The Origins of Markedness
Synchrony and Diachrony

Marc van Oostendorp

May 2, 2005

These are class notes for an undergraduate course I taught together with Clara Levelt at the University of Leiden in the spring of 2005; the notes concern my part of the class only. The students had to read one article in preparation of every class; it will be clear which article this was from the notes.

Contents

1 Introduction

1.1 'Beyond explanatory adequacy'

Consider the following Optimality Theoretic markedness constraint (Prince & Smolensky, 1993):

(1) NoCoda: Syllables do not have codas.

Like most OT constraints (or, as a matter of fact, most phonological generalisations proposed in any framework) this is a markedness tendency rather than a true linguistic universal in the sense that every language obeys it completely. This constraint serves several purposes at the same time. Most importantly, it expresses:

1. the fact that open ('CV') syllables are universal in human language (there is no language which disallows them), while closed syllables are allowed only in a subset
2. the fact that even those languages which do have open syllables, tend to avoid them: E.g. (tautomorphemic) VCV presumably is universally syllabified as V.CV, not as *VC.V.

Every modern theory of phonology uses some mechanism which expresses these generalisations:

- In rule-based theory (Clements & Keyser, 1983), it is assumed that a rule syllabifying CV is basic and universal, whereas rules syllabifiying postvocalic consonants are language-specific and apply in a later module

- In purely representational approaches, like (certain versions of) Government Phonology (Scheer, 2004), it is assumed that CV is the only available syllable type. Something that phonetically looks like CVC phonologically really is $CV_1.CV_2$, where V_2 is an empty vowel, subject to a number of specific constraints and (therefore) marked.

In this course, however, we will concentrate on Optimality Theoretic constraints such as the one in (1). If we use NOCODA in some analysis, we may say that this constraint 'explains' a certain set of facts. For instance, some OT analyses of French liaison will use this constraint to 'explain' those facts. The following tables are an example (from Féry, 2003)

(2) a.

/pəti(t)/ pinson 'finch'	NOHIATUS	NOCODA
a. ☞[pəti] pinson		
b. [pətit] pinson		*

b.

/pəti(t)/ aigle 'eagle'	NOHIATUS	NOCODA
a. [pəti] aigle	*	
b. ☞[pətit] aigle		

The constraints NOHIATUS and NOCODA are part of an explanation of liaison in these tableaux to the extent that they (i) are independently motivated, and (ii) help to derive the observed results.

At the same time, constraints such as these are obviously themselves in need of an explanation: why are our constraints (or our rules, or our representations) the way they are and not otherwise? In the case at hand, why is there constraint NOCODA and not, alternatively, a constraint CODA? In the terminology of Chomsky (2004), we are going 'beyond explanatory adequacy' if we try to answer these questions: we try to explain why UG is shaped the way it is.

It turns out that the answer to this question is very much dependent on our idea of the place of phonology within linguistics, or its relation to phonetics:

1. We might assume that these constraints are 'grounded' in the phonetics. E.g., we know that obstruents, and more specifically stops, are harder to articulate and perceive after a vowel than before it. This gives a motivation for NOCODA, whereas CODA is quite absurd.
2. Alternatively, we might try to find an explanation in the way in which cognitive structures are realized. For instance, we may try to relate the fact about syllable structure to the idea of Kayne (1994) that the syntactic structures of all languages are SVO. If both subjects and syllable onsets are linguistic 'specifiers', we have discovered some similarity to

the two. Obviously, still the question needs to be answered *why* specifiers occur on the left-hand side. Under this view, it might even be possible that coda's are more difficult to perceive, because human beings know that they are less prominent Carstairs-McCarthy (1999).

If we take the 'grounded' position, we have to deal with the question how exactly the phonetics can influence phonology. Here there seem to be roughly three positions:

1. Constraints such as NOCODA are part of Universal Grammar. The problem with this account is one of *duplication*: we have to assume that NOCODA somehow is part of the 'outside world' — the speech signal, the auditory system — and at the same time of the 'inside world' — the innate capacities of human beings. A reason for this might be evolutionary: the language system has adapted over time to the way in which language is used. But it is unclear that there has been enough evolutionary time to get to this point. There is no clear representative of this position ('nativists' seem to usually prefer a cognitive point-of-view).
2. The second option is to assume that the language-learning child constructs constraints such as NOCODA on the basis of what she observes in her own speech and speech errors. The child thus acts as a small experimental phonetician (or 'laboratory phonologist'). This approach has been defended by Hayes (1999), among others (see below).
3. The third option is that these constraints are not part of grammar at all. Their explanation has to be sought elsewhere, and the most obvious place to look is the diachrony: languages change because of misperception or misarticulation, and when children acquire the language, they simply pick up whatever centuries of phonetically initiated change have made out of the system.
   Presently, there seem to be two paths to reach this conclusion. One is by assuming that phonology is only about 'hard universals', hence not about markedness. Phonology is a pure cognitive symbolic system in which there is no place for statistical tendencies. Since virtually no principle in phonology is 'hard', this means we void the theory from many modules that used to be part of it. This is the position defended most forcefully in Hale & Reiss (2000); Hale (2003) and related works.
   More or less the same conclusion has been reached by authors such as Bybee (2001), starting from the assumption that "language is a self-organizing system, and grammar, including both morphosyntax and phonology, is an emergent property of that system" (p. 190). In this view, phonology is all about statistical tendencies, and symbolic systems, if they exist at all, are seen as epiphenomenal. The child just acquires whatever is available, and this material will have been largely subjected to the principles of language use.

> The two positions converge in the sense that most traditional objects of study for the phonology are delegated to a component of statistics. They may differ as to the role of language acquisition: the cognitive view will suppose that the phonetic facts still have to go through a filter of UG (be it one which is much leaner than traditional grammar has it), whereas Bybee (2001) basically assumes a blank slate model of the human mind.

All in all, we can see that the current discussion on markedness has connections to many very basic questions in linguistic theory, such as: what is the division of labour between synchronic and diachronic explanation? What is the role of language acquisition in linguistic change? And how are phonetic and phonological explanations of phenomena to be related?

1.2 The life cycle of sound patterns

The traditional point of view of the phonology-phonetics interface can be summarised as follows.[1] We assume that language change has its origins in phonetics; this origin will most likely (or in some models exclusively) be in the direction of greater ease of articulation and/or ease of perception. After a while, the results of this phonetic change may first become *phonologized*, and later *morphologized* or even *lexicalized*. Bermúdez-Otero (2005) traces this idea back to Baudouin de Courtenay (1895) and summarizes the 'life cycle of sound patterns' as follows:

- *Phase I*
  The life cycle begins when, by Neogrammarian sound change, some physical of physiological phenomenon gives rise to a new cognitively controlled pattern of phonetic implementation. This development, known as *phonologization* (Hyman, 1976), involves the addition of a new phonetic rule to the grammar.
- *Phase II*
  Subsequently, this gradient sound pattern may become categorical. [...][S]uch a change would involve the *restructuring* of the phonological representations that provide the input of the phonological implementation, with the concomitant development of a new phonological counterpart for the original phonetic rule.[...]
- *Phase III*
  Reanalysis can also cause categorical patterns to change. Over time, phonological rules typically become sensitive to morphosyntactic structure, often with a reduction in their domain of application [...] Phonological rules may also develop lexical exceptions [...]

[1] This section relies heavily on Bermúdez-Otero (2005).

- *Phase IV*
  At the end of their life cycle, sound patterns may cease to be phonologically controlled. Thus a phonological rule may be replaced by a morphological operation (*morphologization*), or may disappear altogether, leaving an idiosyncratic residue in lexical representations (*lexicalization*). [...]

As Bermúdez-Otero points out, this view of sound change fits very well into the standard generative view of the synchronic relation between phonology and phonetics, as it is exemplified in models such as Lexical Phonology and Stratal OT. amd which can be summarised as follows:

(3)
Lexical representations
(categorical)
↓
[Phonological rules]
↓
Phonological representations
(categorical)
↓
[Phonetic rules]
↓
Phonetic representations
(gradient)

Under this view, then, sound change moves 'bottom-up' in the grammar: a change which originates in the phonetics may in the course of time end up having an effect only in certain lexical representations.

Although Bermúdez-Otero (2005) does not discuss this point, notice that this view as a matter of fact implies that the explanation of markedness is essentially in the realm of phonetics, because this is where every rule or process will start its life cycle. On the other hand, the process of phonologization (which turns gradient phonetic facts into categorical phonological ones) will be in part the product of phonology. The phonology will then be responsible for the 'universal' aspects and the phonetics for the 'markedness aspects'.

This might be easiest to see within a rule-based framework. Let us suppose that a language L at some point in its history will be subject to a phonetic change by which word-final consonants are gradually reduced. At some point, this might become phonologized to something like:

(4) $C \rightarrow / \_\_ \# C$

What has happened, at this point is that a gradient reduction has turned into something categorical: now word-final consonants are deleted completely.

he same consonant may still show up in some other environment, e.g. before a vowel-initial word. The fact that this rule is a 'natural' rule (an implementation of NoCoda, so to say) is a consequence of the fact that it has originated in the phonetics, however. The only reason why rules such as (4) are (much) more frequent than rules such as (5) is that the latter does not have a plausible phonetic origin; from a purely phonological point of view, there is nothing wrong about it.

(5) C → / __ # V

At the same time, the effects of the phonetics may become obscured in due course, because new rules might follow this one. And then at some point, the rule may become lexicalized: it just happens that some words alternate word-final contexts according to context. (This may be the case of French liaison, which does not affect new words.)

As simple and elegant as this picture may be, there also are a few problems connected to it. Most importantly, it implies that the grammar of every generation is built on the basis of that of previous generations by addition of rules at the end of the grammar. One conceptual problem for this is that this means that every generation of language learners has to be able to see into the heads of their parents directly in order to see the grammars represented there (Hale, 2003).

Also, it is not very clear how this idea can be made compatible with Optimality Theory, precisely. On the one hand, Bermúdez-Otero (forthcoming) shows that the idea of phonologization/lexicalization can be explained more elegantly in OT than it could in rule-based theory, because of a principle of Lexicon Optimization — we will not go into that here. On the other hand, the only thing that can be manipulated in (classical) OT is constraint rankings. This then leaves the source of the constraint NoCoda still unexplained: if it is in the universal constraint set Con, how did it get there in the first place? If this constraint mirrors a phonetic generalisation, how does it do that? The only possibility would be, in fact, to say that Con contains all kinds of constraints, including NoOnset and Coda, and that the only reason why we do not see the latter is that they are unlikely to ever take effect. (And obviously there is always the alternative of rejecting the thought that the origin of sound change should always be phonetic in nature.)

In recent years, an alternative to the traditional view of sound change and the phonetics-phonology interface has been proposed under the rubric of 'exemplar theory'. In this view, lexical items are not categorical — let alone underspecified. Rather, language users store individual phonetic soundshapes of tokens into their memory. These tokens, which are often referred to as 'exemplars', are associated to each other because they are of course phonetically very similar. But in the extreme versions of this theory, they are not categorized in any way. There is no independent phonological representation of a

given word, there is only a network of individual tokens and 'emergent generalisations' (cf. the quote from Bybee (2001) above). Actually semantically, phonetically and morphologically related words will also be connected, albeit with looser ties.

One consequence of this theory is that it predicts that there is no independent phonology: if generalisations can be made, they are due to phonetics, or processing, or other considerations. There is no grammar for individual languages, so by extension there can be no Universal Grammar, and indeed there can be no phonological universals (apart maybe from a few hard restrictions imposed on us by the vocal tract etc.) Also the existence of productive phonological alternations is effectively denied. Final devoicing in Dutch for instance is presumably 'represented' by the fact that all singulars of nouns end in voiceless obstruents, and some plurals have a voiced obstruent in a corresponding position.

Another consequence is that language change can only be gradient and lexically diffusing. The reason why it can only be gradient is because there is no categorial phonology, everything is gradient phonetics. The reason why it can only be lexically diffusing (i.e. affecting item by item, not taking one sound in the language and change it in every place where it occurs) is because there is no such concept of 'a consonant following a vowel': there only are individual occurrences of consonants following vowels in individual tokens of words. This means that the network of words may change in the direction of less and less prominently pronounced coda consonants, but there is no particular reason why the networks of other words should move in the same direction at the same time.

All of this obviously means that the whole idea of the life cycle of phonological rules is completely abandoned, which is a little bit too radical for many scholars, as may be the idea that there are no truly phonological alternations. For this reason, more moderate versions of this approach have been proposed, e.g. by Pierrehumbert. In any case, all of this shows that the studies of markedness, historical phonology and the phonology-phonetics interface are strongly intertwined. By studying them together, we may get a better view on each of them individually. This is what we hope to achieve in this course.

2 'On the need for a separate perception grammar'

2.1 Functionalism in phonology

Paul Boersma is one of the most well-known (and most interesting) modern European representatives of a school of thought which he calls *Functional Phonology* (cf. Boersma, 1998). The guiding hypothesis is that "languages are built according to functional principles of efficient communication. For

speaking, these functional principles are minimization of articulatory effort and minimization of perceptual confusion." (p. 1) The notes that follow do not give a summary of Boersma (1999); rather they give additional information on three topics related to this paper: the notion 'functionalism', the role of perception in phonology, and innateness.

