Typology and acquisition in functional and arbitrary phonology

Paul Boersma, University of Amsterdam, http://fonsg3.hum.uva.nl/paul/
draft, December 15, 1998

In this paper, I will show how an empiricist theory about the substance of phonology accounts in a coherent and typologically adequate way for the acquisition of speech.

1 The relation between innateness and typology

Linguistic theorizing moves between the opposing standpoints of nativism and empiricism.

1.1 Innateness and function

The nativist standpoint of generative linguistics (Fig. 1) holds that the human language capability is laid down in an **innate** Universal Grammar (UG), which contains a number of principles heeded by all languages. Many of these principles must have originated in an evolutionary selective pressure towards efficient communication. For instance, Hurford (1989) shows that if speakers expressed each concept (the *signifié* of Saussure 1916) with the same sound-image (Saussure's *signifiant*) that they have seen listeners interpret as referring to that concept, and if listeners interpreted each word in the way that they have heard others speak with a certain meaning, people would soon fail to understand one another. A more successful strategy would be to for the learner to speak the same words that she hears others speak with a certain apparent meaning, and to base her reactions to speech of others on the meaning apparently intended by the speaker (and not on the meaning apparently distilled by other listeners, which is the third strategy Hurford investigates). Hurford shows that *if* the learning strategy is hereditary and exhibits genetic variation, and successful communicators raise more offspring than others, the successful strategy of using the same words in speaking as well as listening will emerge as an innate property throughout the community after a number of generations. Hurford proposes that this is indeed what happened to humanity, and that this strategy has become part of our innate Language Acquisition Device.

On the other hand, the empiricist standpoint of functional linguistics (Fig. 2) denies the innateness of the details of language behaviour, proposing instead that they are **learnable**. The principles of efficient communication must either still be felt by the speaker/listener, or have been expressed into the shape of the language by previous generations. Substantively, the grammar must consist solely of language-specific, non-innate constraints, which are extracted directly from the language data by general mechanisms of generalization in human perception and by general properties of motor behaviour and its monitoring (Figs. 1 and 2 do not show any arbitrary non-universal

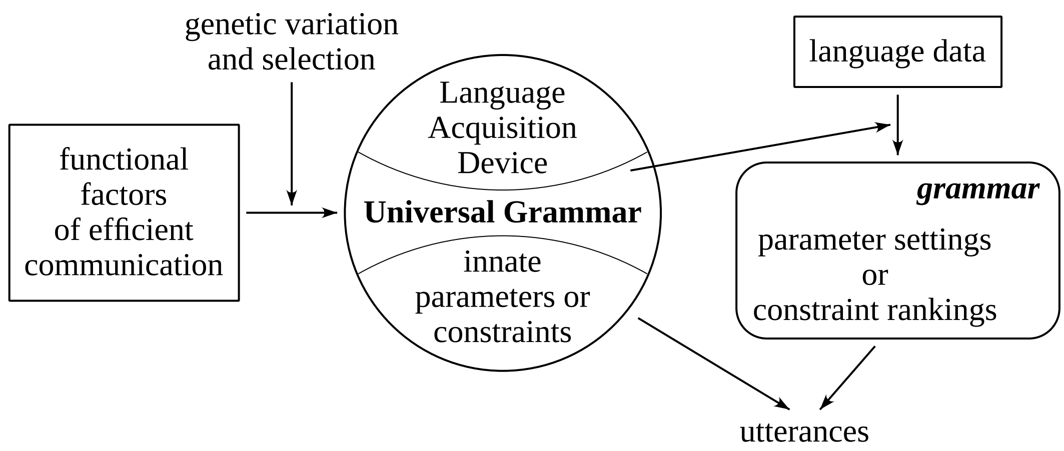


Figure 1
The nativist view of the relation between function and grammar

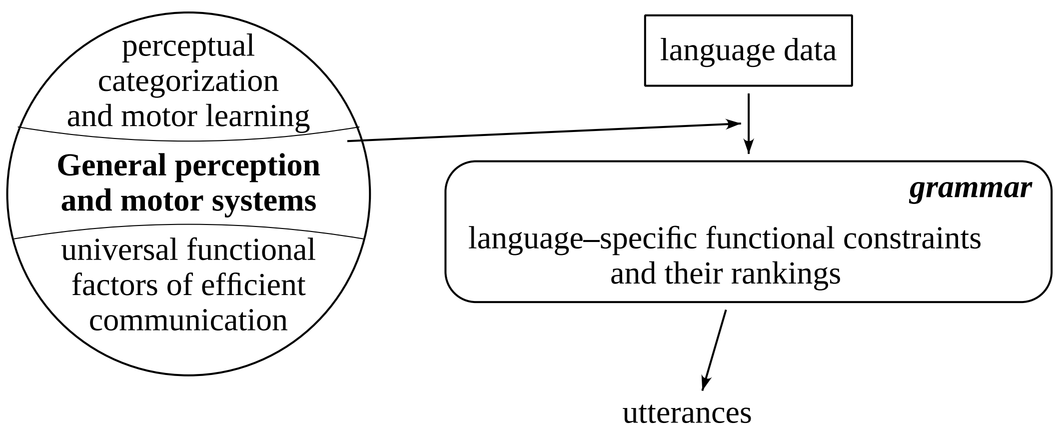


Figure 2
The empiricist view of the relation between function and grammar

things like the lexicon and much of morphology, which must also be there in any theory of phonology).

It is perfectly possible that some parts of the capabilities of human communication are innate and other parts are learnable. But this possibility cannot reconcile the two standpoints. The generative theories of phonology have ascribed every absolute or relative universal to a property of UG, expressing all phonology representations in terms of hybrid cognitive features that do not distinguish between articulation and perception. By contrast, the theory of functional phonology (Boersma 1998c), holds that universals are directly caused by general properties of cognitive organization, motor behaviour, and perception, expresses functional principles directly in the grammar, and makes a principled distinction between articulatory and perceptual representations. Fortunately, the two standpoints make different predictions with regard to what phonological universals and typology should look like, so that it is largely an empirical question which of the two is the more accurate model of reality.

