a. [high F3]$_{VC}$ /coronal/: high F3 values in the VC transitions cue an (anterior) coronal
[silence] /plosive/: a silence cues a plosive
[low F3]$_{CV}$ /ʷ/: low F3 values in the CV transitions cue a labialised approximant release

b. [high F3]$_{VC}$ /+anterior/: high F3 values in the VC transitions cue an anterior coronal
[silence] /plosive/: a silence cues a plosive
[low F3]$_{CV}$ /–anterior/: low F3 values in the CV transitions cue a retroflex
[low F3]$_{CV}$ /$^{ɻ}$/: low F3 values in the CV transitions cue a (retroflex) approximant release

The perception tableaus in (13) show the application of these cue constraints in the modelling of the two generations of listeners.

(13) a. *Perception grammar of a Proto-Athabaskan listener*

| [t$^{w}$] [low F3]$_{VC}$ [silence] [high F3]$_{CV}$ | [low F3]$_{VC}$ /coronal/ | [silence] /plosive/ | [low F3]$_{CV}$ / $^{w}$/ |
|---|---|---|---|
| ☞ i. /t$^{w}$/ | | | |
| ii. /t/ | | | *! |
| iii. /k$^{w}$/ | *! | | |
| iv. /$^{w}$/ | *(!) | *(!) | |

b. *Perception grammar of a Minto-Nenana listener*

| [t$^{w}$] [low F3]$_{VC}$ [silence] [high F3]$_{CV}$ | [low F3]$_{VC}$ /–anterior/ | [silence] /plosive/ | [low F3]$_{CV}$ / $^{ɻ}$/ | [high F3]$_{VC}$ /+anterior/ |
|---|---|---|---|---|
| i. /t/ | *(!) | | *(!) | |
| ☞ ii. /ʈ$^{ɻ}$/ | | | | * |
| iii. /k$^{ɻ}$/ | *! | | | * |
| iv. /$^{ɻ}$/ | | *! | | * |

The candidate lists in both grammars only include segments or feature combinations that actually occur in the respective language, thus Proto-Athabaskan in (13a) does not have any retroflex sounds, and Minto-Nenana no labialised approximant releases. As a result, the cue constraints for Proto-Athabaskan involve unspecified but assumedly anterior coronals and labialised approximants, whereas that for Minto-Nenana employ retroflex and non-retroflex coronals and retroflex approximants.

The cue [high F3]$_{VC}$ is weighted very low or not considered at all by the listeners of Minto-Nenana, and therefore the respective constraint is ranked below the three decisive constraints in tableau (13b). We can assume again that the different weighting did not happen within two generations. It seems very likely that first a re-interpretation of the labialised release as retroflex took place (due to the weak cues involved), and at a later stage an assimilation of the obstruent to its retroflex release happened.

The retroflexion process of Minto-Nenana occurred cross-linguistically very seldom, a number of other Northern Athabaskan languages underwent it, namely Ingalik (or Deg Hit'an), Kolchan, Han, and Eastern and Western Kutchin (or Gwich'in) (Krauss and Golla 1981: 72). Smith River Athabaskan, an Athabaskan language belonging to the Pacific Coast subdivision, has only two retroflexes, a voiceless ejective affricate [ t͡ʂ'] and a voiceless fricative [ʂ], both seem to derive from the same Proto-Athabaskan sounds as the Northern Athabaskan retroflexes (with a collapse of the voicing distinction), see Bright (1964). A reason for the rarity of such a retroflexion process may be the typological unusualness of labialised coronals.

## 6. Conclusion

In the present paper I provided data from two diachronic changes that introduced retroflexes and argued that they happened because learners of both languages interpret the auditory information in a different way than previous generations. I illustrated that both changes belong to what Ohala (1981 et seq.) defines as "hypocorrection". Departing from Ohala's definition, I showed that these diachronic developments do not only involve some kind of re-association of perceptual cues but also a difference in cue weighting, where some cues are given much less (or no) weight by the younger generation than they received by the parent generation. I proposed that such a change in weighting occurs because some cues become less reliable due to variation in their distribution.