Note that it is implicitly assumed that 'efficient communication' is the central (maybe the only relevant) function of language. Since Jakobson (1960) it is usually assumed the following functions can be assigned to human language:

1. *emotive* function, expressing the attitudes or moods of the speaker
2. *conative* function, directing the listener towards some action, a change of attitude, etc.
3. *referential* function, transferring information from one head to the other
4. *phatic* function, keeping the line of contact open between speaker and listener
5. *poetic* function, exploring the properties of language (rhythm, syntactic structures)
6. *metalingual* function, discussing the features of other utterances

It seems clear that some of these functions are more 'important' than others, in particular that the metalingual function is somehow derivative of other functions (it is unlikely that somebody would come up with a theory which would establish that the metalingual function is the most important one). However, it is not a priori clear what the central function is.

By declaring the *referential* function to be primary — which is what 'functionalists', including Boersma, tend to do — we also implicitly take a position on the evolution of language in the human species: it will have developed in order for human beings to convey information in an efficient way. How would this have been the case? It is not so clear that an individual has an evolutionary advantage in conveying the information which he has to others (unless maybe, these are his close relatives). This would then point in the direction of *group selection*: those groups of individuals which can share information are those which are most likely to succeed; hence also the individuals within these groups have the highest chances of reproductive success. As far as I know, this is indeed the line of argumentation that most functionalists would accept. (The other functions may then be seen as epiphenomenal given the complexity of human society.)

It is worth pointing out that in principle other lines of thinking are possible. For instance, Hauser *et al.* (2002) argue that language may be an 'exaptation' of evolution: something which accidentally arose as the result of independent changes, in which case it is obviously not clear why this accident should then have properties which make it particularly suitable for communication.

To take another example, one might take either the poetic of the phatic function as central, and then explaining the evolution of language in terms of so-called *sexual selection*, a term introduced by Darwin, but recently explored e.g. Miller (2000). The idea is that individuals sometimes can profit from displaying features which seem completely useless for survival. The reason that they are useless shows that there owners should have 'good genes', which allow them to display such frivolous extras. A famous example is the peacock's tail, which actually makes it more difficulat for the male peacock to survive. But females can then start using this as a diagnostic for good partners.

Language has a few properties very similar to the peacock's tail, as Miller points out (in an internet interview on `http://www.edge.org/3rd_culture/miller/`):

> Of course the principal way that people court each other is through language.[...]
>
> A project I'm very excited about at the moment is trying to understand why humans have such large vocabularies. The average human knows about a hundred thousand words by adulthood. That requires memorizing arbitrary patterns that relate sound to meaning. It requires memorizing ten words a day every day from age 18 months to age 18 years, and that's a fantastic feat of learning. There's nothing else like it that humans do. The funny thing about that vocabulary is how little of it we use in ordinary conversation. We get by in our day to day speech with just a few thousand words counting for 95% of all the words that we say.
>
> There's are a tremendous number of words we've learned that are not used very often but that we bothered to memorize, that don't seem to be very useful in ordinary day to day life, but that we still sometimes use with each other and those are the words that I want to explain — not the 5,000 most useful words but the 95,000 ornamental words. My prediction is, people mostly use them in courtship. They use them essentially to show off, they use them to show how bright they are, how good their learning ability is, how good their memories are for words. We know in the brain where these words are remembered, roughly in Werneke's area, in certain parts of the left hemisphere, we know that there's specialized brain machinery for learning these words; we know that vocabulary size is an extremely powerful indicator of intelligence — this is why vocabulary items are used in IQ tests; within a few minutes of conversing with somebody you use the vocabulary that they're producing as a pretty good indicator of how intelligent they areso it's a extremely useful thing to use in mate choice. The hypothesis here is that vocabulary size itself has been strongly shaped by sexual selection, and that most of the words

> that we know have been learned not because they're useful for survival, but because they're useful for courtship.

Also within phonology, it is not hard to find elements which do not seem to make sense, and which somehow seem to run counter to the "principles of minimization of articulatory effort and minimization of perceptual confusion": so-called 'crazy rules', all kinds of lexical exceptions, and also the phenomenon of opacity. Note, by the way, that Miller also concedes that:

> language is extremely useful for many functions. You can tell your friends how you're going to hunt an animal and cooperate on tracking it down. Women can tell their friends where the best roots and berries and tubers are growing this season. Parents can tell their kids all sorts of useful information as they're growing up.

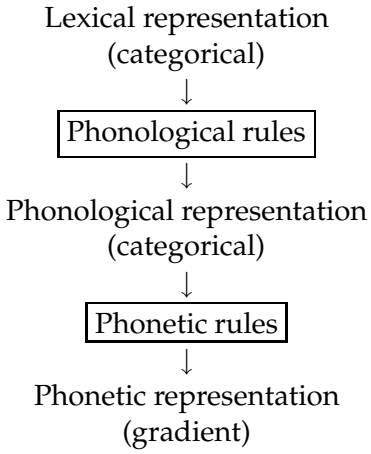
These functions again can be summarised under Jakobson's rubric of the 'referential' function of language. However, under this view it is not so clear that the referential function alone would shape the structure of language. But in this connection the following claim of Boersma (1999) is also relevant (p. 18):

> A related empirically testable claim is that all universal phonology must be directly functional and that all arbitrary phonology must be language-specific, i.e. that there are no arbitrary substantive universals in phonology. This is a rather strong claim, since it is immediately falsified as soon as anyone comes up with a substantive universal that has no direct functional explanation. The evidence to date, however, shows that the reverse falsification is more likely to occur: all proposed substantive phonological universals seem to have directly functional exceptions, i.e., phonology seems to be functionally perfect.

2.2 The role of perception

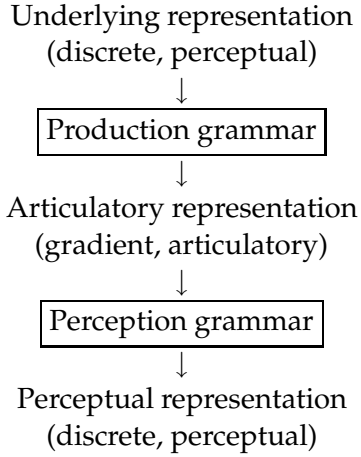
The traditional view of the phonological derivation can be seen as an abstraction of the *production* of human speech for at least two reasons. First, the arrows in the derivation usually point from underlying (lexical) forms to the surface; this is roughly how production would work (the following figure is from the handout of two weeks ago).

(6)



Secondly, (most of) the phonological features are (usually) defined or definable in articulatory terms rather than articulatorily ([coronal], [spread glottis]). In Boersma (1999)'s proposal, there is a grammar for the 'speaker' (or 'production') and also one for the 'listener' (or 'perception'). Still, for most practical purposes we can concentrate on the speaker. The reason for this is that it is assumed that the speaker pays attention to the listener as well: when she has reached an acoustic output for her underlying form, she tries to establish whether the underlying representation will still be reconstructible from this by the listener. All in all we have the following in the speaker's production (simplifying certain matters somewhat, cf. p. 10):

(7)



The model thus involves different representations of different kinds, and faitfulness relations of some sort could hold between all of them, but the *real* faithfulness is between the underlying representation and the output, because these are of the same nature (they are both discrete and also both perceptual). Boersma (1999, p. 10) notes that it is a problem of standard generative grammar (and standard monostratal OT in particular, we may add)

that it assumes that the grammar consists of only a (lexical, discrete) input and a (phonetic, gradient) output. But it is not clear how we could establish realistic 'faithfulness' relations (requiring them to be basically the same) between these two levels, if they are so dissimilar to begin with.

Apart from several metatheoretical arguments, Boersma (1999) also adduces a few empirical points. For instance, he notes (p. 6) that Dutch /r/ can be pronounced in various ways (including [r] and [ɾ]), but this does not seem to affect its phonology. His claim is that these various realisations do not have anything in common from an articulatory point of view, but they are similar from the point of view of perception (they are all trills).

Note that this argument may be in need of refinement, since Dutch /r/'s can be realized in many more ways, and not all of them are obvious trills. But in actual fact, at least one of these realisations shows an additional property of this framework which Boersma (1999) does not discuss in this paper (although he has mentioned it in other work). We mean the 'zero' realisation, where the /r/ is not pronounced at all. This is of course problematic for the specific claim made on page 6 (there is no trill present at all in this realisation), but not necessarily for the framework as a whole. Note that in many cases, there is at least still some trace of the 'underlying' /r/:

(8)		'underlying form'	surface form
	after long vowels	*beer* 'bear' [beːr]	[bɪː]
	after short vowels	*kers* 'cherry' [kɛrs]	[kɛːs]

The long vowels turn from tense to lax before /r/, a well-known but little-understood process which is usually called *colouring*; furthermore long lax vowels only occur in this position in Dutch. Underlyingly short (lax) vowels may lengthen in this position. Both the colouring and the lengthening are preserved even if the realisation of /r/ is zero.

We will now concentrate on colouring, but a similar reasoning will hold for lengthening. Notice that in traditional terms the colouring is 'opaque': there is no surface reason why it would have to take place. Colouring is a process which thus should have affected the form at a different level. This is actually very complicated in traditional terms, because this would mean either (i) that r-drop is phonetic and colouring is phonological (a very undesirable assumption, because colouring is completely allophonic and phonetic and has all other kinds of properties of phonetic processes) or (ii) we have to introduce some complicated mechanisms into our theory which can take care of this opacity.

In Boersma (1999)'s model, the problem is easier to solve, if we assume that colouring means the transferral of some perceptual property of /r/ onto the vowel — which is presumably what we will want to say in any case. This means then that an output [bɪː] is in some straightforward sense more faithful to the input than an alternative output *[beː], while at the same time

satisfying some markedness constraint on the articulation of r's at the ends of syllables.

At the same time, it should also become clear that the framework presented here is extremely powerful. It involves many different representations, and many different constraints. Furthermore, these constraints may originate both in the perception and in the articulation. We have an 'explanation' of some phenomenon ($\alpha > \beta$) as soon as we can show that α is 'better' than β in some phonetic dimension. But there is a problem here — which will return in several classes to come — which is that phonetics is not very restrictive. There is no real phonetic theory, and it is actually hard to find non-trivial phenomena which would be completely excluded by phonetics. It is true that we know that certain things are more likely to happen than others, and it is true that we can do experiments to test hypotheses. But there is as yet no restrictive phonetic theory which could tell us a priori which phenomena we can expect and which ones are completely impossible. One could argue that a real explanation should be couched in terms of a theory, and a real theory should make predictions and exclude certain states of affairs.

Lurking in the background we also still find the problem of 'crazy', non-natural processes and constraints, which Boersma would probably subsume under the heading of "arbitrary facts of the language, plus generalizations, also expressed as constraints." ((36b), p. 18). But if we assume that these will have to be learnt to, it is a question whether the child could not treat all phenomena as 'crazy', including those which happen to have a phonetic explanation. This leads us back again to the problem of what exactly is innate in phonology.

2.3 The role of innateness

As in other schools of phonetically-driven phonology, Boersma (1999) still assumes that the phonetically grounded constraints will end up in the individual grammar. His list of 'what's innate in phonology' is virtually the same as what we find in Hayes' article discussed last week:

(9) *What's innate in phonology* (according to Boersma (1999))
   a. *Peripherals* of speech production (versatile tongue and larynx) and auditory perception (spectrum, periodicity, noisiness, temporal coincidence and ordering, and intensity).
   b. Cognitive capabilities: *categorization* of perceived entities into classes of partial equivalence, *abstraction* of simultaneity relations and sequential relations into higher-level constructs, wild *generalization* and extrapolation, and the *manipulation* of arbitrary symbols (storage, retrieval, access).

c. *Decision making*: stochastic constraint grammars and a gradual learning algorithm.

d. *Functional drives*: the desire to understand and make oneself understood, and laziness.

The 'innateness' of (9a) would presumably not be contested by anybody, since it basically states the way in which the brain controls the body (nobody will assume that this is learned). (9b), (9c) and (9d) give the phonology proper.

- (9b) is the residue of phonological representation and presumably the derivation (assuming that the architecture of the grammar which Boersma gives on p. 1 is also part of these cognitive capabilities). This is really just a skeleton, voided of all substance. For instance, it could be the formalism of 'binary features', without the actual content of those features; and Boersma (1999) also refers to the mechanism of (autosegmental) tiers.
  An argument for not considering the substance of features to be part of universal phonology is that "Sign languages use different features than spoken languages." Notice that this does not solve the problem why within the set of spoken languages and within the set of sign languages we still find many similarities. Also, as far as we know, there are no language which seem to *mix* spoken language and sign language features. In other words, this argument does not lead us directly to the extreme conclusion that *no* substance whatsoever can be innate or even universal.
- (9c) basically describes the Optimality Theoretic architecture of the system (or Boersma (1998)'s flavour of this). The child obviously still has to know that it has to set up an OT system, and not e.g. a set of SPE-type rewrite rules. But the suggestion is that systems such as these might make up more general cognitive capacities, beyond phonology or even beyond grammar.
- (9d) might be the most controversial part of Boersma (1999)'s claims. The "desire to understand and make oneself understood, and laziness" should not only be part of the child's inherent psychological make-up, the language learner should also be able to somehow use these 'drives' in setting up her constraints.
  A potential problem for this assumption is that it will have to deal with some type of 'doubling', if we assume that language change is often provoked by (ungrammaticalized) 'laziness' etc. (note that Boersma (1999) explicitly mentions Passy (1891), a study on 'changements phonétiques', as a source of inspiration). If the diachrony can already explain the data as they are, it is not clear why the child would still have to act as a phonetician in order to set up 'plausible', 'natural' constraints.