1.2 Innateness and universals

I will assume that the two standpoints share the following positions:

(1) *Standpoints shared between nativism and empiricism*
  a. All innate properties of language are universal.
  b. All arbitrary universals are innate.

Statement (1a) is expressed in Fig. 1 as the universality of UG, and is compatible with Fig. 2 because that figure does not allow any innate properties, so any statement about their attributes is true. As for statement (1b), Fig. 1 says that *all* universals must be innate, so this must hold a fortiori for the arbitrary ones; Fig. 2 says that all universals must be functional, i.e. non-arbitrary, so that arbitrary universals do not exist, so that all those that do exist may as well be innate.

This means that there are two empirical tests for distinguishing the two standpoints: finding a counterexample to an alleged universal would reduce the number of possible innate principles; and finding a universal that cannot be related to a communicative or other function, would increase the number of certainly innate principles. Unfortunately, neither the nativist nor the empiricist standpoint is easily falsified. When confronted with an exception to a proposed universal, generative grammarians tend to modify their view of UG in such a way that it incorporates the newly attested facts. Likewise, when confronted with a seemingly arbitrary universal, functional grammarians tend to answer that a functional explanation must exist even if they cannot find it on the spot. Less cynically, we could hope that there is some agreement to the idea that any necessity of unbridled extensions to UG would weaken the case for nativism, and that any large-scale postponing of functional explanations would weaken the case for empiricism.

1.3 Universals: sometimes arbitrary or always functional?

While many of the principles proposed in UG must still show their functionalism as a result of biological evolution, it is crucial that some of these principles are expected to be *arbitrary* or used outside the realm in which they are directly functional. For instance, it seems arbitrary that relative clauses but not complement clauses should be islands for *wh*-movement, and an arbitrary re-use of the sonority hierarchy in various unrelated phonological processes would clearly indicate that this hierarchy is innate.

The empiricist standpoint, on the other hand, must hold that **no universals are arbitrary** and all of them are immediate results of functional principles like minimization of confusion and effort. For instance, Boersma (1998a; 1998c: §19.2.2) shows that the “sonority” hierarchies for syllabification and the transparency to nasal spreading are not exactly equal, and that the difference is precisely where we would expect it to be from functional considerations; specifically, the segment /h/ patterns with the obstruents in its preference for being in the onset of a syllable rather than in the nucleus, and it patterns with the sonorants in its transparency to the spreading of nasality, two facts considered paradoxical by Gnanadesikan (1995) but easily explained by the idea that speakers want syllable nuclei to be voiced (for purposes of the implementation of stress, tone, and rhythm) and that listeners do not hear much difference between nasalized and non-nasalized glottal or pharyngeal fricatives. As another example, the concept of mora has been applied in theories of syllable weight as well as contour tones; but the set of weightbearing units typically includes *any* consonant, because any consonant tends to contribute to the duration of the rhyme, whereas the set of tone-bearing units understandably tends to include only sonorant segments, so that rhymes of short vowel plus obstruent in e.g. Lithuanian and Limburgian must be considered monomoraic. Thus, sonority hierarchies and prosodic hierarchies seem to be shaped in such a way that they come out as predictable on the basis of functional considerations that depend on their use in the language at hand.

So we have seen her two case in which the generative theories of phonology have expected arbitrariness while it does not in fact occur. If it is an empirical question which of the two standpoints is correct, these two examples of clearly central phonological phenomena weaken the case for detailed innateness.

1.4 Arbitrariness: sometimes innate or always learned?

Logically reversing the empiricist standpoint, we see that all arbitrary parts of grammar must be language-specific, which entails that **all arbitrary grammar must be learned**.

Even in a theory that expresses phonetic principles directly in a phonological grammar (Boersma 1998c), not all phonology can be reduced to phonetics: "only" the generative theories about the non-arbitrary parts of phonology, i.e. autosegmental phonology and feature geometry, can be replaced by a theory of interacting articulatory and perceptual constraints. As for largely arbitrary things like the morphology of English strong verbs, these must be accounted for in any theory by an appeal to a general capability to generalize.

In §6, I will show how a child can learn the past tense of English verbs, exhibiting phenomena of lexicalization, generalization, and overgeneralization.

1.5 Minimal innateness of phonology

The current paper will develop an empiricist account of phonological acquisition. The following properties of the learner will not be explained, and may be regarded as innate. First, there are two innate peripheral systems:

(2) *Innate peripherals*
  a. *Speech production peripherals*: the versatility of the tongue and larynx.
  b. *Auditory perception mechanisms*: the ability to detect spectral frequency distributions, periodicity, noisiness, temporal coincidence and ordering, and intensity.

Then, there are four relevant cognitive capabilities:

(3) *Innate cognitive capabilities*
  a. *Categorization*: the entities perceived in the world (dogs or acoustic features) are put into classes of partial equivalence. This is a prerequisite for the following three capabilities.
  b. *Abstraction*: higher-level constructs are recognized and classified: simultaneity relations, sequential relations, programs (Powers 1973).
  c. *Wild generalization and extrapolation* on perceived data. Predictive power is a clear evolutionary advantage.

d. *Arbitrary symbols*: their storage, retrieval, and access.

It should be noted that none of these cognitive capabilities is specific to speech. As for (3a): regarding the short time that dogs have been around on the earth, the concept of "dog" cannot be innate with humans, though it forms a clear category for us all. As for (3d): though sign languages use a lot of seemingly iconic symbols, these languages are not generally mutually intelligible, so that the language-specific lexicon of signs must be regarded as arbitrary (this is uncontroversial). Unless we can regard the set of phonological signing features as innate as well (perhaps because our genes remember a long period of sign language that existed before the human race started to speak), this domain-aspecificity is clear evidence of the fact that the lexicon does not have to refer to a universal set of phonological features.

And the capabilities in (3abc) are not specifically human: even chinchillas are known to show perceptual categorization of human speech (Kuhl & Miller 1978). The extensive use of arbitrary symbols, however, seems to be restricted to the human race.