It was furthermore shown that the association of cues with phonological categories is an integral part of explaining speech perception and sound change. This association is not trivial, and involves both our phonetic and phonological knowledge. For this reason, I argued that speech perception is phonological and can be modelled by linguistic means, namely with a perception grammar in the linguistic framework of BiPhon (Boersma 2006 et seq.). The perception grammars for two generations of Norwegian and Minto-Nenana listeners yielded a straightforward formalisation of a difference in cue weighting and of its results for the categorisation of an incoming auditory signal.

Not only the mapping between cues and phonological categories were shown to differ from generation to generation. The phonological categories and features themselves were assumed to be emerging and not universal, and therefore also variable across generations. An assumption of universal categories and features, by contrast, raises the question whether the mapping between cues and such categories is also universal (this would not allow changes in mapping like the ones we provided evidence for) or, if it is not, how the learning infant can activate the correct universal category when confronted with varying cues in the input.

The reader might wonder whether the present model disposes of everything that might be universal in sound change. This is not the case. It is well-known that a large number of frequently occurring diachronic changes are due to universal phonetic characteristics such as the acoustic similarity between segments (e.g., Winitz, Scheib and Reeds 1972) or aerodynamic preferences (e.g., Ohala 1983). Such universals grounded in the vocal tract or the hearing apparatus can be included in a formal model of phonetics as general restrictions or preferences, and are not influenced in any way by the language-specificity of the phonological categories. This independence of phonetics and phonology is expressed in BiPhon by separate representations, and their interaction can be modelled by a parallel evaluation of the representations. The independence but cooperation of phonetics and phonology makes BiPhon resistant to criticism voiced repeatedly that phonological accounts of perceptually-motivated diachronic processes reduplicate phonetic information or incorporate phonetic markedness into phonology (see, e.g., Blevins 2004, Ohala 2005 and Howe and Fulop 2005[15]).

A remaining question is whether the present proposal of the listener as the initiator of sound change generalises from the discussed cases of hypocorrection to other types of diachronic developments. Not all sound changes seem to be initiated by the learning child. Let us look at two cases where adults seem to be actively involved. The first one is the well-known example of the British queen, who altered her pronunciation over the last 50 years (Harrington, Palethorpe and Watson 2000). We know that the queen did not initiate a sound change but instead slightly adjusted her own pronunciation to the Received Pronunciation of English spoken nowadays (see the illustration of the change in the articulation and perception of /u/ across three generations by Harrington, Kleber and Reubold 2007). This example only illustrates that even adult

15. Howe and Fulop (2005) argue "that sound change has its source in phonetics, not phonology […] and that phonetics and phonology are distinct modules, each with its privileged principles and elements, including features" (p.1). Their acoustic features [grave] and [flat], however, seem to duplicate the task of phonological abstraction within phonetics.

speakers continuously adjust the rankings of their cue constraints on the basis of the perceived input.[16] The actual sound change is the change in the input, which is to a large part produced by adults. A better example for sound change initiated by adults might be adult misperceptions, such as Ohala's hypercorrection elaborated in section 2. Adult misperceptions, however, do not usually lead to change. Listeners are used to misperceive all the time (due to background noise and other factors), but mostly correct their percept with the help of semantic context and the mental lexicon. Only highly infrequent words stand a chance of being permanently stored with a representation that departs from that of other speakers. Still, an idiosyncratic pronunciation of a highly infrequent word by a single speaker does not constitute a sound change. In order to be considered a sound change, such a pronunciation of a word has to spread in the population of speakers and to the pronunciation of other words, and this seems to involve again infant learners. It remains a topic for future studies to test whether adults alone can initiate and propagate a sound change.

## *References*

Barnes, Michael
2005 The standard languages and their systems in the 20th century II: Faroese. In: Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann and Ulf Teleman (eds.), *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, 1574–1584. Berlin: Walter de Gruyter.

Beddor, Patrice Speeter, and Rena Arens Krakow
1998 Perceptual confusion and phonological change: how confused is the listener? In: Benjamin K. Bergen, Madelaine C. Plauché and Ashlee C. Bailey (eds.), *Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society*, 320–334. Berkeley: Berkeley Linguistics Society.

Bermúdez-Otero, Ricardo
1999 *Constraint Interaction in Language Change: Opacity and Globality in Phonological Change*. Ph.D. dissertation, University of Manchester.