3 Phonetic implementation

3.1 'Word-specific phonetics'

We can study the division between phonology and phonetics from a number of perspectives:

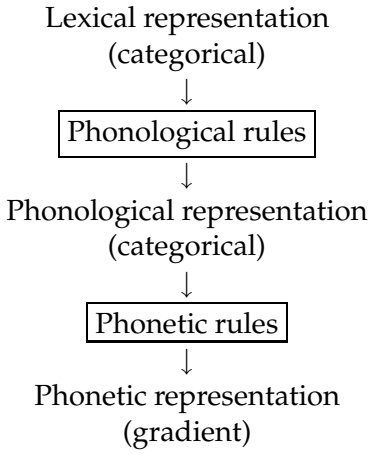
- From a metatheoretical point of view: do we need phonological explanations for phenomena which can also be explained phonetically, or vice versa? (cf. this course, *passim*)
- From the point of view of the analyst: how do we decide whether a given phenomenon should be given a phonological or a phonetic explanation? (Cf. Myers' article last week).
- From the point of view of language acquisition and diachrony: how does phonetic markedness get into the individual grammar (cf. Hayes (1999))
- From the point of view of the architecture of the grammatical model: how do phonology and phonetics interact within the grammar? What is the nature of 'phonetic interpretation'? (cf. (Boersma, 1999))

The different questions are of course interrelated; if somebody has decided that the answer to the first question is that there is no separate phonology, the answers to the following questions become trivial, for instance.

Pierrehumbert (2002) is mainly an attempt to deal with the last issue, although it has things to say about the other issues as well. The model presented in this article assumes that there are indeed two separate modules of grammar, and it even adopts a view of the organization of the grammar that is in some sense fairly traditional, viz. what Pierrehumbert calls a 'modular feedforward' model of phonetics. Yet she tries to combine it with something which is often seen as a radical departure of that model, viz. exemplar theory. Because her concern is with the architecture of the grammar, and she views the grammar as a realistic model of processing and perception, she makes more references to the psycholinguistic literature than other authors (and hardly any to literature on articulation or acoustics).

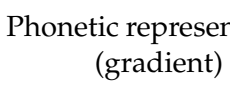
Here is a picture of the traditional (once again):

(10)



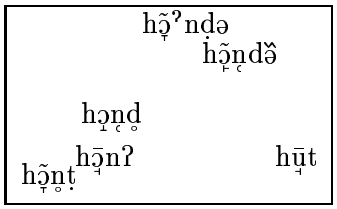
We have already seen that in essence this model has been with us for over a century now. It is contested by those who want to abandon a separate phonological representation, and in particular by radical 'exemplar'-based theories. Those theories are essentially monostratal; they can be pictured as follows:

(11)



In as far as there are any regularities, they are 'emergent': the phonetic representations of individual items stand in a relation of mutual correspondence, so that we get the following picture (or a very small fragment thereof):

(12)



Correspondence relations connect all of these forms: the closer they are pictured here, the thicker the lines (I did not include lines depicting these relations, since they make the picture even more obscure.)

Both models have advantages and Pierrehumbert (2002) tries to combine them. Starting from (10), Pierrehumbert (2002)'s proposal suggests the following emendations:

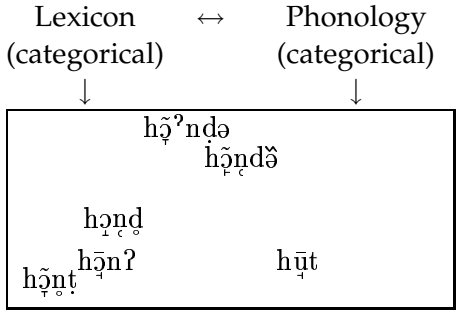
1. The picture in (10) suggests that phonetic and phonological 'rules' are qualitatively of the same type. They take one representation and change things on the basis of this one representation alone. This is not true for Pierrehumbert (2002): there is no explicit theory of phonological rules

here, but they are presumably cognitive, discrete and abstract. On the other hand, the phonetic 'rules' are statistic devices incorporating frequency, sociolinguistic, stylistic and other effects (including the effect of adaptation to the last speaker you heard), i.e. they are of the type in .

2. In the traditional model, phonetics is always across-the-board by definition: it does not have access to lexical information, but only to the phonological output (which itself has been subjected to Bracket Erasure). It is furthermore even assumed that the phonetics is 'universal': a given phonological representation will be pronounced in the same way evebn independent of the language. If there are language- or word-specific differences, they should be already present in the phonology.
   One of Pierrehumbert (2002)'s main points is that phonetics can be, to some extent, word-specific (hence also speaker- and language-specific). This means that information about phonetic implementation can also be stored.
   One of the most spectacular examples of this is the experiment of Goldinger (2000), which ingeniously demonstrates that people tend to be influenced in the details of their specific pronunciation by the way they have heard other people pronounce it recently (or even five days ago). But this means that indeed pronunciations of individual lexical items are stored, if not in the lexicon, then at least in some part of memory which can be reached from the lexicon:

All in all, the following is my reconstruction of how Pierrehumbert (2002) might picture the grammar:

(13)



This picture is too simple in at least three ways: first, as in other exemplar-based theories, the representations in the phonetics cannot really be matched to IPA symbols (since these are still categorical, even if they draw more distinctions than a normal phonological representation; they are stored forms of purely phonetic representations. Secondly, the lexicon and the phonology define exemplar spaces which are to some extent independent. Phonological representations define the exemplar space of e/g/ all /i/'s, or all underlying /in/ sequences. The lexicon has competing lexical spaces for e.g. the words *seen* and *limousine*. It is not defined how this competition should work ex-

actly, even though certain parts of the theory are formalized in mathematical equations.

3.2 What is a phonological fact?

As we stated above, Pierrehumbert (2002) wants to build a model which is psycholinguistically real: it should be able to explain the results of psycholinguistic (as well as sociolinguistic) experiments. Those results, rather than native speaker intuitions, data acquired from fieldwork, etc., count as the primary data which the theory has to account for. Furthermore, it is assumed that one model accounts for both perception and production facts. (Pierrehumbert counts as one of the originators of the Laboratory Phonology programme, and we can now see why.)

Pierrehumbert (2002) for (and against) modular feedforward are mostly based on data of this sort:

- From a number of experiments it appears that there is a 'phonological buffer': phonological structures are first built up before they are interpreted phonologically. This explains speech errors, but also the fact that people seem to take time to compute predictable information, rather than retrieving everything from lexical storage, and that "assembly of longer phonological plans takes longer than assembly of shorter plans".
- The phonetic realisation of e.g. intonational or even lexical tones can only be computed if we know the context in which these occur: a specification for high tone means that the tone is high *relative to the context in which it occurs*, not that it is pronounced at so-and-so many Hz. This only makes sense if the lexical storage does not contain information about the absolute height, but only something more abstract, such as a feature [+High Tone]. Something similar might be true for segmental allophony in general.

There are also arguments of a different kind, however:

- We can distinguish logically between two types of knowledge: *declarative memories*, i.e. lexical/phonological representations ("I know *that* the word *man* consists of one stressed syllable") and *procedural memories*, i.e. phonetic know-how ("I know *how* to pronounce /m/: by closing my lips and then letting the air escape through the nose"). "The mere fact of this distinction provides a basic argument for a modular theory" (Pierrehumbert, 2002, p. 3). This we might call an **ontological argument** — because we can logically distinguish between two things, therefore it follows that they are distinct. It is not very clear what the force of this argument is: because we can distinguish between phonology and phonetics, we should do so?

- "A related fact is the outcome of Neogrammarian sound changes (which enter the language as allophonic processes and may eventually become fossilized across the entire vocabulary)." Across-the-board changes such as these are hard to capture in a model which does not allow for any kind of abstraction. If we need to say that e.g. in a language all plosives have changed into corresponding fricatives, we need to be able to refer to abstract categories such as plosives and fricatives, and we need to be able to also formalize the equally abstract notion "corresponding". This is a **typological argument** – we need to be able to explain patterns (of language change in this case) which occur frequently in natural language. (More on this particular argument, which might also be problematic, in section 3.3.)

Similarly, the arguments for exemplar theory are mostly drawn from the psycholinguistic stock:

- The pronunciation of the vowel in CVC words depends on 'neighbourhood density'. The neighbourhood of a word W are the words which are similar to it (*man* has in its neighbourhood the words *ran, mac, mine*, etc.) If this neighbourhood is dense — if it contains many words — people tend to make a more 'extreme' version of the vowel. So an /i/ in a dense neighbourhood word is pronounced higher, more fronted etc. than an /i/ in a different kind of word. In a classical modular feedforward theory, this would have to mean that such words have a different feature value, which is unlikely. (This might remind you of the difference between English and Dutch /i/ which we discussed last week and which likewise might be reduced to the fact that English has more front vowels — but not in ordinary feature theory.)
- Similarly, there is more reduction in frequent words than in infrequent words. We do not want to say that such words have different phonologies, but that means that the phonetic implementation needs to have access to lexical information (it needs to be able to discriminate between words, and it needs to be able to count the frequency of certain words). Exemplar theory gives us both.
- We already mentioned the Goldinger (2000) experiment, showing that speakers are influenced in their production of words by recent memories of other speakers pronouncing the same words.
- Phonetic details sometimes seem sensitive to morphosyntactic structure. For instance, there is more glottalisation before the *a* in *realign* (morphologically related to *align*) than before the *a* in *realize*.

Also in this case, we can also find arguments from outside the laboratory:

- Language change sometimes is sensitive to lexical frequency: certain changes occur earlier in higly frequent words than in less frequent words

(Bybee, 2001, to be discussed next week). They are also sensitive to lexical class: in Québec French, some vowel shift affected all words, except for thos "representing organs of the church, the military and the schools". It is hard to evaluate this argument if we do not really see how language change is modelled in this theory. This is again a typological argument; we return to it in section 3.3 below.

- Phonetic implementation may be due to sociolinguistic factors. E.g. "Oprah Winfrey displayed the ability to shift her speech style between a more AAVE influenced style to more mainstream style, depending on the subject matter she was speaking about"; or "Her Majesty Queen Elizabeth II [...] gradually shifted her pronunciation in the direction of the Southern British English which has become fashionabel with younger speakers" in radio speeches over the past decades. This is a **sociolinguistic argument** but not a very strong one, given the fact that most formal theories trying to combine sociolinguistics and grammatical theory – such as variable rule theory — have been couched in a modular feedforward framework.

We thus see that the non-psycholinguistic arguments which Pierrehumbert (2002) provides are never very strong. Yet they also do not argue against this particular model, it is just the case that they are not very decisive.

3.3 Implications for language acquisition and change

Given that the questions we posed at the outset of this class are interrelated, it may be helpful to study Pierrehumbert (2002)'s from some of the other dimensions of the phonology-phonetics relation.

Let us first consider the case of Her Majesty Queen Elizabeth II. Apparently, she has gradually changed her phonetic implementation over the year, and she did this in the general direction of society surrounding her. Supposing that she and Oprah Winfrey are representative for humanity as a whole, this means that an individual can acquire new phonetics during her lifetime, and also that at least at the phonetic level a language can change for the whole community over the course of years. (This has of course been well-known in the sociolinguistic literature for a long time, see for instance Labov (2001); the fact that Pierrehumbert (2002) does not refer to studies of this type is significant for the sociology of the field of linguistics.) These facts can be gracefully modelled within exemplar theory.

One implication of this is that there is no *critical period* for language acquisition, and that is indeed the consequence of radical versions of exemplar theory, but it also seems not to be in accordance with many things we know about how language acquisition works. Pierrehumbert (2002)'s hybrid model might solve this problem to some extent. We could assume for

instance that the critical period holds for the cognitive/categorical aspects grammar (hence for phonology) but not for the fluid facts of phonetics.

This might still not be completely satisfactory, since even phonetic implementation is not always completely adapted to new circumstances. Pierrehumbert (2002)'s model would predict, for instance, that somebody who has lived in region A for the first 10 years of his life, and then moved to region B to live there for 30 years, without any contact with the original region, will have lost all phonetic traces of his dialect of origin. This may happen to some people, but it does not seem to happen to everybody (from the sociolinguistic literature just mentioned we know for instance that there is a difference between men and women). Yet, in spite of problems like these, the hybrid feedforward/exemplar model seems to be giving a reasonable approximation.[2]

Notice on the other hand that there is no worked out theory about the interaction between phonology and phonetics in language acquisition — as there was in Hayes (1999). It is not clear how the child is going to decide what is phonological and what is phonetic in the data she hears around her. The main reason for this unclarity is of course that this article does not present a worked-out theory of phonology. (This can never be a fatal criticism of course: we cannot expect one author to cover everything in a 20-page article.) Exactly what is in the phonology, other than an inventory of phonemes, and how this relates to the stuff phonetics is made of, is a matter to be sorted out.