And there are two probably innate mechanisms for decision making (Boersma 1997, 1998c, to appear a):

(4) *Innate decision making*
  a. *Stochastic constraint grammars*: decision-making systems are implemented as constraint grammars (§2.4) with noisy evaluation (§2.6).
  b. *Gradual learning algorithmn*: error-driven learning by gradual reranking of constraints (§4.1) is realistic (§4.2) and robust (Boersma 1998c: 320).

These mechanisms are not specifically human. Evaluation noise explains probabilistic behaviour in animals as well, and the learning algorithm explains probability-matching behaviour in animals. For instance, a rat will have a probability of about 75 percent of searching for food where food has a history of turning up 75 per cent of the time. Probability-matching behaviour leads automatically to maximum-likelihood decision criteria (Boersma 1998c: 339), with obvious advantages for the survival of the being. Note that there is an element of non-functional arbitrariness, i.e. overt innateness, in this: probability-matching behaviour may be functionally non-optimal since it does not lead to decisions that maximize likelihood (outside the criteria).

Finally, there are several innate functional drives:

(5) *Innate functional drives*
  a. *The desire to communicate*, i.e. to understand and to make oneself understood. The evolutionary advantage of minimizing confusion in world full of dangers is clear.
  b. *Laziness*. In a world with limited resources, animals and plants who economize on the expenditure of the available food are usually in an advantage, everything else being equal.

In §3, I will show that all of these ten innate properties are needed to understand the acquisition process. The ultimate list may be longer, but does not include the following phenomena, which must be considered learnable:

(6) *Learned elements of phonology*

a. *Phonological features*. Languages may distribute vowels at random points in the perceptual spectral space (F1-F2), or they may order them in a symmetric manner with a small number of perceptual heights (F1) and front/back places (F2), especially if the number of vowels is large. Also, the learnability of sign languages shows that other than oral phonological feature values can be learned with ease.

b. *Phonological feature values*. These originate because of the finite number of acquired perceptual categories and the finite number of acquired articulatory gestures. (no selection pressure; on the contrary).

The points (2) through (6) may not seem controversial at all. Nevertheless, much of the theorizing in phonological theory has centred around discovering the universal properties of phonological features and their values; it is these substantive theories of autosegmental phonology and feature geometry that the theory of functional phonology aims to replace. But the range of application may be larger than that, as we will see in the next section.

1.6 Example: functional, universal, arbitrary, and language-specific stress

In Boersma (1998c), I did not discuss any subject outside the realm of single phonological features and their combinations, so I will now explain what *stress* systems would look like if they adhered to general principles of human motor behaviour and perception.

The first, functional, use of stress is to raise important parts of the utterance above the background noise. Important parts of the utterance, like new information, focus, or contrasts, should receive extra prominence, and unimportant parts, like given or repeated information, may be destressed. This relates to the functional principle of minimization of confusion: in a world in which speech often has to travel large distance through background noise, sounds that are louder, higher, or longer, have less chance of being missed or misinterpreted by the listener. For instance, the human ear is more sensitive to higher frequencies, especially if the loudness level is low (Fletcher & Munson 1933). This means that the following factors can contribute to perceived relative prominence:

(7) *Auditory factors contributing to perceived prominence*

a. Intensity stress.

b. Vowel length.

c. High F0: high tones are perceived better than low tones (although a low F0 will give better resolution of vowel height).

d. High F1 (open vowels): more open vowels have higher first-formant frequencies (openness also contributes synergistically to intensity).

e. High F2: one hears higher second-formant frequencies in going along the sequence [u-o-a-e-i], as can be seen in Fig. 3. These become increasingly more salient perceptually. Passy (1891: 81) ascribes the discovery of this pattern to M. Trautmann.

Note that this list does not mention any of the effects caused by the iconic correlation between high frequency and high energy, which arises because louder noises will have higher frequencies and a more rising spectral slope (Stevens 1971). These effects, which

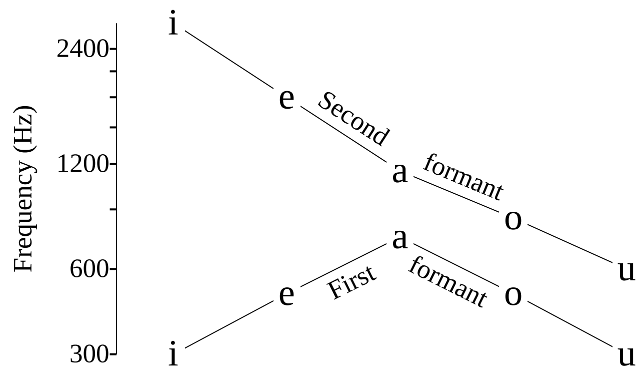


Figure 3
The two main auditory spectral properties of the peripheral vowels.
The vertical scale is linear in ERB units.

do not contribute much to minimal confusion by themselves, include the contribution of stronger high formants to perceived stress (Sluijter 1995) and the perception of stress as a result of the vicinity of a rising or falling pitch movement.

We expect the properties in (7) to coincide in the language-specific implementations of functional sentence stress. In Dutch (Van Bergem 1993), we see that words that bear no sentence accent have more closed vowels than stressed words (though this could have other causes: coarticulation with neigbouring consonants, or a free ride on the implementation of increased loudness or enhanced height contrast).

Apart from the just-mentioned functional use of perceived prominence, there are at least two thinkable arbitrary applications: first, for contrasting lexical items (arbitrary lexical stress); second, for marking constituent structure (arbitrary predictable stress in words or phrases). If used arbitrarily, the aspects that could contribute to perceived prominence (like tone and length) can be used independently: in Czech, lexical vowel length is unrelated to the fixed arbitrary initial intensity stress. However, we often find positive correlations even in arbitrary stress:

(8) *Auditory features positively correlated with arbitrary stress*

a. Stress-to-length: Italian lengthens vowels in lexically stressed syllables.

b. Weight-to-stress: in Latin, penultimate long syllables must be stressed. There is no need for an innate constraint WEIGHT-TO-STRESS. Rather, the correlation is functional (and symmetrical), because weight and stress both contribute to a single perception of prominence, in the same way that e.g. tongue-backing and lip-rounding contribute to a single perception of low F2 (perceptual backness).

c. Stress-to-F1: in Italian (Leoni, Cutugno & Savy 1995), /a/ in unstressed position has the same F1 as /ɛ/ and /ɔ/ in stressed position.

d. Stress-to-F2: in Proto-Indo-European, lexically accented syllables had /e/ (Greek /léipɔː/ 'I leave'), which changed to /o/ in case of a stress shift (/léloipa/ 'I have left'). This /e/ - /o/ ablaut shows a falling F2, with an F1 that remains the same (Fig. 3). This explanation contradicts that of Hirt (1939:219) and Lehmann (1993: 139), according to whom the language had a *pitch* accent system when this ablaut arose. The relation between F2 and stress is functional, that between F2 and pitch is either arbitrary or derived as a secondary correlation based on the stress-F2 and stress-pitch relations.