16. Updating the cue ranking on the basis of the input has less drastical effects the more experienced (i.e., older) the listener gets.

Best, Catherine T.
1993 Emergence of language-specific constraints in perception of non-native speech: a window on early phonological development. In: Bénédicte de Boysson-Bardies, Scania de Schonen, Peter W. Jusczyk, Peter F. MacNeilage and John Morton (eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year*, 289–304. Dordrecht: Kluwer Academic Press.

Best, Catherine T., and Robert A. Avery
1999 Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. *Psychological Science* 10: 65–69.

Best, Catherine T., Gerald W. McRoberts and Elizabeth Goodell
2001 Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109: 775–794.

Bhat, D. N. S.
1973 Retroflexion: an areal feature. *Working Papers on Language Universals* 13: 27–67.

Blevins, Juliette
2004 *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boersma, Paul
1997 How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21: 43–58.

Boersma, Paul
1998 *Functional Phonology: Formalizing the Interaction between Articulatory and Perceptual Drives*, Ph.D. dissertation, University of Amsterdam, The Hague: Holland Academic Graphics.

Boersma, Paul
2003 The odds of eternal optimization in Optimality Theory. In: D. Eric Holt (ed.), *Optimality Theory and Language Change*, 31–65. Dordrecht: Kluwer.

Boersma, Paul
2006 Prototypicality judgments as inverted perception. In: Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky and Ralf Vogel (eds.), *Gradience in Grammar*, 167–184. Oxford: Oxford University Press.

Boersma, Paul
2007 Some listener-oriented accounts of *h*-aspiré in French. *Lingua* 117: 1989–2054.

Boersma, Paul
to appear A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: Anton Benz and Jason Mattausch (eds.), *Bidirectional Optimality Theory*. Amsterdam: John Benjamins.

Boersma, Paul, and Paola Escudero
2004 Learning to perceive a smaller L2 vowel inventory: an Optimality Theory account. *Rutgers Optimality Archive* 684.

Boersma, Paul, Paola Escudero and Rachel Hayes
2003 Learning abstract phonological from auditory phonetic categories: an integrated model for the acquisition of language-specific sound categories. In: Marie-Josep Solé, Daniel Recasens and Joaquín Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 1013–1016. Barcelona.

Boersma, Paul, and Silke Hamann
2008 The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.

Boersma, Paul, and Silke Hamann
this volume Introduction: models of phonology in perception.

Bradley, Travis G.
2007 Morphological derived-environment effects in gestural coordination: a case study of Norwegian clusters. *Lingua* 117: 950–985.

Bright, Jane O.
1964 The phonology of Smith River Athapaskan (Tolowa). *International Journal of American Linguistics* 30: 101–107.

Broselow, Ellen
2004 Language contact phonology: richness of the stimulus, poverty of the base. *Proceedings of the North-Eastern Linguistic Society* 34: 1–22.

Broselow, Ellen
this volume Stress adaptation to loanword phonology: perception and learnability.

Brosnahan, Leonard Francis, and Bertil Malmberg
1970 *Introduction to Phonetics*. Cambridge: Heffer & Sons.

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

Cook, Eung-Do, and Keren D. Rice
1989 Introduction. In: Eung-Do Cook and Keren Rice (eds.), *Athapaskan Linguistics: Current Perspectives on a Language Family*, 1–62. Berlin: Mouton de Gruyter.

Cutler, Anne, and Dennis Norris
1979 Monitoring sentence comprehension. In: William E. Cooper and Edward C.T. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, 113–134. Hillsdale: Erlbaum Associates.

Cutler, Anne, Jacques Mehler, Dennis Norris and Juan Segui
1987 Phoneme identification and the lexicon. *Cognitive Psychology* 19: 141–177.

Dehaene-Lambertz, Ghislaine, Christophe Pallier, Willy Serniclaes, Liliane Sprenger-Charolles, Antoinette Jobert and Stanislas Dehaene
2005 Neural correlates of switching from auditory to speech perception. *NeuroImage* 24: 21–33.

Dorman, Michael F., Michael Studdert-Kennedy and Lawrence J. Raphael
1977 Stop-consonant recognition: release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics* 22: 109–122.