Similar remarks apply to the theory of language change if seen from Pierrehumbert (2002)'s model. We have seen that various diachronic arguments are presented in order to support the model presented here. Yet we can only evaluate such arguments if we understand how language change would work. Roughly speaking we can mention the following properties of the two component theories of which this model is a hybrid:

- Feedforward models seem to be particularly good in describing processes of Neogrammarian sound change, i.e. sound changes which are across-the-board and affect individual sounds rather than individual words
- Exemplar models seem better in describing processes of lexical diffusion in which some individual words change, after which other words may follow, etc.

Given that this is a hybrid model, we might hope that it would be able to describe both phenomena, but it is not very clear (to me) whether it is as succesful in describing Neogrammarian sound change as more traditional models are. In the case of across-the-board sound change, the question arises how

[2]It should be noted that the mathematical formulas used in this article give a sense of exactness which is not really empirically justified; it also seems hard to test whether they are completely correct, given the absurdly complicated nature of the experiments that would have to be involved. But on the other hand, we could derive experiments from this, which could then test at least certain aspects of the formulas involved.

'phonologization' works: at some point the phonetic alternations have to enter the phonology, but in order to do this they have to completely change their nature in this model. Furthermore, if a change is indeed across-the-board, it means that it should be categorical, i.e. phonological, in this model. But that runs counter to the traditional idea that such sound changes originate in the phonetics and become phonological at the moment in which they also tend to become less 'across-the-board'.

Pierrehumbert (2002) can be seen as the culmination point of present thinking on the phonology-phonetics interface. It needs to be complemented though by at least a theory of phonological representations, and a theory of diachrony. The latter will our the topic for the next few weeks.

4 The force of grammar in sound change

4.1 The Neogrammarians and theory

The Neogrammarians — 'Junggrammatiker', 'Leipziger Schule' — continue to be a source of inspiration for every theoretically inclined historical linguist. This is so much the case that people with radically opposing views in the present debate can all refer to the Neogrammarian heritage (e.g. the generativist Kiparsky, the 'substance-free' phonologist Hale, the sociolinguist Labov, the 'evolutionary phonologist Blevins — they cannot all be right at the same time, yet they all claim to be heirs to this tradition).

It is no exaggeration to say that the relative 'scientific' prestige that linguistics has among the humanities is due in large part to Neogrammarians. They introduced 'hard' methods into the study of language, bringing rigid logical thinking to bear on it; and they have influenced almost all important linguists of the 20th century and beyonf (including Saussure, Chomsky, Labov). 19th Century linguistics became a prestigious field also outside the humanities. Darwin, for instance, stated in his book *The Descent of Man* that:

> The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel.

And after this he quoted 15 parallels between language change and biological evolution (Labov, 1993).

Like many successful branches of humanities in the 19th century, neogrammarian linguistics was primarily historical. The principles that were discovered were historical ('diachronic') principles, as we will see. We will also see what might have been the largest problem for this type of linguistics: the fact that the ontological status of its object of study, 'language', was not clear. One could describe how language changed, but not what it was, exactly, that changed when language changed.

However this may be, "two assertions [...] are assumed in modern literature on change to be the hallmark of Neogrammarianism: sound change is *regular* and *purely phonetically conditioned*" (Hale, 2003, p.343). Apparent exceptions are due to later sound changes, to 'analogy', to lexical borrowing from other languages after the sound change took effect, from 'minor sound changes' etc.

Hale's program is one of trying to find the cognitive 'core' of language and language change, by 'purifying' the data from those factors which can and ought to be explained by other factors. Some of these distinctions are routinely made by most phonologists, even if they have not been made explicit. For instance, we distinguish between:

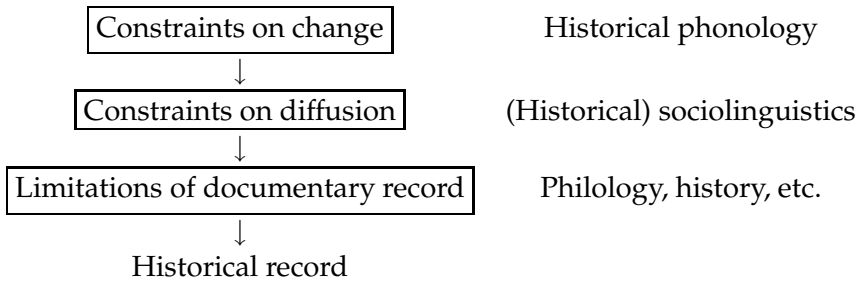
1. the **initiation**: some individual X starts to construct a phonological form α for something which is β for other members of the community.
2. the **diffusion**: the form α spreads through the community to the expense of β.

The study of diffusion of sound change is usually considered to be the task of sociolinguistics; the initiation phase was the business of Neogrammarians and also of most present-day historical phonologists. However, not everybody seems aware of the fact that for virtually every sound change we know has been diffused (otherwise it is very hard to even become aware of the fact). But this sometimes means that factors which are considered to be 'constraining' the initiation as a matter of fact are only relevant in the diffusion stage. There may be many possible initiations which never arise, simply because they cannot be diffused.

The model is slightly more complicated, because even diffusable sound changes are not all known to us, because for most historical studies we are limited by the accidental properties of the media (written sources) which transmit the crucial information. It can be hard to study e.g. the development of tone in a language of which the writing system does not reflect tonal properties. This means that sometimes we will not be able to see that a sound change has taken place, even though it was initiated and diffused. This topic is studied in (certain) branches of philology.

All in all, Hale (2003) presents the following 'modularity' of scholarly explanation for the existing historical record:

(14)



The figure on p. 363 more or less outlines the same line of thinking: the set of attested human languages is a subset of those which are the possible result of language transmission, which in turn is a subset of all 'computationally possible' human grammars.

Hale (2003) assumes that the 'diachronic filter' is mainly responsible for the statistical effects that are known as 'markedness'. "The failure to recognize the critical role played by diachrony in shaping the set of attested human languages has consistently led phonologists astray, engendering the attribution of epiphenomenal effects of extragrammatical interface systems to UG itself." This is in line with what we have seen for other authors in this class so far; but it is different from most of them in that it assumes that UG still has an explanatory role to play. UG in this view is truly universal; there is hardly any role for variation (parameters, etc.)

4.2 Language acquisition as a source of change

What exactly is a language change? Since we do not see language as an organism independent from its speakers, and since we are only interested in the initiation of a language change, language change can be defined as: the grammar G_1 of one (younger) individual having changed into G_2 of the next individual. Seen this way, all language change takes place during the process of language acquisition.

Furthermore, we have to distinguish between two types of 'change':

1. **Change of lexical representations.** One example: the grandmother of Mark Hale says *davenport* and Mark Hale says *couch* when they talk about the same concept: there has been a change in underlying representations.
   Notice by the way that from this example we can already see that things are a little bit more complicated than Hale (2003) suggests: it is not the case that the form *davenport* has disappeared completely in the new generation. As a matter of fact, the new speakers (like Hale, but presumably also non-linguists of his generation) have a piece of sociolinguistic knowledge available to them: *davenport* is a word used by an older generation. I think this is an important factor in understanding trends in language change stretching over generations. (We will return to this below.)
   'Minor' sound changes should also be understood in this way. If Proto-Polynesian **lango* has turned into Maori *ngaro* rather than *rango*, displaying an irregular metathesis, we assume that the lexical entry for this particular form has simply changed.
2. **Phonological change**, which can be described in terms of a diachronic rule $X \rightarrow Y / Z$, such that every X in the context Z in the original grammar G_1 has turned into a Y in the next generation grammar G_2.

What are the possible sources of language change, i.e. of change of one grammar to another? Hale (2003, p. 349) states that "change can *only* result from the acquirer being exposed to primary linguistic data (PLD) which differs in some way from the PLD which were presented to the source during the source's own acquisition." He then goes on to distinguish between two "primary forces giving rise to such differences":

i. The unique subset of data presented to the acquirer of G_2 may, and in virtually every case will, be different from that presented to the acquirer of G_1 either in *scope* or in *sequence* [...]
ii. The aquirer may mistake the effects of the speaker's production system (A), of ambient effects in the acoustic stream (B), or of his or her own perceptual system (C) as representative of G_1-internal representations or computations. [...]

Blevins (2004) (to be discussed from next week on) is a representative of a modern school of thought which takes ii as central to language change. But Hale (2003)'s claims that such changes will usually only lead to changes in lexical representations. Point i is more important: since children only hear a finite output produced by a grammar which can have an infinite number of outputs, they will all hear a different sample, and furthermore all in a different order.

It is not completely worked out in this paper — or anywhere else in the universe — what are the precise boundaries between the change of individual lexical representations and real sound changes. At some point, the child will no longer have evidence that there are any XZ's in the language, only YZ's. At that point the phonological grammar (rather than anything else will have changed). It seems very likely to me that under Hale (2003)'s approach a typical sound change would start by way of lexical diffusion.

We already mentioned in passing that Hale (2003) is a Saussurean in at least one respect (Saussure, 1916): he assumes that the corpus of data is sociolinguistically unanalysed. The child takes all the data indiscriminately to construct a grammar. Diachronic trends in themselves are invisible. Typically, Hale (2003) gives an abstraction of the acquisition process in which there is exactly one source grammar G_1 and no more. From (Labovian) sociolinguistics we know that this is not necessarily true. If children are able to observe that certain phonetic shapes belong to certain style registers, this may make it possible for them to construct a direction of linguistic change, and set a further step along this path.

4.3 Two opposing views

Hale (2003) contrasts his own views with those of Kiparsky (2003), who represents the classical generative viewpoint, summarised already in previous

classes. The most important point of criticism against such an approach is that it describes language changes directly in terms of changes of grammar — a rule is added to the end of a grammar, it moves upward, rules are reordered, etc. This implicitly presupposes that children will have direct access to their parents' grammars in the course of language acquisition, but there is no evidence that this is the case, nor is it easy to see how this would work. Rather, it seems more plausible that children produce grammars on the basis of what they hear plus their innate predispositions only.

Kiparsky (2003) aims to show how the classical view of generative phonology — insight in synchronic grammar can help to understand diachrony, and language change is driven by the needs of grammars to become simple — can help us solve some of the problems of the Neogrammarian doctrine. Hale (2003) discusses some of these (and the following is again a selection).

One point is (again) the issue of lexical diffusion, where a process spreads from one word to another. One example is *æ* Tensing in Philadelphia English. In this dialect, the lax vowel *æ* tenses to [E] in certain environments:

(15) a. pl[E]n, l[E]gh, m[E]ss
b. pl[æ]net, j[æ]zz, b[æ]g

We can formulate the rule leading to the original sound change as follows:

(16) æ→ [+tense] / { f, s, θ, n, m} $

æ Tensing seems to be subject to a process of lexical diffusion, however. Some words ending in e.g. an /l/ or /d/ also seem to be undergoing it (*m[E]d*, *p[E]l*), and so are some words in which the following consonant is not tautosyllabic (*m[E]nage*).

Kiparsky (2003)'s analysis of this is that the rule in (16) is subject to a process of formal simplification; in particular, gradually the context of this process is becoming simpler and simpler (untill we will have reached a point in which the context is null and all *æ*'s turn into [E].

In the mean time, words which have not yet undergone the process will have to be marked as exceptions. This is done by giving them a lexical specification [-tense], in contrast to other underlying *ae*'s, which are supposed to be unmarked (i.e. they have no value assigned to the feature [tense] at all). In the course of time, marked segments will tend to become simplified, in the sense that they loose their feature specification. Lexical diffusion in Kiparsky (2003)'s model is thus the result of two types of simplification operating in tandem:

- Rules tend to become simpler in that they refer to a smaller number of symbols
- Underlying representations tend to become simpler in that they contain a smaller number of symbols

Hale (2003) does not accept lexical diffusion (something spreading from one word to the next) as a process, or even internal forces within the grammar as an explanatory factor in language change.[3] For him, this is "putting the cart before the horse: the rules are posited by the grammar constructor on the basis of the analysis of the surface forms, and not vice versa", and he argues that the "child cannot change the rules — this would involve first correctly deducing what they were, then 'unlearning' the required system".

In my view, this gives a caricature of Kiparsky (2003)'s actual position. It is of course true that the children do not have access to the rules of their caretakers directly. But if we assume that children have a tendency to always choose the most optimal grammar when confronted with a certain set of data, it is very likely that they will come up with a set of 'rules' and 'underlying representations' not unlike that of the previous generation, only a little bit more optimal. This does not seem to contradict Hale (2003)'s scenario of imperfect learning in any way, except that it requires us to stipulate that some grammars are more optimal than others.

Hale (2003)'s own analysis of these tensing facts is that "whatever acoustic feature of *ae* licensed *ae*-Tensing is a gradient feature, present in various environments to various degrees. It was clearly most salient in the 'core environment' [i.e. the context represented in (16)], present but less prominent in the 'extended environments', and absend or not salient at all in excluded environments." So what happens is that representations turn from æto [E] in individual words. It is statistically more likely that the words with a 'core environment' in which the tensing is most 'salient' will be affected first. (This arguments depends on various factors which are rather vague, such as how exactly to define salience, and how we can understand that some dialects of English have no tensing at all, whereas in others the process has already affected so many words.)