The factors may be perceived independently as separate perceptual features like [tone] and [vowel length], or combined into a single perceptual feature like [stress], depending on the language (or person). We see that some phonological features must be language-dependent, as proposed in the discussion (§1.5) of the alleged innate property (6a), and there seem to be no arbitrary universal principles of cooccurrence.

2 Functional phonology

Compared to generative theories of phonology, the theory of functional phonology (Boersma 1998c) differs with respect to the nature of phonological representations, the grammar model, the determination of the optimal output, and the way in which typological predictions are made.

2.1 Articulatory versus perceptual representations

The theory must make a principled distinction between articulatory and perceptual representations, because otherwise we could not evaluate to what extent utterances honour the functional principles.

Generative phonology commonly uses hybrid phonological features, switching between articulatory and perceptual correlates as comes in handy. Pinker (1994: 166), for example, correctly explains that vowels can be characterized by spectral qualities, and that the common order of i-a in such phrases as /nˈɪknæk/ ‘knick-knack’ and /flˈɪpflɒp/ ‘flip-flop’ is to be explained as a lowering of some part of the spectrum (namely, the second formant, F2). However, Pinker then (p. 167) explains the relation between articulation and perception by saying that the ‘small resonant cavity’ of [i] ‘amplifies some higher frequencies’ and the ‘large resonant cavity’ of [a] ‘amplifies some lower frequencies’. This reasoning is already somewhat dubious, since in going from [a] to [u] the second formant continues to fall (Fig. 3), while the resonant cavity becomes smaller again. And indeed, Pinker (p. 168) concludes that ‘the rationale is based on how the tongue produces the vowels’. This is incorrect, since longer phrases tend to include a vowel that approaches [u]. Pinker mentions /tʰˌɪktʰˌæktʰˈou/ ‘tic-tac-toe’, and in Dutch we have /pˌifpˌɑfpˈuf/ ‘bang, bang!’. While [i-e-a] and [u-o-a] could be identified with an opening jaw (though more easily with a raising F1), the sequence [i-e-a-o-u] has no such direct articulatory correlate (perhaps lip-rounding activity?), but it has a straightforward perceptual correlate: a falling F2, as seen in Fig. 3. These F2 sequences are used in languages in places other than reduplication. In §1.6 we have seen an example of their functional use, namely as a way to contribute to perceived relative prominence of parts of the utterance. An example of arbitrary use of this phonetic continuum is the zero-i-a-u-u ablaut in Arabic (Guerssel & Lowenstamm 1996).

Thus, the failure to distinguish between articulation and perception obscures the straightforwardness of the organization of human speech, and leads to a lot of confusion and unnecessary proposals, as we will see again in §3.1.

The hypothesis of functional phonology is that the distinction between articulation and perception plays a role in the language user’s organization of spoken language. The perceptual result of lowering the velum, for example, heavily depends on the degree of simultaneous oral constriction: nasalization has a much smaller perceptual influence on a

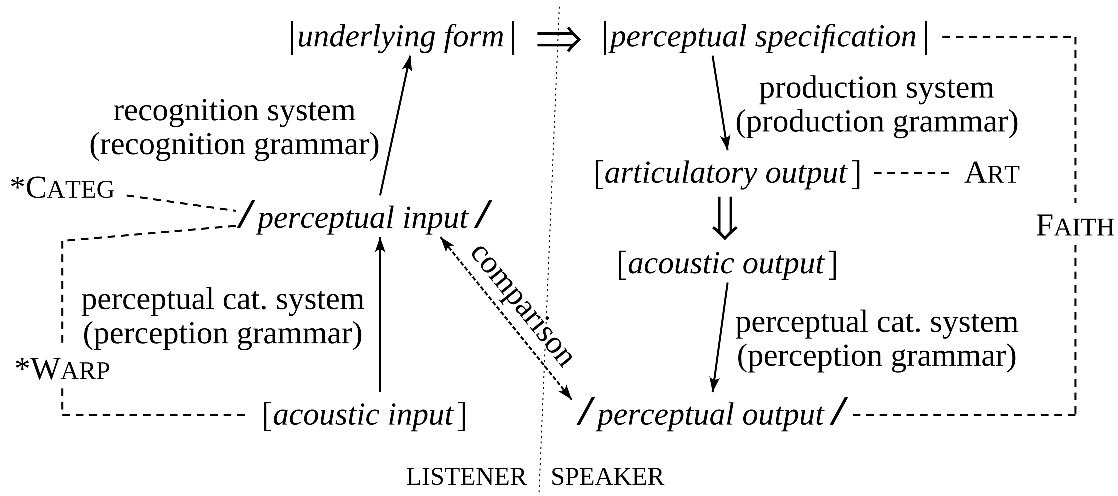


Figure 4
The grammar model of Functional Phonology

vowel, [h], or a liquid, than on obstruents like [t] or [s], which would turn into the nasal stop [n]. This asymmetry between articulation and perception leads to a perception-based hierarchy of transparency to nasal spreading (§1.3).

2.2 Grammar model: representations

Our basic model of the speaker's production grammar is that of a specification-articulation-perception triad (Fig. 4):

(9) *The three representations of the production grammar*
a. *Perceptual specification*. The underlying form is specified in perceptual terms derived from acoustic quantities, and is equated to the perceptual specification of the utterance. The perceptual dimensions include timbre (frequency spectrum), loudness (intensity), pitch (periodicity), noisiness (harmonicity), and their simultaneous and sequential combinations.
b. *Articulatory output*. The speaker can try to implement a given perceptual specification by any of a number of candidate articulations. The articulatory dimensions include the lengths, tensions, and positions of the lung, larynx, pharynx, velum, tongue, and lip muscles, and their simultaneous and sequential combinations.
c. *Perceptual output*. Each articulatory candidate leads to a perceptual result, which is the automatic acoustic result of the articulation, as perceived by the speaker's perceptual categorization system (perception grammar).