Douglas, Wilfrid H.
1981 Watjarri. In: Robert M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian Languages*, Volume 2, 197–273. Amsterdam: Benjamins.

Eliasson, Stig
1986 Sandhi in Peninsular Scandinavian. In: Henning Andersen (ed.), *Sandhi Phenomena in the Languages of Europe*. Berlin: de Gruyter.

Eliasson, Stig
2005 Phonological developments from Old Nordic to Early Modern Nordic IV: a typological and contrastive survey. In: Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann and Ulf Teleman (eds.), *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, 1116–1128. Berlin: Walter de Gruyter.

Escudero, Paola, and Paul Boersma
2003 Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In: Sundha Arunachalam, Elsi Kaiser and Alexander Williams (eds.), *Proceedings of the 25th Annual Penn Linguistics Colloquium* (Penn Working Papers in Linguistics 8.1), 71–85.

Escudero, Paola, and Paul Boersma
2004 Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551–585.

Flege, James E.
1987 The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65.

Fowler, Carol A.
1986 An event approach to the study of speech perception. *Journal of Phonetics* 14: 3–24.

Gess, Randall
1996 *Optimality Theory in the Historical Phonology of French*, Ph.D. dissertation, University of Washington.

Gess, Randall
2003 On re-ranking and explanatory inadequacy in a constraint-based theory of phonological change. In: D. Eric Holt (ed.), *Optimality Theory and Language Change*, 67–90. Dordrecht: Kluwer.

Green, Antony Dubach
1997 *The Prosodic Structure of Irish, Scots Gaelic, and Manx*, Ph.D. Dissertation, Cornell University.

Green, Antony Dubach
2001 The promotion of the unmarked: representing sound change in Optimality Theory. Manuscript, University of Potsdam.

Grierson, George
1908 *Linguistic Survey of India.* Volume III: *Tibeto-Burman Family*. Delhi: Motilal Banarsidass.

Ham, William H.
1998 A new approach to an old problem: gemination and constraint reranking in West Germanic. *Journal of Comparative Germanic Linguistics* 1: 225–262.

Hamann, Silke
2003a Is retroflexion licensed by VC cues only? Observations from Norwegian. *Nordlyd* 31: 63–77.

Hamann, Silke
2003b *The Phonetics and Phonology of Retroflexes*, Ph.D. dissertation, Utrecht University.

Hamann, Silke
2005 The diachronic emergence of retroflex consonants in three languages. *LINK: Tijdschrift voor Linguistiek te Utrecht* 15: 29–48.

Hamann, Silke, and Heriberto Avelino
2007 An acoustic study of plain and palatalized sibilants in Ocotepec Mixe. In: Jürgen Trouvain and William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, 946–952. Saarbrücken.

Hamann, Silke, and Susanne Fuchs
to appear Retroflexion of voiced stops: data from Dhao, Thulung, Afar and German. *Language and Speech* 53.

Harrington, Jonathan, Sallyanne Palethorpe and Catherine I. Watson
2000 Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association* 30: 63–78.

Harrington, Jonathan, Felicitas Kleber and Ulrich Reubold
2007 /u/-fronting in RP: A link between sound change and diminished perceptual compensation for coarticulation? In: Jürgen Trouvain and William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, 1473–1477. Saarbrücken.