Hale (2003) provides an interesting empirical argument for his account, based on a language change in New Caledonia (p. 355–356). In some dialects only word-final *k* is lost, in others also *c* is lost, in still others *p* has also disappeared and, finally, in the most 'innovating' dialects all final stops {k, c, p, t} have disappeared. The change is one of lexical diffusion; there may be dialects in which most *k*'s and several *c*'s have been deleted, rather than all of these segments.

It is of course absurd to say that deletion of segments can be seen as the result of feature filling and in that sense, New Caledonian dialects do indeed provide an argument against the specifics of Kiparsky (2003)'s proposal. On

[3] Hale (2003) also gives a more technical argument against Kiparsky (2003)'s analysis: "Concerning *ae*-Tensing, is the environment 'before *f, s, θ, m, n, l, d*' more optimal or simpler, in any meaningful sense, than the environment 'before *f, s, θ, m, n*' but not *l* and *d*?" (p. 353) The answer to this question ultimately depends on our view on the representations of the segments involved.

the other hand, the general trend of grammar simplification might be still seen at work in this case:

- The rule is simplified from
  - [+plosive, +peripheral, +velar, -anterior] → ∅ to
  - [+plosive, +peripheral, +velar] → ∅ to
  - [+plosive, +peripheral] → ∅ to
  - [+plosive] → ∅

  in coda position
- The underlying representations turn from e.g. /Vk/ to /V∅/ (=/V/) over time, hence also start to contain a smaller, hence simpler, set of specifications

Especially if we assume that children are able to observe that a language is moving in a certain direction — and Hale (2003) himself observes that the dialects in question are in close contact — these facts do not seem to be in disagreement with a view of these changes as driven by the grammar at all. As a matter of fact, what we have here, is a classical definition of markedness at work.

This is also connected to another point of comparison between the two approaches. Kiparsky (2003) argues that phonological changes are 'structure-dependent', i.e. they tend to create structures already preferred by the language. Long-term drifts such as that of Slavic towards open syllables are understood in this way, and so are diachronic facts such as the following (we mention the fact that is not empirically contested by Hale): "total assimilation of consonant clusters resulting in geminates seems to happen primarily (perhaps only?) in languages that already have geminates (Finnish, Ancient Greek, Latin, Italian). Languages with no pre-existing geminates prefer to simplify clusters by just dropping one of the consonants (English, German, French, Modern Greek)." We do not necessarily have to assume that the children have access to the phonological representations of their parents to follow these guidelines. It could simply be the case that they can only make misanalyses of this type if they have independent evidence for geminates in other parts of the language. (It remains of course a question how a language could ever develop geminates in the first place.)

4.4 Pure phonology (UG)

The main issue distinguishing Kiparsky (2003)'s position from Hale (2003)'s thus is that the former attributes a more important role to internal grammatical factors in language change. His opponent is right in pointing out that it is hard to see how one grammar can change into anohter grammar directly, since there always has to be transmission through E-language, but none of

these problems seem unsolveable. Whether or not grammar does indeed sometimes act as a driving force is a different, empirical, issue altogether.

The fact that Hale (2003) does not acknowledge the driving role of UG in language change — he could not do that, since all language satisfy UG by definition, so they do not need to change in his view — does not mean that it does not have a role to play: it can certainly filter out certain diachronically possible changes. Take the following language change forever:

(17) [-sonorant, -voice] → [+continuant]

This is a well-known type of language change, attested e.g. as part of Grimm's Law in Germanic (***p**ater-**f**ather*). However, this change cannot lead to a violation of a universal of phonology:

(18) Every language has voiceless plosives.

In the case of Grimm's Law, this problem is solved because at the same time other segments (the voiced obstruents) turned into voiceless stops (***d**ens-**t**and*). One could imagine that a child, when confronted with phonetic evidence of a type which would lead her to construct the sound change in (17), would be forced by the UG principle in (18) to analyse other segments as voiceless plosives.

Hale (2003) does not give concrete examples of the constraining power of UG of this type. It is not at all clear that he would accept (18) as part of UG. The fact that it is universal might after all also be attributed to statistical markedness; with the one peculiar (but not impossible) property that the markedness of voiceless obstruents happens to be 0% rather than some higher number. It is therefore not altogether clear what remains to be studied for the phonologist once we have removed the diachronic filters in the model in (14).

5 The diachrony of laryngeal features

5.1 The principles of Evolutionary Phonology revisited

According to Blevins (2004), (almost) all sound patterns in natural language can be explained by reference to language diachrony, i.e. as the result of a combination of the following three regular sound changes:

(19) General typology of sound change in Evolutionary Phonology (S=speaker, L=listener)

a. **Change**: The phonetic signal is *misheard* by the listener due to perceptual similarities of the actual utterance with the perceived utterance.

☞Example: S says [anpa], L hears [ampa]

b. **Chance**: The phonetic signal is accurately perceived by the listener but is intrinsically phonologically ambiguous, and the listener associates a phonological form with the utterance which differs from the phonological form in the speaker's grammar.
☞Example: S says [ʔa̰ʔ] for /aʔ/, L hears [ʔa̰ʔ] and assumes /ʔa/

c. **Choice**: Multiple phonetic signals representing variants of a single phonological form are accurately perceived by the listener, and due to this variation, the listener (a) acquires a prototype or best exemplar of a phonetic category which differs from that of the speaker; and/or (b) associates a phonological form with the set of variants which differs from the phonological form in the speaker's grammar.
☞Example: S says [kakáta, kăkáta, kkáta] for /kakata/, L hears [kkáta, kăkáta, kakáta] and assumes /kkata/

It is thus assumed that the underlying representations are the things that change in language change. There is not a lot of discussion in the chapters we read (or elsewhere in the book) of the repercussions of this on alternations, i.e. on phonological grammar. For instance, the present chapter discusses final devoicing. In many languages, like in Dutch this diachronic process has led to a situation where there are alternations throughout the grammar:

(20)				
	some stage of Dutch a long time ago	/hɑnd/ [hɑnt] ↓	/hɑndən/ [hɑndən] ↓	hand/hands
	Modern Dutch	/hɑnt/	/hɑndə/	

If we take the slashes to literally represent underlying forms, we miss a generalization: there are no words in Modern Dutch which end in a voiceless stop. Nonsense words or recent loanwords also are subject to final devoicing immediately. In terms of Optimality Theory, we still need a constraint:

(21) FINALDEVOICING: Word-final obstruents cannot be voiced.

However, we no longer need to explain this constraint itself in synchronic terms, nor do we need to refer it to UG. The shape of this constraint is arbitrary from a synchronic point of view; it is explained by diachrony.

Notice that this in turn implies that we do not necessarily have to take the slashes in Modern Dutch (20) to refer to real underlying representations; rather, we could take them to designate the outputs of phonology (rather than phonetics). But now we have turned Blevins (2004)' model into one of 'phonologization', and this seems to be too limited for her ambitions.

5.2 Laryngeal features

The chapter we are discussing today wants to give an account of the typology of laryngeal feature behaviour, in particular position-specific neutralisation, and to some extent also assimilation. Different from some other scholars whose work we have studied, Blevins (2004) still assumes universal phonological features for laryngeality:

- [voice] (gives, roughly, voicing)
- [constricted glottis] (gives, roughly, laryngealization)
- [spread glottis] (gives, roughly, aspiration)

The reason Blevins (2004) gives for using features are traditional, but still worth repeating. The most important argument concerns the fact that certain clear phonetic distinctions never seem to play a role in the phonology. One such distinction is the one between preaspiration and postaspiration, which "clearly constitute different acoustic and auditory categories" but still are never contrasted in languages of the world. If a language has both, like Icelandic, they always occur in complementary distribution.

A learner classifies a given sound as belonging to one of these categories (or more than one, or none). An important question now is, of course, where these three phonological categories originate. Since they are universal, they presumably will be part of the innate endowment of all human beings, as are the ways to link these abstract categories to phonetics; note that this means that there still is some substance to (synchronic) phonology.

Phonetic implementation of phonological categories furthermore itself is argued to be "the transparent result of phonetically motivated sound change." (p. 92) Notice however that this type of implementation does not really fit into the general scheme of sound changes in (19) (which, as we have just seen, talks about underlying representations or, at most, the output of phonology).

In any case, even though the phonological and presumably the phonetic categories are assumed to be universal, this is not true for any of the universals involved, according to Blevins (2004). These are the result of "parallel evolution", in her view; because they occur in so many typologically and geographically unrelated languages. She uses a typological argument for this:

> If there is no principled difference between word-final devoicing and word-final voicing, both being the result of chance events, then how are we to explain the high frequency of the first pattern in contrast to the low frequency of the second?

But note that Blevins (2004)'s own proposal does not account for this frequency distribution without further assumptions. There *are* a few languages which seem to have final voicing, and indeed languages of this type should be able to arise (section 4.5.1, p. 108–111), for instance in the following way:

- Suppose we have a language with open syllables and a voiced-voiceless contrast everywhere, except intervocallically (thus with words like [kaga, gaga, paba, baba, tada, dada] but not [gaka, kaka], etc. This seems a natural state of affairs.
- Suppose the language then suffers loss of the final vowel (a natural process, to be understood, e.g. by the principle of Choice)
- The result will be a language with final voicing: [kag, pab, tad] but not [gag, gak, kak], etc.

Somehow it has to be made plausible that this type of situation does not happen very often, or that these languages will be immediately subjected to a further change. But in that case it seems to me that the instability of this particular configuration should still be a topic for study.

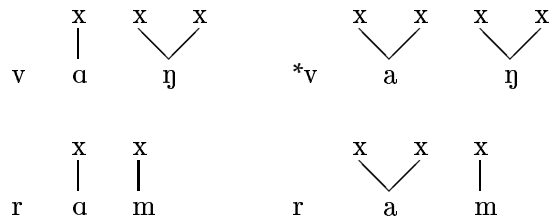
In order to see how Blevins (2004)'s proposal works, let us take a closer look at Frisian. In this language, final devoicing is a relatively recent phenomenon: it did not take place diachronically until well into the 19th century (and for some dialects it does not seem completed yet even now). According to Blevins (2004), the main motivation for this process should have been 'phrase-final lengthening'. Now long consonants are produced and perceived as voiceless. In the first place it is hard to sustain voicing across a relatively long consonant. In the second place, voiceless stops are usually longer than their voiced counterpart so that, inversely, consonant length can give the impression of voicelessness (when listeners misinterpret the phonetic cue). The whole change is then seen as the result of a combination of the principles Choice and Change (4iii).

One prediction of this model is that phonetic devoicing will be most common in velar consonants, and less in coronals and labials (since the former are longer independently). This is supposed to be attested in Frisian, where "final devoicing of *b, d, g* is of recent origin" and where *g* seems to have devoiced before the other stops did.

Another clue that Frisian is supposed to provide for the correctness of this ananlysis is that in Frisian devoicing seems to have started after long vowels, falling diphthongs and liquids, and only later after short vowels and rising diphtongs. The reasoning is that since these vowels already provide some length to the syllable, it might cue the consonant as long as well.

A few things can be said about this. On the positive side, there are a few indications indeed that velar consonants in West-Germanic behave as long. This is clearest for the velar nasal, which can only occur after a short vowel. This can be understood if we assume that ŋ occupies two positions, whereas e.g. m takes only one; and that the rhyme takes three positions:

(22)



The form vaːŋ would be excluded, because it would be too long. On the other hand, this phenomenon of course would get a totally different diachronic explanation (involving the fact that ŋ derives from ŋg.

Furthermore, note that in this case, the *shortness* of the vowel is the cue to the length of the consonant, not the length of the vowel, as it is supposedly the case in Frisian. In general, it seems more likely that long vowels will make the following consonant shorter rather than longer. We should also note that it is not clear at all that indeed final devoicing started at the ends of phrases or sentences in Frisian (although this point can be taken as an interesting empirical challenge as well.)

There are more problems with this particular analysis. For instance, it tacitly assumes that fricatives and stops should have been devoiced at the same time in Frisian (since both are equally subject to lengthening), but this presumably is not the case. Furthermore, it is not clear at all that the early devoicing of ɡ should be attributed to its length only. In the first place, ɡ is a marked segment in Frisian, anyhow — it only occurs in initial onsets in present-day Frisian — and furthermore, it is not unreasonable to assume that devoicing in Frisian was influenced by intensive contact with Dutch, in which devoicing already happened quite some time ago, and which does not have ɡ at all. As is often the case in large-scale typological studies, the particulars become more complicated once we know the language more intimately.

6 Il n'y a que des différences

6.1 Contrast and the organizing system

We have seen in previous classes that diachronic phonology can serve the purpose of explaining many cross-linguistic tendencies, but also that we probably still need to take recourse to synchronic explanation: the changes eventually have to fit into a system.

An important question now is what are the organizing principles of this system. In this class we examine two possibilities for systems of individual sounds:

- Sound systems are organized according to a principle of dispersion: sounds tend to be maximally different.

- Sound systems are organized according to a principle of feature economy: they tend to use a minimal set of features to describe a maximal number of sounds.

Both theories are theories of 'systemic markedness' as opposed to 'paradigmatic markedness' in the words of Itô & Mester (2003). The latter refers to the markedness of e.g. a specific consonant in a specific position (voiced consonants in coda); the former to the markedness of a specific consonant with respect to a whole system.

The two theories make different typological predictions. Under the former we expect systems such as (1); otherwise we expect systems such as (2):

1. {t̪, g, ɱ, ɸ, x̟}
2.