As an example, |r| could be a shorthand for perceptual specifications of [trill] and [place: coronal] (the perceptual place feature corresponds to spectral properties of the perceived sound). It can be implemented as [r], i.e. a certain alveolar approximation of the tongue tip, or as [ʀ], i.e. a certain uvular approximation of the tongue body, both in combination with contracting lungs, adducted glottis, and open epiglottis and lips. We could write the perceptual results of these articulations as /r/ and /ʀ/, respectively. Both of these notations are shorthands for a perceptual feature [trill], though their values for the continuous perceptual feature [place] are different.

The solely perceptual underlying form is the same as the 'image acoustique' of Saussure (1916), which he defines as 'l'empreinte psychique [du] son [matériel], la représentation que nous en donne le témoignage de nos sens' (p. 98). The idea of a solely perceptual specification is a special case of Powers' (1973) more general theory of behaviour as the control of perception, and is contrary to the generative idea of combined articulatory-perceptual (but mainly articulatory) features (Chomsky & Halle 1968), and also contrary to Flemming's (1995) and Steriade's (1995) idea of a mixed articulatory/perceptual representation.

Because of the classificatory nature of perception, both the perceptual specification and the perceptual output are discrete representations. The traditionally separate modules of postlexical phonology and phonetic implementation sit together side by side in a single production grammar, and yet there is a discrete phonological surface structure, which is the sound of the utterance, as classified into a finite number of perceptual categories by the speaker's perception grammar. This idea is congruent with generative theories of discrete feature values (Chomsky & Halle 1968), but transcends the idea of direct manipulation of acoustic cues (Jun 1995; Steriade 1995; Boersma 1998: ch.10), which disregards the human capabilities of categorization.

2.3 Assessment of output candidates: functional constraints

The speaker will assess every output candidate on the basis of the two innate functional principles identified in §1.5:

(10) *The implementation of functional principles as constraints*
  a. Faithfulness constraints implement minimization of perceptual confusion.
  b. Gestural constraints implement minimization of articualtory effort.

The speaker determine the extent to which the perceptual result is similar to the perceptual specification. By favouring candidates with a great perceptual similarity to the specification, she honours innate property (6a), the desire to make herself understood. In the production grammar, this drive for minimization of perceptual confusion is implemented as a plethora of faithfulness constraints (summarized as FAITH in Fig. 4). For instance, both candidates [r] and [ʀ] produce a perceptual result that is faithful to the [trill] part of the perceptual specification, i.e. both honour the constraint *DELETE (trill). This example shows that two very different articulatory gestures can faithfully implement the same perceptual specification. The common occurrence of [ʀ] in the speech of children whose adult environment says [r], is readily explained in a theory that distinguishes between articulatory and perceptual representations. A typical full-fledged family of faithfulness constraints is:

(11) *REPLACE (*feature*: $value_1$, $value_2$ / *condition* / *left-env* _ *right-env*):
Do not replace a specified value ($value_1$) on a perceptual tier (*feature*) with a different value ($value_2$), under a certain *condition* and in the environment between *left-env* and *right-env*.

Other faithfulness constraints militate against insertion of surface material and deletion of underlying material, or against the loss of specified simultaneous and sequential relations between features.

The speaker will also measure the articulatory effort associated with each articulatory candidate. By favouring candidates that require only a small effort, she honours innate property (6b), the desire to be lazy. In the production grammar, this drive for minimization of articulatory effort manifests itself as a number of gestural constraints (summarized as ART in Fig. 4). A typical full-fledged family of these is:

(12) *GESTURE (*articulator*: *gesture* / *distance*, *duration*, *precision*, *velocity*):
Do not perform a certain *gesture* with a certain *articulator*, along a certain *distance*, for a certain *duration*, and with a certain *precision* and *velocity*.

Other articulatory constraints militate against the synchronization of two gestures or against the coordination of two simultaneous gestures.

Thus, far from being extralinguistic, production and perception factors make up the non-arbitrary features of the grammar, possibly including some universal ones.

2.4 Constraints are violable and ranked along a continous scale

As in the original Optimality Theory (Prince & Smolensky 1993), most of the constraints can be violated in speech, as we see in the usual rendering of |sa| as [sa], which violates, among others, *GESTURE (tongue blade: groove), or in the common rendering of |an| as [am] before [pa], which violates *REPLACE (place: coronal, labial) since [am] (i.e. velum lowering plus lip closure) is heard as /am/ (i.e. nasal and labial place). From the rendering of |sa| as [sa], which is probably perceived as /sa/, we see that the speaker apparently prefers to stay faithful to the underlying form, at the cost of violating the anti-tongue-groove constraint. Thus, it seems that a constraint like *DELETE (sibilant) must dominate that gestural constraint. We account for this by proposing that every constraint has a ranking value along a continuous ranking scale: the higher this ranking value, the more important this constraint.

The assumption of a continous ranking scale is different from the original Optimality-Theoretic conception of the grammar as determined solely by the ranking *order* of the constraints. This move enables us to account for variation, gradual learning curves, gradient grammaticality judgments, and frequency-dependent time orderings in the process of acquisition.