Haspelmath, Martin
1999 Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18: 180–205.
Haugen, Einar
1976 *The Scandinavian Languages: An Introduction to their History*. London: Faber and Faber.
Haugen, Einar
1982a *Scandinavian Language Structures: A Comparative Historical Survey*. Tübingen: Niemeyer.
Haugen, Einar
1982b *Oppdalsmålet: Innføring i et Sørtrøndsk Fjellbygdmål*. Oslo: Gyldendal.
Holt, D. Eric
1997 *The Role of the Listener in the Historical Phonology of Spanish and Portuguese: An Optimality-Theoretic Account*, Ph.D. dissertation, Georgetown University.
Howe, Darin, and Sean A. Fulop
2005 Acoustic features in Athabaskan. Paper presented at the *Annual Meeting of the Linguistic Society of America*, Oakland.
Hume, Elizabeth, and Keith Johnson (eds.)
2001a *The Role of Speech Perception in Phonology*. San Diego: Academic Press.
Hume, Elizabeth, and Keith Johnson
2001b A model of the interplay of speech perception and phonology. In: Elizabeth Hume and Keith Johnson (eds.), 3–26.
Jacobs, Haike
1995 Optimality Theory and sound change. *Proceedings of the Annual Meeting of the North-Eastern Linguistic Society* 2: 219–232.
Jahr, Ernst H., and Ove Lorentz
1981 *Fonologi/Phonology*. Oslo: Novus.
Jakobson, Roman, Gunnar Fant and Morris Halle
1952 *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge (Massachusetts): MIT Press.
Kimura, Doreen
1961 Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology* 15: 166–171.
Koshal, Sanyukta
1982 *Conversational Ladakhi*. Delhi: Motilal Banarsidass.
Krauss, Michael E.
1962 Minto phonology. Manuscript, Alaska Native Language Center.
Krauss, Michael E.
1973 Na-Dene. In: Thomas A. Sebeok (ed.), *Linguistics in North America*, (Current Trends in Linguistics, Volume 10,) 903–978. The Hague: Mouton.

1979 Na-Dene and Eskimo-Aleut. In: Lyle Campbell and Marianne Mithun (eds.), *The Languages of Native North America: An Historical and Comparative Assessment*, 803–901. Austin: University of Texas Press.

Krauss, Michael E., and Victor K. Golla
1981 Northern Athapaskan languages. In: June Helm (ed.), *Handbook of North American Indians*, 67–85. Washington: Smithsonian Institution.

Krauss, Michael E., and Jeff Leer
1981 Athabaskan, Eyak and Tlingit sonorants. *Alaska Native Language Center Research Papers* 5.

Kristoffersen, Gjert
2000 *The Phonology of Norwegian*. Oxford: Oxford University Press.

Ladefoged, Peter, and Ian Maddieson
1996 *The Sounds of the World's Languages*. Oxford: Blackwell.

Larsen, Amund B.
1907 *Kristiania Bymål*. Kristiania: Cammermeyer.

Leer, Jeff
1996 The historical evolution of the stem syllable in Gwich'in (Kutchin/Loucheux) Athabaskan. In: Eloise Jelinek, Sally Midgette, Keren Rice and Leslie Saxon (eds.), *Athabaskan Language Studies: Essays in Honor of Robert W. Young*, 193–234. Albuquerque: University of New Mexico Press.

Leer, Jeff
2005 How stress shapes the stem-suffix complex in Athabaskan. In: Sharon Hargus and Keren Rice (eds.), *Athabaskan Prosody*, 277–318. Amsterdam: John Benjamins.

Lehiste, Ilse
1964 Acoustical characteristics of selected English consonants. *International Journal of American Linguistics* 30.

Liljencrantz, Johan, and Björn Lindblom
1972 Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48: 839–862.

Lindau, Mona
1985 The story of /r/. In: Victoria Fromkin (ed.), *Phonetic Linguistics: Essays in Honour of Peter Ladefoged*, 157–168. Orlando (Florida): Academic Press.

Lindblom, Björn
1986 Phonetic universals in vowel systems. In: John J. Ohala and J.J. Jaeger (eds.), *Experimental Phonology*, 13–44. Orlando: Academic Press.

Maye, Jessica
2000 *Learning Speech Sound Categories from Statistical Information*, Ph.D. dissertation, University of Arizona.

Maye, Jessica, and LouAnn Gerken
2000 Learning phonemes without minimal pairs. *Proceedings of the 24th Annual Boston University Conference on Language Development*: 522–533.

Maye, Jessica, Janet F. Werker and LouAnn Gerken
2002 Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.

McCarthy, John J., and Alan Prince
1993 Prosodic morphology I: Constraint interaction and satisfaction. Manuscript, University of Massachusetts and Rutgers University.

McKay, Graham
2000 Ndjébbana. In: Robert M. W. Dixon and Barry Blake (eds.), *Handbook of Australian Languages*, Volume V, 155–354. Oxford: Oxford University Press.

McQueen, James, and Anne Cutler
1997 Cognitive processes in speech perception. In: William J. Hardcastle and John Laver (eds.), *The Handbook of Phonetic Sciences*, 566–586. Oxford: Blackwell.