		[labial]	[coronal]	[velar]
[-son]	[-vc]	p	t	k
	[+vc]	b	d	g
[+son]		m	m	ŋ

This way of presentation is somewhat biased: it seems obvious that (2) is much more plausible than (1). However, if we do the same for vowels, things become much less clear:

1. {a, e, i, o, u}
2.

	[labial]	[coronal]
[-low]	u	i
[+low]	o	e

It is not altogether clear that (2) – lacking *a* — is a more natural system here than (1). In general, dispersion theories seem often more succesful in explaining vocalic systems, while feature economy systems are more succesful for consonants. In this class, we discuss one example from both schools.

6.2 Dispersion theory

Itô & Mester (2003) discuss "a specific development in the historical phonology of Japanese concerning sibilants preceding front vowels." The main point they want to analyse is the following:

(23) While [i] triggers palatalisation of (among others) a preceding /s/ to [ʃ], this is not true for [e]. What is more, [e] sometimes seems to trigger depalatalisation of [ʃ].

The latter point is illustrated by loanword adaptation, e.g. examples such as the following:

(24)

Yale	eeru
Los Angeles	rosu anzerusu
shepherd	sepaado

In all cases, the Japanese form has done away with the palatal element before the /e/. If the original language has a palatal+/i/ sequence, the vowel is sometimes lowered to protect the consonant from palatalisation:

(25)

digital	dezitaru
dixie	dekiʃii

(One could of course wonder why a similar solution (vowel raising) is not chosen in the examples in (24), yielding [jiiru] for *Yale*. This is not discussed in the present paper.) In theory, loanword adaptation could still be attributed to something extragrammatical. But this is not true for fully Japanese alternations such as the following:

(26)

	transitive	intransitive	gloss
/tob/	tob-asu	tob-eru	'fly'
/sam/	sam-asu	sam-eru	'cool down'
/moj/	moj-asu	mo-eru	'burn'
/taj/	taj-asu	ta-eru	'extinguish/be extinct'

Itô & Mester (2003) point out that in terms of paradigmatic markedness this fact requires a curious constraint:

(27) DEPAL(mid, front): no palatal(ized) segment before /e/

The constraint is unnatural for a number of reasons. In the first place, palatalisation before /e/ is phonetically grounded. Secondly, a constraint such as this runs against the (phonetically grounded) constraint in (28):

(28) PAL(mid, front): palatalize segments before /e/

It seems unwanted within OT to allow constraints both of type X and of type ¬X, because this gives excessive power (and basically reintroduces the notion of a parameter). Notice, by the way, that we do find a few constraint pairs of this type, e.g. TROCHEE and IAMB.

The effects of (28) is found in other languages and (perhaps) more relevantly, even in dialects of Japanese, for instance in Western and Eastern Kyushu and in Tohoku (Hokurikudo). We find examples such as the following:

(29)

dialect form	standard form	transcription
[ʃenʃee]	[sensee]	'teacher'
[kaʒe]	[kaze]	'wind'

Itô & Mester (2003) show that this type of pronunciation was probably much more widespread in 16th century Japan.

Notice that this line of argumentation presupposes first phonetic grounding and second a universal constraint set: we have evidence for palatalisation and we have evidence for depalatalisation, and since we do not want to have directly conflicting constraints, we go with the former. If we do not accept phonetic grounding, there is no a priori argument for this choice; if we do not accept a universal constraint set, it might be that Modern Japanese simply uses a different constraint from earlier periods.

However, this may be, according to Itô & Mester (2003), the relevant constraint is one of systemic markedness. They give the following 'sibilant systems' for both pre-Edo and Modern Japanese (I leave out the system with /a/, since it is irrelevant):

(30)

pre-Edo

*si	ʃi	su	ʃu
*se	ʃe	so	ʃo

Modern

*si	ʃi	su	ʃu
se	*ʃe	so	ʃo

The idea is that by allowing only non-palatals before e, we enhance the contrast in the 'sibilant system': "overall, [ʃi,se] is a better contrast than [ʃi,ʃe]."

In order to formalize this, we need to take a rather unusual step within Optimality Theory. We no longer compare one candidate output of a given input with all other candidates, but we rather compare two 'systems' such as those in (30). This means we also need to have an 'underlying' system, which in the case at hand is supposed to be a full system consisting of {si, se, ʃi, ʃe} (very little arguments are given for this assumption, and it is not clear how it relates to OT's principle of Richness of the Base.) It is also not very clear how the evaluation here proceeds with respect to

We now need two sets of constraints on inventories: markedness and faithfulness constraints:

(31)
a. Systemic markedness: SPACE$_{x>1/n}$; Potential minimal pairs differing in property x must differ in x by at least $1/n$th of the available space
b. Systemic faithfulness: NOMERGE; underlying contrasts may not be merged.

SPACE$_{x>1/n}$ will ensure that in the output we will only find maximally contrasting segments; NOMERGE makes sure that underlying contrasts are not deleted. (It is not clear to me that the latter constraint plays a significant role in the analysis; all its effects could be assumed by regular faithfulness constraints.)

Apart from the fact that this type of analysis brings a lot of extra computational power to the theory (we are now comparing whole systems rather than individual forms), this analysis also raises many questions. We mention two:

- First, there is no real answer to the question why the complementary distribution between [s] and [ʃ] involves only the front vowels. According to this theory, a distinction between {ʃu, so} would also be preferable. Technically, there is an answer to this: there is a highranking CONTRAST constraint for front vowels, but not for back vowels. But this cannot count as a very insightful answer of course to the question what is the relation between palatalisation and preservation of contrast. In principle, this theory could deal with the ontogenesis of palatal contrast before back vowels (or the pair {ʃu, sa}) just as well.
- Secondly, the notion of contrast in this case is built on the 'CV mora': the unit of a consonant followed by a vowel. It is not clear why we would want to compare this unit rather than anything else, or more precisely, whether this will be a universally available unit of comparison and, if so, whether we can also compare single segments, full syllables, feet, etc.

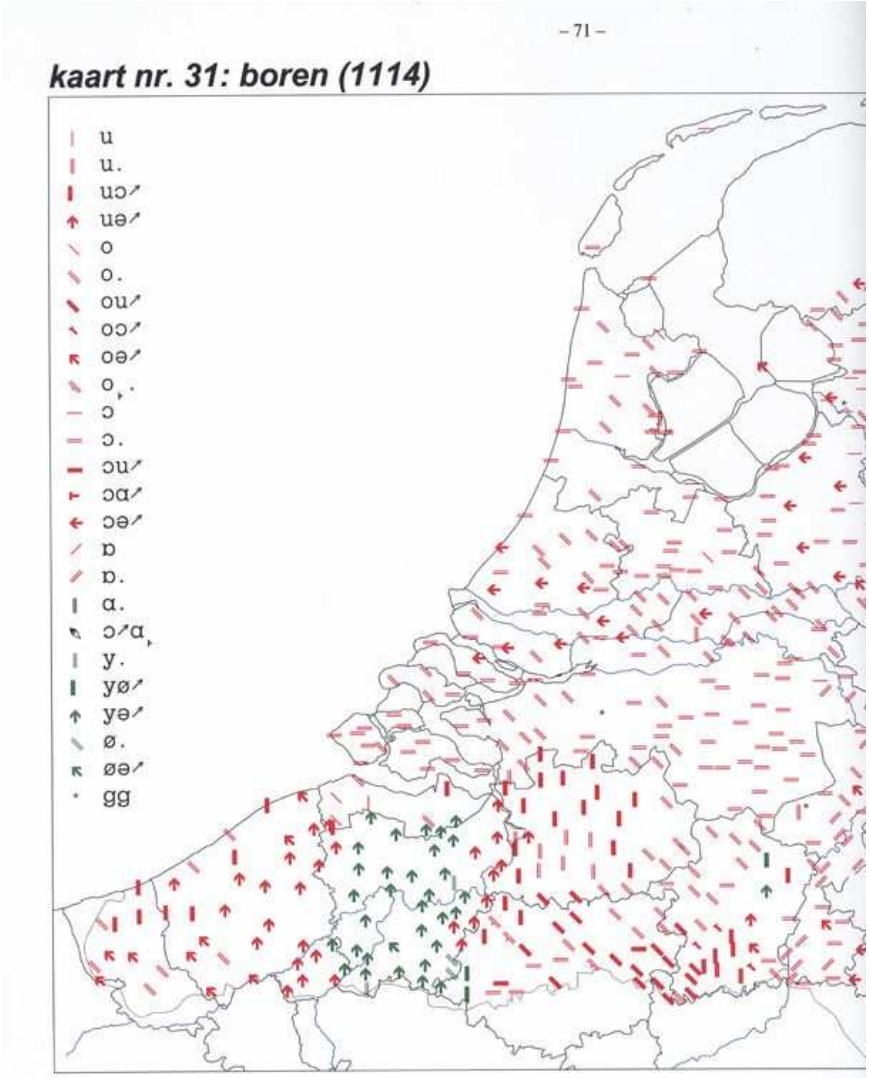
In section 5, Itô & Mester discuss an older proposal of their own making (Itô & Mester, 1995). This proposal is representational, and it depends on the assumption of a different segmental structure for /e/ and /i/. I think they dismiss this proposal to readily; the only argument is that "palatalization before mid front vowels is not at all uncommon" and "in light of this, it seems ill-advised to depend on a theory where an assimilated sequence such as [Se] violates a putative universal representational constraint". But in the first place, palatalisation before high vowels and not before mid vowels is not uncommon either, and more importantly, the OCP constraint seems independently necessary (e.g. for Dutch, where *[ji] sequences are ruled out independently — notice by the way that [je] is very marginally possible in a few names; contrast does not seem to play a role here).

6.3 Feature Economy

Feature Economy (Clements, 2003) is the hypothesis that "if a feature is used once in a system, it will tend to be used again". Clements has shown this hypothesis to be a valid generalisation for cross-linguistic databases; we will examine its value for Dutch dialectology. For this we use the Goeman-Taeldeman-Van Reenen Database, containing a large survey of phonological and morphological variation in the Netherlands. The results seem to confirm the Feature Economy Hypothesis, and by extension, a feature-based view of phonological structure.

Structure of FAND

- Based on the Goeman-Taeldeman-Van Reenen Project (GTRP):
- A rather traditional, historically oriented, atlas
- Volumes I (on short vowels) and II (on long vowels) appeared in 2000 and 2002
- Volume III (on consonants) is due in 2005
- FAND is strongly oriented towards history, we do not necessarily know what is e.g. the consonantal system of a modern dialect
- It is hard (if not impossible) to study inventories on the basis of maps of individual segments



Boren (1114)

Verklaring van de symbolen in de legenda:

u	[u]
u.	[u.], [u:], [u˕:], [0.], [0:,]
uɔ↗	[uɔ↗], [u.ɔ↗], [u˕.ɔ↗], [u:ɔ↗], [u˕:ɔ↗]
uə↗	[u˕ə↗], [u˕ə↗], [u˕ə˕↗], [u.ə↗], [u˕.ə˕↗], [u˕˕.ə↗], [u:ə↗], [u˕:ə↗], [u˕:ə↗]
o	[o], [o˕], [o˕], [u˕]
o.	[o.], [o˕.], [o˕.], [o˕.], [o:], [o˕:], [o˕:], [o˕:], [o˕:], [o˕˕:], [u.], [u:]
ou↗	[o˕u↗], [o:u↗], [o˕:u↗]
oɔ↗	[o:ɔ↗], [u:ɔ↗]
ɔu↗	[ɔu↗], [ɔ˕u]
ɔɑ↗	[ɔ˕.ɑ↗], [ɔ˕.a↗]
ɔə↗	[ɔə↗], [ɔ.ə↗], [ɔ˕.ə↗], [ɔ:ə↗]
ɒ	[ɒ˕]
ɒ.	[ɒ:], [ɒ˕:], [ɔ˕:], [ɔ˕˕:]
ɑ.	[ɑ:]
ɔ↗ɑ˕	[ɔ↗a˕.]
y.	[y:]
yø↗	[y:ø↗]
yə↗	[yə↗], [y˕ə↗], [y.ə↗], [y˕.ə↗],

Using a database

Principles of GTR

- a database of approximately 1,000,000 words
- from 612 traditional dialects spoken in the Netherlands and Flanders.
- data are collected between 1980-1995
- Database can be accessed at `http://www.meertens.knaw.nl/projecten/mand/`
- Project had as its ideal (in 1978) to answer "research questions for dialectologists in the next few decades" (Goeman & Taeldeman, 1996)
- These were supposed to be questions from 'formal linguistics and language typology'

A few example questions from 1978:

1. To what extent are sound systems organized following principles of harmony/economy? (N.B. this is an old research question dating back to the Prague circle)
2. How frequent do we find certain distinctions (e.g. in the case of vowels: [front]-[back] and in the case of front vowels: [±round])
3. Are there any recurrent patterns of allophonisation?
4. Do we find a correlation between the functional load of a phoneme and its stability?