2.5 Noisy evaluation

From all the candidate articulations, the speaker will choose a single optimal one, called the WINNER. To determine this winning candidate, the speaker first adds an amount of normally distributed noise, different for every evaluation, to the ranking value of each constraint:

(13) *Noisy evaluation*
disharmony = *ranking value* + *noise*

The speaker subsequently sorts the constraints according to decreasing disharmony. For instance, if the ranking value of constraint A is 90, and that of B is 80, and the noise for these constraints at this particular evaluation is –2.1 and +0.7 respectively, the disharmonies will add up to 87.9 and 80.7 respectively, and the constraints will thus be sorted in the order AB. The evaluation of the candidates will then proceed in the normal Optimality-Theoretic way, with a strict ranking of the violable constraints A >> B. As an example, suppose that the underlying form is |*spec*|, and there are two articulatory candidates [art_1] and [art_2] with respective perceptual results /$perc_1$/ and /$perc_2$/. If the first candidate violates the constraint B and the second violates A, our evaluation tableau becomes

(14) *Evaluation of articulatory candidates and their perceptual results*

|*spec*|	A (90–2.1)	B (80+0.7)
☞ [art_1] /$perc_1$/		*
[art_2] /$perc_2$/	*!	

Note that two representations show up for each candidate, because some constraints evaluate the articulations [art_i], and some constraints evaluate the similarity of the perceptual results /$perc_i$/ to the specification |*spec*|. In (14), the first candidate wins because it violates only the lower-ranked constraint. We can see that if the ranking values of the constraints were closer, e.g. 82 and 80, respectively, the ranking of the constraints at evaluation time will sometimes be the reverse of (14), depending on the accidental values of the nose component. For instance, of the noise values are –2.1 and +0.7 respectively, we get

(15) *Variation caused by noisy evaluation*

|*spec*|	B (80+0.7)	A (82–2.1)
[art_1] /$perc_1$/	*!	
☞ [art_2] /$perc_2$/		*

In a certain percentage of the evaluations, B will outrank A although it is ranked somewhat lower, and the second candidate wins. If the mean of the noise is 0 and its standard deviation is 2, the first candidate will win in 76 percent of the cases, and the second candidate in 24 percent (Boersma 1997; 1998c: 332).

Thus, noisy evaluation leads to variation in the output. This is exactly what we need to model the acquisition process realistically. In fact, we need three things:

(16) *Requirements for a realistic model of acquisition*
- a. A continuous ranking scale (§2.4).
- b. Noisy evaluation (§2.5).
- c. A learning algorithm that modifies the ranking values by small amounts at a time (§4.1).

In §4 and §6, we will see how a model with these ingredients manages to learn functional and arbitrary phonology, respectively.

2.6 Constraints are not universal

It will come as no surprise that a theory that does not recognize innate phonological features, must also propose that the constraint set is not innate and must be learned anew by every child. In fact, I would like to maintain that the far majority of constraints are not even universal.

For instance, Chomsky & Halle (1968: 5) identify a universal constraint against simultaneous occurrences of the feature values [–consonantal] and [+strident]. All theories that use hybrid phonological features have had to express this fact formally as a universal redundancy rule or as a universal constraint. In generative Optimality Theory (Prince & Smolensky 1993), for instance, an input specification of simultaneous [–consonantal] and [+strident] must be possible according to the maxim of richness of the base, so that the non-occurrence of this combination in the output must be attributed to a universally undominated filter constraint like *[–consonantal, +strident]. These filters can be done away with by distinguishing between articulatory implementations and perceptual output representations: since it is phonetically almost impossible to produce sounds that are perceived as strident and sonorant at the same time, no articulatory candidate will produce a perceptual output that contains this simultaneous combination, which means that the optimal candidate will automatically violate at least one of the constraints *DELETE (strident) or *DELETE (sonorant). Exeunt universal grammatical co-occurrence restrictions.

Now that the theory of functional phonology has been introduced, I will show how it expresses typology (§3), acquisition (§4), and the preservation of typological relations throughout the acquisition process (§5).

3 Typology: the local ranking principle

Later on in this paper, I will show how an empiricist theory of phonology can account for the conservation of typological universals throughout childhood. First, however, I must show how such a theory predicts typologies, i.e. how it predicts what kinds of languages are possible and what kinds of languages are impossible.

Apart from positing a set of fixed principles, generative grammar allows a set of binary parameters or rankable innate constraints to account for crosslinguistic variation. Together, the principles and constraints or parameters predict a typology of possible and impossible languages. In a principles-and-parameters approach (Chomsky 1981), a typology is generated by choosing all possible parameter values. For instance, a system of 30 independent binary parameters yields 2^{30} different grammars. In an Optimality-Theoretic approach, a typology is generated by permuting across all possible constraint rankings. For instance, a system of 30 constraints yields a factorial typology of 30! different grammars. The actual number of generated languages is much lower than that, because of the lack of interaction between many pairs of constraints and because of some constraint families must be ranked in a universal way. For instance, Prince & Smolensky (1993) propose a universal constraint hierarchy for nucleus avoidance as a function of segment type. Boersma (1998c) proposes a principle for predicting what pairs of constraints must be ranked universally and what pairs can be ranked in a language-specific way:

(17) *Local ranking principle (LRP)*
Two constraints whose arguments, conditions, or environments differ in a single respect, may be universally ranked. All other pairs are ranked in a language-specific way.

This proposal leads to the common property of humans that different scales of articulatory effort and perceptual contrast are to a large degree incommensurable. The universal rankings are established on the basis of local hierarchies:

(18) *Local ranking within functional constraint families*
a. Articulatory constraints are ranked by articulatory effort, but not according to a global measure.
b. Faithfulness constraints are ranked by the degree of perceptual confusion, but not according to a global measure:

For the prototypical gestural constraint (12), this is made explicit as follows:

(19) *GESTURE is ranked higher:
a. if the *distance, duration, precision*, or *velocity* is higher,
b. and everything else is equal;
otherwise, these constraints can be ranked in a language-specific manner.

For the prototypical faithfulness constraint (11), the LRP is made explicit as follows:

(20) *REPLACE is ranked higher:
a. if *value*$_1$ and *value*$_2$ are further apart:
*REPLACE (ɛ, i) >> *REPLACE (ɛ, e)
(it is worse to implement |ɛ| as /i/ than to implement |ɛ| as /e/)
b. or if the *condition* enhances discriminability (e.g. release burst):
*REPLACE (place: *x*, *y* / plos) >> *REPLACE (place: *x*, *y* / nas)
(e.g., it is worse to implement |t| as /p/ than to implement |n| as /m/)
c. or if the *environment* enhances discriminability (e.g. release into vowel):
*REPLACE (n, m / _ V) >> *REPLACE (n, m / _ C)
(the loss of a specified place is worse in onset than in coda)
d. or if *value*$_1$ is less frequent:
*REPLACE (place: lab, cor / nas) >> *REPLACE (place: cor, lab / nas)
(it is worse to produce a specified labial as a coronal than the reverse)
e. and everything else is equal;
otherwise, these constraints can be ranked in a language-specific manner.