Mielke, Jeff
2003 The interplay of speech perception and phonology: experimental evidence from Turkish. *Phonetica* 60: 208–229.

Mielke, Jeff
2004 *The Emergence of Distinctive Features*, Ph.D. dissertation, Ohio State University.

Molde, Ann-Kristin
2005 *Innføringen av Tjukk l og Retroflekser i Østlandsk: En Empirisk Avklaring og Optimaitetsteroretisk Analyse*, M.A. thesis, Universitetet i Bergen.

Nittrouer, Susan
1992 Age–related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics* 20: 351–382.

Nittrouer, Susan, Carol Manning and Carol Meyer
1993 The perceptual weighting of acoustic cues changes with linguistic experience. *Journal of the Acoustical Society of America* 94: 1865.

Ohala, John J.
1974 Experimental historical phonology. In: John M. Anderson and Charles Jones (eds.), *Historical Linguistics II: Theory and Description in Phonology*, 353–389. Amsterdam: North-Holland Publishing.

Ohala, John J.
1981 The listener as a source of sound change. In: Carrie Masek, Roberta A. Hendrik and Mary Frances Miller (eds.), *Proceedings of the Chicago Linguistics Society* 17, 178–203. Chicago.

Ohala, John J.
1983 The origin of sound patterns in vocal tract constraints. In: P.F. MacNeilage (ed.), *The Production of Speech*, 189–216. New York: Springer.

Ohala, John J.
1986 Phonological evidence for top-down processing in speech perception. In: Joseph S. Perkell and Dennis H. Klatt (eds.), *Invariance and Variability of Speech Processes*, 386–401. Hillsdale: Erlbaum Associates.

Ohala, John J.
1989 Sound change is drawn from a pool of synchronic variation. In: Leiv Earl Breivik and Ernst Håkon Jahr (eds.), *Language Change: Contributions to the Study of its Causes*, 173–198. Berlin: Mouton.

Ohala, John J.
1993a Sound change as nature's speech perception experiment. *Speech Communication* 13: 155–161.

Ohala, John J.
1993b The phonetics of sound change. In: Charles Jones (ed.), *Historical Linguistics:Problems and Perspectives*, 237–278. London: Longman.

Ohala, John J.
1995 The perceptual basis of some sound patterns. In: Bruce A. Connell and Amalia Arvaniti (eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, 87–92. Cambridge: Cambridge University Press.

Ohala, John J.
2005 Phonetic explanations for sound patterns: implications for grammars of competence. In: William J. Hardcastle and M. Beck (eds.), *A Figure of Speech: A Festschrift for John Laver*, 23–38. London: Erlbaum.

Ohala, John J., and Elizabeth E. Shriberg
1990 Hypercorrection in speech perception. *Proceedings of the International Conference on Spoken Language Processing*, 405–408. Kobe.

Ohala, John J., and Deborah Feder
1994 Listeners' identification of speech sounds is influenced by adjacent "restored" phonemes. *Phonetica* 51: 111–118.

Ohala, John J., and Grazia Busà
1995 Nasal loss before voiceless fricatives: a perceptually-based sound change. *Rivista di Linguistica* 7: 125–144.

Ohde, Ralph N., and Katarina L. Haley
1997 Stop-consonant and vowel perception in 3- and 4-year-old children. *Journal of the Acoustical Society of America* 102: 3711–3722.

Pater, Joe
2004 Bridging the gap between perception and production with minimally violable constraints. In: René Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 219–244. Cambridge: Cambridge University Press.

Pierrehumbert, Janet B.
2003 Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–154.

Prince, Alan, and Paul Smolensky
2004 *Optimality Theory: Constraint Interaction in Generative Grammar*. London: Blackwell. First appeared as Technical Report of the University of Colorado Computer Sience Department [1993].

Recasens, Daniel
1991 On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics* 19: 267–280.

Recasens, Daniel, and Maria Dolors Pallarès
1999 A study of /flap/ and /r/ in the light of the ”DAC” coarticulation model. *Journal of Phonetics* 27: 143–169.