Structure of the database

The following is an example of an entry in GTRP:

(32)

1.	2.	3.	4	5
E192p	130	6n dr7a2.t	'n draad	a thread
E192p	131	dr7a2.<d6n	draden	threads
E192p	132	6n dr7a.>ts2i	'n draadje	a thread-DIM
E192p	133	6n dr7o5po2l4	'n druppel	a drop
E192p	134	6n do7y_f	'n duif	a pigeon
E192p	135	do7y_v6n	duiven	pigeons
E192p	136	6n do7y_fi	'n duifje	a pigeon-DIM

The three leftmost columns are actually in the database. In column 1 we find the so-called Kloeke-code, a geographical code which represents a city, town or village in the Netherlands (Kloeke-code E192 represents Utrecht), and in column 2 we find the question number (these thus are questions 130-136). In column 3 there is the actual answer to the question in K-IPA, a pre-SAMPA rendering of IPA in ASCII letters.

Letters correspond to IPA symbols, numbers and other symbols are diacritics (except '6'=ə)

The fourth column gives the actual 'question', i.e. the Dutch word which had to be translated into the dialect, and the fifth column gives an English gloss.

Feature economy

Principles

Clements (2003) revives 'an old research question dating back to the Prague circle' on the basis of macrotypology (i.e. the UPSID database, cf. van de Weijer & Hinskens (2004)):

(33) a. "A sound S will have a higher than expected frequency in languages that have another sound T bearing one of its features, and vice versa"'

b. If a language L_1 had /p/ and L_2 does not, the chance that L_1 also has /b/ is larger than the chance that L_2 has the same sound

Compare the following two hypothetical consonant inventories:

1.

		[labial]	[coronal]	[velar]
[-son]	[-vc]	p	t	k
	[+vc]	b	d	g
[+son]		m	m	ŋ

2. {t̪, g, ɱ, ɸ, ɬ}

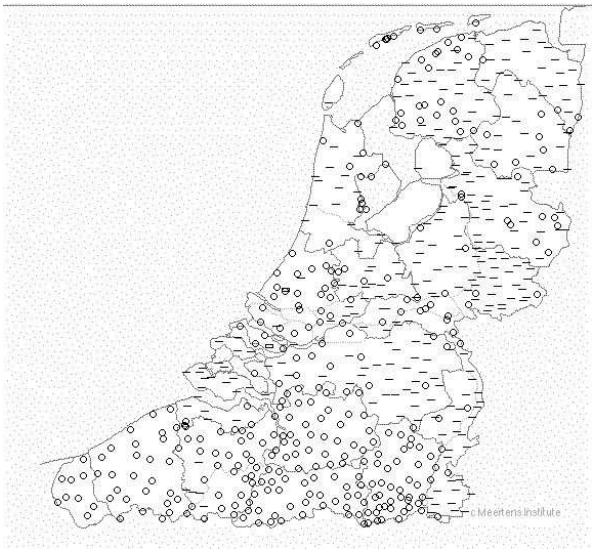
- If phonetics (perception) would be an organizing principle of sound inventories, one might expect (2) to be predominant
- Assuming Feature Economy, we expect (1) to be predominant
- This issue cannot be studied on the basis of one language system alone, it is a typological question by definition

Q: To what extent are the consonant inventories of Dutch dialects organized according to principles of feature economy?

Method

- We automatically derived phonetic segment inventories for each of the 612 dialects in the GTRP database
- Average size of inventories: 208 distinct segments
- Standard Dutch has approximately 50 'phonemes'
- Therefore, the transcriptions are presumably phonetic

Examining the inventories, we discover there is a difference between the Netherlands and Flanders (dots represent dialects with fewer transcriptions than average, horizontal lines dialects with more transcriptions than average):



Probably this is due to a difference in methodology between transcribers. Some of them draw 70 distinctions on average (i.e. a phonological distinction, others over 350.)

The structure of a segment

- There is no reason to expect that Dutch dialects will differ in the number of e.g. /b/'s
- We will therefore have to concentrate on more fine-grained distinctions if we want to study Feature Economy effects

A simple way of doing this, given the structure of the database, is by looking at diacritics.

- One diacritic is { for dentality (i.e. K-IPA t{ corresponds to IPA t̪)
- This only occurs on coronals {t, d, s, n, l, z}

Given Feature Economy, we hypothesize the following:

(34) The chance of having d̪ is bigger in dialects with t̪, than it is in dialects without t̪?

Chance of finding a pair of dentals

Using a set of specifically designed (Python) scripts to search the database of phonetic segment inventories derived from GTRP. Given this, we found the following:

- The chance that we find d̪ is (14+11) / 612 = 0.04.
- The chance that we find t̪ is (11+41) / 612 = 0.09.
- Therefore, the chance that we find d̪+t̪ should be 0.04 x 0.09

- We thus should find 0.04 x 0.09 x 612 = 2 instances of the combination. But we find 14.

Frequency of dentals:

	t̪	¬t̪
d̪	14	11
¬d̪	41	546

χ^2 = 70.43
p <0.001

Obviously, most dialects do not have any dental segment at all. But disregarding these (which are not informative as to our present question), we find that indeed the combination of d̪ and t̪ is much more frequent than we expect on the basis of the relative frequencies of each of these sounds individually (but notice that this does not mean that there are more dialects with the combination than there are dialects which have only one of the two dentals).

Outlook

Because there are such huge differences among transcribers, we are not sure that the effect we found is not due to them.

- A following step should be based on the following definition of Clements:
  - "Feature economy can be quantified in terms of a measure called the economy index.
    Given a system using F features to characterize S sounds, its economy index E is given by the expression E = S/F "
- Are big inventories more economical than big small?
- Initial calculations suggest that they are: for large (>203) segment inventories, the economy index is 5.1 on average; for small ones (<204) this is 2.6

7 Franconian tones: phonology vs. phonetics

7.1 Introduction

Most Limburg dialects of Dutch, as well as neighbouring dialects of German — the dialect area is usually referred to as 'Franconian' — display a lexical contrast between two types of tone, a falling tone (traditionally called *stoottoon*) and a level high tone (*sleeptoon*). The following examples are from Maasbracht:

(35)

falling tone	*level high tone*
míǹ 'minus'	míń 'vile'
dǽǹ 'fir'	dǽń 'then'
klám̀ 'trap'	kláḿ 'hardly'
bî 'bee'	bíí 'with'
zî 'side'	zíí 'she'
pîp 'to squeak'	píip 'pipe'

The tonal contours can also be used to express (inflectional) affixation, e.g. for singular/plural pairs:

(36)

falling tone	*level high tone*
bɛ́ìn 'legs'	bɛ́ín 'leg'

The following two pictures represent the F0 values for these two tones (for a speaker from the Roermond dialect, very close to Maasbracht[4]):

(37) falling tone level high tone

The 'falling' tone is characterised by a clear downward movement; the 'level high' tone also moves slightly downward, but then goes up again towards the end. There are several ways to translate this into the phonology, but many analysts have converged on the following (see Gussenhoven, 2004, for an authoritative overview):

(38)

falling tone	level high tone
H L	H (L) H
| |	| |
μ μ	μ μ
| |	| |
m ɪ n	m ɪ n

Limburg tones have several properties which can be readily explained in a phonetically-based framework of phonology. Take for instance the following observations:

1. Tonal contours are only expressed on a long vowel or a V+sonorant ('VR') rhyme.

[4]The data were analysed with the *Praat* programme; `http://www.praat.org/`. The data are almost identical to those presented in Gussenhoven (2000).

2. Every word has at most tone contour, which is expressed on the syllable carrying main stress.

Zhang (2004) gives a few typological generalisations which show that Limburg is not special in these respects:

(39) (= Zhang (2004)'s (11)) Contour tones occur more freely:
   a. ...
   b. on CVV and CVR in 66 languages (Kiowa, Nama, Fuzhou Chinese)
   c. on stressed syllables in 21 languages (Xhosa, Jemez, Lango)
   d. ...

Generalisations such as these are supposed to follow from phonetics. For instance, the reason why we can hear contours on long vowels but not on short vowels is that "given the same pitch excursion, the longer the duration of the vowel, the more 'contour-like' the tone is perceived by the listener. [Greenberg & Zee (1979)] also show that listeners cannot perceive pitch changes reliably when the duration is below 90 ms." Because stressed syllables are longer, their attraction to tones may also be explained in this way (as well as by the fact that stress also corresponds to relative loudness).

As for the reason why coda sonorants, but not obstruents, can support tone, Zhang (2004) observes that "the main perceptual correlate of tone is the fundamental frequency (f_0). All harmonics serve as a cue to f_0, since they occur at integral multiples of f_0. However [...] the spectral region containing the second, third and fourth harmonics is especially important in the perception of fundamental frequencies in the range of speech sounds. Since sonorants possess richer harmonic structures than obstruents, including the crucial second to fourth harmonics, sonorants are better tone bearers than obstruents."

Seen from this angle, the motivation behind some of the choices made in Limburg dialects may therefore be phonetic. Still, the way it works out here (as well as in other tone languages) has to be phonological. The reason is that the ban on tones outside of the main stress and syllables with a long vowel or sonorant rhyme is absolute. This applies also to loanwords, which will always have one of the two tones on the stressed syllable, given appropriate conditions. This means that analyses in which there is no role for a (phonological) system are difficult to maintain. Zhang (2004) formalizes the analysis in terms of a family of OT constraints:

(40) CONTOUR(x_i)-C$_{contour}$(y_i)
no contour tone x_i is allowed on syllable with the C$_{contour}$ value of syllable y_i or smaller

The constraints in this family are inherently ranked. "These rankings reflect the speaker's knowledge that a structure that is phonetically more demanding should be banned before a structure that is less so; and that a syllable should be able to host a tone with a lower complexity before it can host a tone with higher complexity." This places Zhang (2004) in the school of thought which was exemplified in our course by Hayes.

Of greater interest is the fact that the Limburg dialects also seem to run counter to at least some of the proposed universals. In some cases the problem may be only apparent. Take, for instance, the following (Zhang (2004)'s example (13a)):

(41) If a language has contour tones, then it also has a level tone.

Under the analysis presented above, Limburg dialects only have contour tones HL and HLH. However, there is actually even more than one way out of this problem. For instance, we may analayse HLH as actually being a level high tone, although this would be a rather abstract analysis in many cases — one which we might want to avoid if we follow the lines of Zhang (2004). A more viable option might be to consider that other factors are interfering here. For instance, the fact that the first mora of the stressed syllable might be attributed to independent functionally motivated constraints, such as the fact that stress tends to be interpreted as high toned in any case. This would leave the choice of lexical specification restricted to the second mora, which could be only high or low. (41) (which is by the way called a 'strong implicational tendency', not a universal by Zhang (2004)) would thus be violated for a reason in Limburg.

Yet, the question always is how many ways out we allow ourselves. Notice that Zhang (2004)'s arguments have all been taken from the realm of articulatory ease, but the argument about the high tone on stress has to come from some other area (maybe processing). It is the question how much knowledge about how many areas the child may bring to bear on language acquisition (and to what extent this gives us a parsimonious theory).

In the following sections, we will go into more detail into two more specific problems, having to do with the interaction with consonants and vowels respectively. In both cases, the generalisations we can make seem to run counter to what we would expect on the basis of phonetic facts only. We will see how these seem to rule out certain accounts.

7.2 Interaction with consonants

Lexical tone interacts with laryngeal features on consonants in a number of ways; they all have in common that a falling tone prefers not to be followed by a voiceless obstruent. Here is an example; if a stressed vowel is followed by a sonorant plus a voiced obstruent, we have a potential contrast:

(42) Roermond contrast in voiced environment

Falling tone	Level tone
bændə 'gang'	dʊndər 'thunder'
hærdər 'shepherd'	mærɣəl 'marl'

If the last consonant of this sequence is voiceless, however, only the level tone is allowed:

(43) Roermond gap in voiceless environment

Falling tone	Level tone
(missing)	pɪmpəl 'booze'
(missing)	hærsəs 'brains'

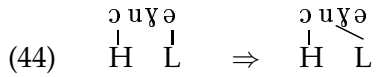
This gap suggests the following generalisation:

☞ *Consonant-tone interaction in Roermond*
A voiceless consonant disprefers a low tone on a consonant at its left.

Now interactions between laryngeal features and tone are well-known from other languages. In Suma (Bradshaw, 1999), for instance, imperfective verbs start with a High tone (*éé* 'leave behind', *kírí* 'look for'), except when they begin with a voiced obstruent, in which case the first tone is rising (i.e. Low-High: *bŭsi* 'be bland'). Diachronically, the effect is known as well from other tone languages, where voicing contrasts on onset consonants turn into tone contrasts on following vowels. A second type of interaction is that the spreading of a high tone is blocked by an intervening voiced obstruent and the spreading of a low tone is blocked by an intervening voiceless obstruent (Hyman & Schuh, 1974). Botswana Kalang'a (Bradshaw, 1999) is a language where the former process applies. In this language high tones normally spread to toneless (i.e. phonetically low-toned) syllables in a following word (/tʃipó + tʃipó/ ← [tʃipótʃípó] 'your gift'), except if a voiced obstruent intervenes (/zwipó + zwipó/ ← [zwipózwipó] 'your (pl.) gifts').

In all known cases, it is the following vowel that is affected by the consonant, and any theory has to account for this. But all of these theories are then facing a serious problem with the Limburg facts. This is specifically true for the phonetic optimisation approach: if it is optimising to link an laryngeal feature in the onset to a following vowel, the Limburg facts seem 'crazy' and in need of a separate constraint.