Thus, local confusion and effort rankings are strict and can reduce factorial typology. On the other hand, *global* measures of confusion and effort (e.g. Boersma 1998c: eqs. 4.24 and 7.4) may still be useful for predicting cross-linguistic MARKEDNESS. By determining a cross-linguistic average ranking value, these measures predict statistical tendencies, sometimes with a skewing so severe that the current number of languages on earth that exhibit a certain phenomenon is likely to be zero (Boersma, to appear b).

3.1 Example: typology of place assimilation

Consider the following implicational universals for the assimilation of plosives and nasals to a following consonant (Mohanan 1993):

(21) *Universals of place assimilation*
  a. If non-coronals assimilate, so do coronals (with the same manner, in the same environment);
  b. If plosives assimilate, so do nasals (with the same place, in the same environment).

These universals generate the following typology of labial and coronal stops that can be targets of assimilation:

(22) *Typology of place assimilation*

		p	t	m	n
(I)	nothing assimilates	–	–	–	–
(II)	only coronal nasals assimilate	–	–	–	+
(III)	all nasals assimilate	–	–	+	+
(IV)	all coronals assimilate	–	+	–	+
(V)	all nasals and coronals assimilate	–	+	+	+
(VI)	all plosives and nasals assimilate	+	+	+	+

Thus, of the 16 logically possible patterns, only six are predicted by (21) to occur. I will now show that theories with hybrid constraints have trouble accounting for these facts, and that theories that distinguish between articulatory and faithfulness constraints accurately predict the attested typology.

3.1.1 *Serialism translated to constraints*. In derivational phonology, place assimilation of nasals is expressed with a feature-changing rule like “the place of a nasal consonant changes to the place of a following consonant”, or with an equivalent feature-filling and/or autosegmental formulation (“spread place”). Such a rule can be translated directly into an imperative constraint like:

(23) NASSIM
  The place of a nasal stop is equal to the place of any following consonant.

This constraint, or an equivalent like SPREAD(PLACE) (Padgett 1995), is clearly a constraint rooted in the interaction of functional principles: a nasal assimilates to the following consonant because the slight disadvantage of increased perceptual confusion ([m] and [n] are hardly discernible anyway before [p]) is outweighed by the advantage of not having to produce a closing and opening gesture of the tongue blade (and the bilabial gesture needed for [m] merges with that of [p]). But it is also a *hybrid* constraint in that expresses an interaction instead of the principles themselves. As such, it is a good candidate for an innate universal constraint, created in the process of biological variation and selection (Fig. 1). The constraint shares this property with Archangeli & Pulleyblank’s (1994) GROUNDING CONDITIONS, whose Optimality-Theoretic counterparts (Pulleyblank 1993) are considered innate, as is apparent from the associated principles-and-parameters-based learning algorithm (Pulleyblank & Turkel 1996).

But NASSIM is not enough to explain typology (22). We have to express the analogous tendency for coronals to assimilate, and to account for the possibility that everything assimilates:

(24) CASSIM
The place of an underlyingly coronal stop (or a stop underspecified for place) is equal to the place of any following consonant.
ASSIM
The place of a stop is equal to the place of any following consonant.

The effects of these three assimilatory constraints arise in competition with a single faithfulness constraint:

(25) PARSE (place)
Preserve the specified place in the output.

If this constraint outranks all three assimilation constraints, there is no assimilation (type I). If ASSIM outranks PARSE (place), all stops assimilate (type VI). Now if ASSIM is ranked low, assimilation of nasals only (type III) is found with the ranking NASSIM >> PARSE >> CASSIM, and assimilation of coronals only (type IV) occurs if the ranking is CASSIM >> PARSE >> NASSIM. Finally, the ranking of both NASSIM and CASSIM above PARSE (place) causes both nasals and coronals to assimilate (type V). But one type is missing: the constraint set (23) – (25) cannot generate languages where only the coronal nasal assimilates (type II). Examples of such languages are Catalan (Recasens 1991) and Dutch. Of course, such problems can be remedied by populating UG with constraints like CNASSIM, but that goes against the generative maxim of maximum simplicity in the formulation of constraints, as formulated by McCarthy & Prince (1993: 31) in their rejection of ONSET(EXCEPT). The general idea is that if a constraint has a composite formulation, some generalization has been missed.

3.1.2 *Phonetically-based constraints*. Myers (1997) tries to express explanations based on perceptual confusion directly into markedness constraints:

(26) *Myers' approach*
ALIGN-CV
Place of a consonant that isn't immediately before a more open articulation is difficult (to distinguish).
*PL/NAS
Place exclusively associated with nasal is difficult (to distinguish).
MAX-IO (Pl)
A place target in the input must have a correspondent in the output.

In this way, Myers separates the factors of being nasal and being before a consonant. The constraints have to be conjoined in order to account for pre-consonantal place neutralization of nasals: MAX-IO (Pl) must be dominated by the conjoined constraint ALIGN-CV-&-*PL/NAS, which is violated only if both ALIGN-CV and *PL/NAS are violated, i.e. by nasals not in onset.

Typologically, (26) will predict all of the types in (22) if extended with a constraint *PL/COR, which says that 'place associated exclusively with coronal is difficult (to distinguish)’. The Catalan and Dutch type is then described by a high ranking of the triple conjunction Align-CV-&-*Pl/Nas-&-*Pl/Cor and low rankings of the single constraints and the simple conjunctions. However, (26) predicts the existence of languages with a high-ranked *Pl/Nas, which would delete place from all nasals, even in onsets. This is due to a failure to express that the confusion (or perceptual effort) of nasals is position-dependent. We need a better separation of the phonetic principles.