Repp, Bruno
1979 Perceptual trading relation between aspiration amplitude and VOT. In: Jared J. Wolf and Dennis H. Klatt (eds.), *Speech Comunication Papers*, 53–56. New York: Acoustical Society of America.

Repp, Bruno
2004 Language contact, phonemic inventories, and the Athabaskan language family. *Linguistic Typology* 8: 321–383.

Sandøy, Helge
2005 The typological development of the Nordic languages I: phonology. In: Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann and Ulf Teleman (eds.), *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, 1852–1871. Berlin: Walter de Gruyter.

Seip, Didrik Arup
1955 *Norsk Språkhistorie til Omkring 1370*. Oslo: Aschehoug, 2nd edition.

Simonsen, Hanne G., Inger Moen, and Steve Cowen
2008 Norwegian retroflex stops in a cross linguistic perspective. *Journal of Phonetics* 36: 385–405.

Skjærvø, Prods
1989 Modern East Iranian languages. In: Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicum*, 370–383. Wiesbaden: Reichert.

Solé, Maria-Josep
1999 Production requirements of apical trills and assimilatory behavior. In: John J. Ohala, Yoko Hasegawa, Manjari Ohala, D. Granville and Ashlee C. Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, 77–94. San Francisco.

Stager, Christine L., and Janet F. Werker
1997 Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388: 381–382.

Steriade, Donca
1995 Positional neutralization. Unpublished manuscript, University of California, Los Angeles.

Steriade, Donca
2001 Directional asymmetries in place assimilation: A perceptual account. In: Elizabeth Hume and Keith Johnson (eds.), 219–250.

Stevens, Kenneth N.
1998 *Acoustic Phonetics*. Cambridge (Massachusetts): MIT Press.

Stevens, Kenneth, and Sheila Blumstein
1975 Quantal aspects of consonant production and perception: a study of retroflex stop consonants. *Journal of Phonetics* 3: 215–233.

Strange, Winifred
1995 Cross-language studies of speech perception: a historical review. In: Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 3–45. Timonium: York Press.

Studdert-Kennedy, Michael, and Donald Shankweiler
1970 Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America* 48: 579–594.

Tharp, George
1972 The position of Tsetsaut among the Northern Athapaskans. *International Journal of American Linguistics* 38: 14–25.

Torp, Arne
1983 Nokre tankar om inndeling av folkemålet i Skandinavia. *Folkmålsstudier* 28: 45–92.

Torp, Arne, and Lars S. Vikør
1996 *Hovuddrag i Norsk Språkhistorie*. Oslo: Gyldendal.

Tuttle, Siri
1998 *Metrical and Tonal Structures in Tanana Athabaskan*, Ph.D. dissertation, University of Washington.

Walley, Amanda C., and Thomas D. Carrell
1983 Onset spectra and formant transitions in the adult's and child's perception of place. *Journal of the Acoustical Society of America* 73: 1011–1022.

Werker, Janet F., and Richard C. Tees
1984 Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development* 7: 49–63.

Werker, Janet F., and John S. Logan
1985 Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37: 35–44.

Werker, Janet F., and Judith E. Pegg
1992 Infant speech perception and phonological acquisition. In: Charles A. Ferguson, Lise Menn and Carol Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*, 285–311. Timonium: York Press.

Werker, Janet F., John H. V. Gilbert, Keith Humphrey and Richard C. Tees
1981 Developmental aspects of cross-language speech perception. *Child Development* 52: 349–355.

Whalen, Doug H.
1981 Effects of vocalic transitions and vowel quality on the English /s/ - /ʃ/ boundary. *Journal of the Acoustical Society of America* 69: 275–282.

Whitney, William Dwight
1889 *Sanskrit Grammar*. Cambridge: Harvard University Press.

Winitz, Harris, M. E. Scheib, and James A. Reeds
1972 Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *Journal of the Acoustical Society of America* 51: 1309–1317.

Wright, Richard
2001 Perceptual cues in contrast maintenance. In: Elizabeth Hume and Keith Johnson (eds.), 251–277.

Żygis, Marzena, and Silke Hamann
2003 Perceptual and acoustic cues of Polish coronal fricatives. In: Marie-Josep Solé, Daniel Recasens and Joaquin Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 395–398. Barcelona.