A diachronic (evolutionary) approach for part of the Limburg facts is possible, however. Most scholars assume that the rise of tone is connected in one way or another to a process of schwa apocope by which e.g. *ouge* 'eye' turned into *oug*. Although there is no consensus on what this relation is exactly, this gives us the opportunity to give a typologically more plausible analysis to the Limburg phenomenon: we may assume that the Low tone originated on the schwa following the voiced obstruent:

(44)


The origin of this phenomenon could thus have been a phonetically natural one (whatever causes the misperception of voicing on obstruents as low tone on the following vowel), but it needs to be noted that the system has been subsequently heavily phonologized in ways that stretch far beyond these phonetic origins.

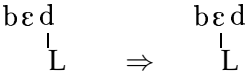
One example of this is from the Moresnet dialect (Jongen, 1972). Different from most other dialects, short vowels followed by an obstruent also get tone. This tone is falling if the vowel is followed by voiced obstruent, but level if followed by a voiceless obstruent. Thus, in a word like *bedde* 'bed-PL' the first vowel has a falling tone, whereas in words like *teppich* 'carpet' and *kes* 'casket' it is level.

Interestingly, words that are devoiced by Final Devoicing have a falling tone, according to Jongen. Thus, *bet* 'bed-SING' has a falling tone. This seems to suggest that the tone of the short vowel is determined at the underlying level. That is, since *bet* has a voiced consonant underlyingly, the vowel preceding it has a falling tone, even at the surface level.

It is hard to see how a synchronic or diachronic account which is based directly on phonetics is going to make sense of this, because the effect is clearly very opaque: something happens only to devoiced segments, not to segments which are underlyingly voiceless, or voiced on the surface. But 'devoiced', referring to the phonological derivation, is obviously not a phonetic category.

The only approach which has some hope of shedding light on this, is the one by Boersma: the tonal change on the vowel acts as a recoverability marker for the underlying voicing. The hearer gets a cue from the 'wrong' tone about the underlying status of the consonant. Yet, if tone on a vowel is usually a cue for the voicing of the *preceding* segment, it is still not clear how this would work exactly.

Notice that a more abstract approach is available; there have been various places in the literature where it is suggested that from a phonological point of view, Low tone and [+voice] are the same animal, which we will represent here as L. Under such an approach, the tonal shift is a type of faithfulness:

(45)


Phonetically low tone and voicing are not obviously the same thing. We thus need this level of abstractness to maintain this analysis.

7.3 Interaction with vowels

The interaction between tone and vocalic height is very puzzling from a typological point of view. In the first place, the fact that there is any phonological interaction between these two dimensions at all sets Limburg apart from all other known tone languages. There is some phonetic interaction, but this is never phonologized — in itself an observation that is worth considering.

In Limburg, the connections between vowel quality and tone are many, but the basic observations are that we find diphthongization under a falling tone — e.g. long /eː/ or /iː/ turning into [ɛi] — and monophthongization under a level high tone — e.g. /ɛi/ turning into [eː]. Other cases involve lowering of diphthongs and mid vowels under a falling tone — e.g. /ɛi/ changing to [ɑi], and /eː/ to [ɛː] — and raising of mid vowels under a level high tone — changing e.g. /ɛː/ to [eː]. A further problem with these developments is that they go in the wrong direction: if there is a phonetic effect, it is that high tones want to go with high vowels; but here the level high tones seem to prefer low vowels and vice versa.

By way of an example, the following gives an overview of vocalic changes which happened to the Maastricht dialect (de Vaan, 2004).

(46)

West-Germanic vowel	Falling tone	Level tone
iːC, yːC, uːC	ɛi, øy, ou	iː, yː, uː
iːr, yːr, uːr	eː, øː, oː	iː, yː, uː

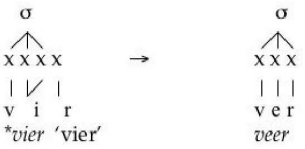
In order to explain effects such as these (we will concentrate on the lowering), which run in the reverse direction of what we want Gussenhoven & Driessen (2004) propose the following phonetic explanation. Falling tones have to be perceptually short. Now normally high vowels are actually shorter than mid vowels (and it is more difficult to keep to the length of high vowels from an articulatory point), so this would again normally run in the wrong direction, but for this reason Gussenhoven & Driessen (2004) assign a special role to the listener. Since she knows that high vowels are usually shorter, she will automatically always add something to their length; and in an experiment they proved that if we present an [iː] and an [eː] of exactly the same length, say 220 ms, to a set of speakers, they will perceive the former as longer than the latter. (If we present an elephant and a mouse of the same absolute height, the elephant will be perceived as very small and/or the mouse as very big.) This is why a vowel which is usually short can be used as a cue for length.

Various things can be said about this approach; for instance, it runs the danger of paving the way for very unrestricted analyses. If we find an effect A, we can attribute this to articulation; but if we find ¬A, we can attribute this to the fact that the listener subtracts ease of articulation. We can thus explain one effect and the contrary effect at the same time, which is not a very desirable state of affairs.

Again, a more abstract approach presents itself. Let us assume that the basic idea of Gussenhoven & Driessen (2004) is correct, viz. that falling tones in Maastricht want to be represented as short. This seems particularly attractive in this dialect, where length is the most important cue for the tonal distinction (Gussenhoven & Aarts, 1999). But let us also assume that this is a *phonological* generalisation.

Now we can observe that Maastricht contrasts short [i] to long [iː] (there are minimal pairs), but it has only one tense mid vowel [e]. This means, then, that [iː] is really, phonologically, long, whereas [eː] is phonologically only tense. Seen from that perspective, the change from [iː] to [e] is thus a phonological shortening:

(47)



Again, it is hard to see how this could be captured directly in phonetic terms (since [e] is not necessarily shorter than [iː]).

We thus see that the Limburg tones interact with consonants and with vowels in ways that are very hard to capture in terms of phonetics directly. If phonetic generalisations have played a role at all, it is most likely that they have done so in a diachronic sense, more or less in the way this is depicted in Evolutionary Phonology. But after these generalisations have entered the grammatical phonological system, they have been subjected to rules of this cognitive game. That seems to be the most important lesson to take home after this course.

Bibliography

Baudouin de Courtenay, Jan (1895). *Versuch einer Theorie phonetischer Alternationen: Ein Capitel aus der Psychophonetik*. Strassburg/Crakow: Trübner.

Bermúdez-Otero, Ricardo (2005). 'Diachronic Phonology'. In: de Lacy, Paul (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.

——— (forthcoming). *Stratal Optimality Theory*. Oxford: Oxford University Press.

Blevins, Juliette (2004). *Evolutionary phonology. The emergence of sound patterns*. Cambridge: Cambridge University Press.

Boersma, Paul (1998). *Functional Phonology*. Ph.D. thesis, University of Amsterdam.

——— (1999). 'On the need for a separate perception grammar'. URL `http://www.fon.hum.uva.nl/paul/papers/percgram.pdf`. Manuscript, University of Amsterdam.

Bradshaw, Mary (1999). *A Crosslinguistic Study of Consonant-Tone Interaction*. Ph.D. thesis, Ohio State University.

Bybee, Joan (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Carstairs-McCarthy, Andrew (1999). *The origins of complex language: an inquiry into the evolutionary beginnings of sentences, syllables and truth*. Oxford: Oxford University Press.

Chomsky, Noam (2004). 'Beyond explanatory adequacy'. In: Belletti, Adriana (ed.), *Structures and Beyond: The Cartography of Syntactic Structures, Vol. 3*. Oxford: Oxford University Press.

Clements, G.N. (2003). 'Feature Economy in Sound Systems'. *Phonology*, 20, 3: 287–333.

Clements, Nick & Samuel Jay Keyser (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.

Féry, Caroline (2003). 'Liaison and syllable structure in French'. URL `http://www.sfb632.uni-potsdam.de/homes/fery/pdf/liaison%20actual.pdf`. Manuscript, Potsdam University.

Goeman, Ton & Johan Taeldeman (1996). 'Fonologie en morfologie van de Nederlandse dialecten. Een nieuwe materiaalverzameling en twee nieuwe atlasprojecten'. *Taal en Tongval*, 48, 1: 38–59.

Goldinger, S.D. (2000). 'The role of perceptual episodes in lexical processing'. In: Cutler, Anne, J.M. McQueen, & R. Zondervan (eds.), *Proceedings of SWAP (Spoken Word Access Processes)*, pp. 155–159. Nijmegen: MPI.

Greenberg, Steven & Eric Zee (1979). 'On the perception of contour tones'. In: *UCLA Working Papers in Phonetics*, vol. 45, pp. 150–164.

Gussenhoven, Carlos (2000). 'The lexical tone contrast of Roermond Dutch in Optimality Theory'. In: Horne, Merle (ed.), *Intonation: Theory and Experiment*, pp. 129–167. Dordrecht: Kluwer. URL `http://www.let.ru.nl/gep/carlos/roermond.pdf`.

——— (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos & Flor Aarts (1999). 'The Dialect of Maastricht'. *Journal of the International Phonetic Association*, pp. 155–160.

Gussenhoven, Carlos & Wilske Driessen (2004). 'Explaining two correlations between vowel quality and tone: The duration connection'. In: *Prosody 2004*. Nara, Japan: Prosody.

Hale, Mark (2003). 'Neogrammarian Sound Change'. In: Joseph, Brian & Richard Janda (eds.), *The Handbook of Historical Linguistics*, pp. 344–368. Oxford: Blackwell.

Hale, Mark & Charles Reiss (2000). 'Phonology as cognition'. In: Burton-Roberts, N., Philip Carr, & Gerard Docherty (eds.), *Phonological Knowledge: Conceptual and Empirical Foundations*, pp. 161–184. Oxford: Oxford University Press.

Hauser, M.D., N. Chomsky, & W.T. Fitch (2002). 'The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?' *Science*, 298: 1569–1579.

Hayes, Bruce (1999). 'Phonetically-driven phonology: the role of Optimality Theory and grounding'. In: Darnell, Michael, Edith Moravscik, Michael Noonan, Frederick Newmeyer, & Kathleen Wheatly (eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, pp. 243–285. Amsterdam: John Benjamins. URL `http://www.linguistics.ucla.edu/people/hayes/phonet.htm`.

Hyman, Larry (1976). 'Phonologization'. In: Devine, A.M. & Laurence D. Stephens (eds.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*, pp. 407–418. Saratoga: Anma Libri.

Hyman, Larry & R. Schuh (1974). 'Universals of tone rules: Evidence from West Africa.' *Linguistic Inquiry*, 5: 81–115.

Itô, Junko & Armin Mester (1995). 'The core-periphery structure of the lexicon and constraints on reranking'. *University of Massachusetts Occasional Papers*, 18: 181—210.

——— (2003). 'Systemic Markedness and Faithfulness'. Manuscript, UCSC.

Jakobson, Roman (1960). 'Linguistics and Poetics: Closing Statement'. In: Sebeok, Thomas (ed.), *Style in Language*. Cambridge, MA: Massachusetts Institute of Technology.

Jongen, René (1972). *Phonologie der Moresneter Mundart; Eine Beschreibung der Phonologie der Moresneter Mundart. Eine Beschreibung der segmentalen und prosodischen Wortformdiakrise*. Assen: Van Gorcum.

Kayne, Richard S. (1994). *The Antisymmetry of Syntax*. Cambridge: The MIT Press. Linguistic Inquiry Monograph Twenty-Five.

Kiparsky, Paul (2003). 'The Phonological Basis of Sound Change'. In: Joseph, Brian & Richard Janda (eds.), *The Handbook of Historical Linguistics*, pp. 313–343. Oxford: Blackwell.

Labov, William (1993). *Principles of Language Change. I: Internal Factors*. Blackwell: Oxford.

——— (2001). *Principles of Language Change. II: Social Factors*. Blackwell: Oxford.

Miller, Geoffrey (2000). *The Mating Mind : How Sexual Choice Shaped the Evolution of Human Nature*. Doubleday.

Passy, Paul (1891). *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin Didot.

Pierrehumbert, Janet (2002). 'Word-specific phonetics'. In: Gussenhoven, C. & N. Warner (eds.), *Laboratory Phonology VII*. Berlin: Mouton de Gruyter. URL `http://babel.ling.northwestern.edu/~jbp/LP7box/LP7.html`.

Prince, Alan & Paul Smolensky (1993). 'Optimality Theory: Constraint Interaction in Generative Grammar'. Manuscript, Rutgers University and University of Colorado at Boulder. Available at ROA.

Saussure, Ferdinand (1916). *Cours de Linguistique Générale*. Paris: Editions Payot.

Scheer, Tobias (2004). *A Lateral Theory of Phonology. Volume I: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.

de Vaan, Michiel (2004). 'Wgm. *ī en *ū vóór *r* in Zuid-Limburg'. *Taal en Tongval*, 54, 2: 171–182.

van de Weijer, J.M. & F. Hinskens (2004). 'Markedness and Segmental Complexity: A Cross-Linguistic Study'. In: *GLOW Phonology Workshop, Thessaloniki*.

Zhang, Jie (2004). 'The role of contrast-specific and language-specific phonetics in contour tone distribution'. In: Hayes, Bruce, Robert Kirchner, & Donca Steriade (eds.), *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.