Despite their promise, both of the constraints Align-CV and *Pl/Nas still express the problem indirectly: there is no indication of *why* we should neutralize under these conditions, i.e. no expression of the advantage of not moving the tongue blade. Apparently, Myers seeks to maintain the hybrid representations of place and nasality that do not distinguish between articulation and perception. This is most clearly seen in his parentheses around ‘to distinguish’, which signal a shift from an account in terms of perception to one in terms of effort. That’s *perceptual* effort, of course, but Myers wants to have it that it is violated in [np] but not in [mp], although in reality the perceptual effort of reconstructing an underlying /n+p/ is larger for the realization [mp] than for [np]. Hence, (26) is not a direct expression of functional principles.

3.1.3 *Minimal directly functional constraints and rankings*. Jun (1995) cleanly separates the articulatory and perceptual constraints, and gives fixed rankings for some of them, based on considerations of hierarchies of perceptual confusion:

(27) *Jun’s approach*

Weakening
: Conserve articulatory effort.

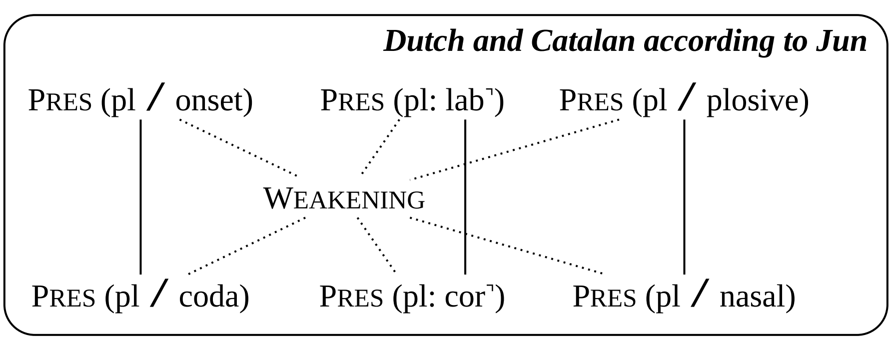
Pres (pl / plosive) >> Pres (pl / nasal)
: Preserve the perceptual place cues for plosives (or nasals).
: Universally ranked because plosives have stronger place cues than nasals.

Pres (pl: lab˺) >> Pres (pl: cor˺)
: Preserve the place cues of unreleased labials (or coronals).
: Universally ranked because unreleased labials have stronger place cues than unreleased coronals (mainly due to the rapidity of coronal stops).

Pres (pl / onset) >> Pres (pl / coda)
: Preserve the place cues in onset (or coda) position.
: Universally ranked because plosives and nasals have stronger place cues in onset than in coda position.

This set can generate Catalan and Dutch assimilation (type II):

(28)



In this grammar of crucial rankings, depicted in the style of Boersma (1998c), high constraints are drawn at the top and low constraints at the bottom; solid lines denote universal rankings, dotted lines language-specific rankings. If at least one of the universally dominated constraints is ranked above WEAKENING, this constraint set generates a type-I language. If, starting from (28), PRES (pl: lab˺) falls below WEAKENING, we get type III; if PRES (pl / plosive) falls, we get type IV. However, there is no way to generate type V, since if we let both PRES (pl: lab˺) and PRES (pl / plosive) fall below WEAKENING, there will be assimilation of any coda consonant (type VI). The only way to derive a type-V language is to add a conjoined constraint PRES (pl: lab˺) & PRES (pl / plosive).

The constraint set in (28) can generate some more grammars, namely those with PRES (pl / onset) ranked below WEAKENING. The remaining faithfulness constraints will then compete for the direction of assimilation. If only PRES (pl: lab˺) is ranked high, there is forward assimilation of all coronals to labials, as in Limburgian |kom+t| 'come+3SG' → /kømp̄/ 'comes' (this is restricted to a small group of inflectional morphemes). If only PRES (pl / plosive) is ranked high, there is forward assimilation of all nasals to all plosives. And if both are ranked high, nasals will assimilate but plosives will not, as in Low Saxon |loːp+n| 'walk-INF' → /loːpm/ '(to) walk', but |loːp+t| 'walk-PL' → /loːpt/ '(we/you/they) walk'.

Apart from having problems generating the type-V assimilation predicted by (21), Jun's account still represents an idealization of phonetic reality: his faithfulness constraints are generalizations across conditions and environment.. For instance, the separation of faithfulness relations into three independent hierarchies, hinges on a coincidence: the fact that place cues are stronger for labials than for coronals, holds for plosives as well as nasals. Likewise, the dependence of faithfulness of nasals on place or of coronals on manner is not expressed. While this may reflect an arbitrary innate condition on constraint types, we can go one step further in the direction of direct expression of functionality.

3.1.4 *Directly functional constraints and rankings*. According to Boersma (1998c), the conditions and environments for faithfulness constraints contribute additively to the ranking of the constraints. This obviates the need for constraint conjunction:

(29) *Additive ranking*

a. *GESTURE

Do not move the tongue tip (or the lips).

b. *REPLACE (place: lab, cor / plosive / _C) >>
*REPLACE (place: lab, cor / nasal / _C)

Do not implement a specified value [labial] on the perceptual place tier as the value [coronal] in the output, if the segment is a plosive (or nasal) and stands before another consonant. The fixed hierarchy is based on the relative strengths of the place cues.

Shorthand: *ap+na → atna >> *am+ta → anta

c. *REPLACE (place: cor, lab / plosive / _C) >>
*REPLACE (place: cor, lab / nasal / _C)

Same as previous, but against implementing a coronal as a labial. The relative weakness of the place cues for labials applies here, too.

Shorthand: *at+ma → apma >> *an+pa → ampa

d. *REPLACE (place: lab, cor / plosive / _C) >>
*REPLACE (place: cor, lab / plosive / _C)

Do not implement a labial plosive as coronal before another consonant. The fixed hierarchy is based on the swiftness of the tongue blade or on frequency of occurrence (Boersma 1998c: 183).

Shorthand: *ap+na → atna >> *at+ma → apma

e. *REPLACE (place: lab, cor / nasal / _C) >>
*REPLACE (place: cor, lab / nasal / _C)

Same as previous, but for nasals instead of plosives. Both the swiftness consideration and the frequency bias point to the same fixed hierarchy.

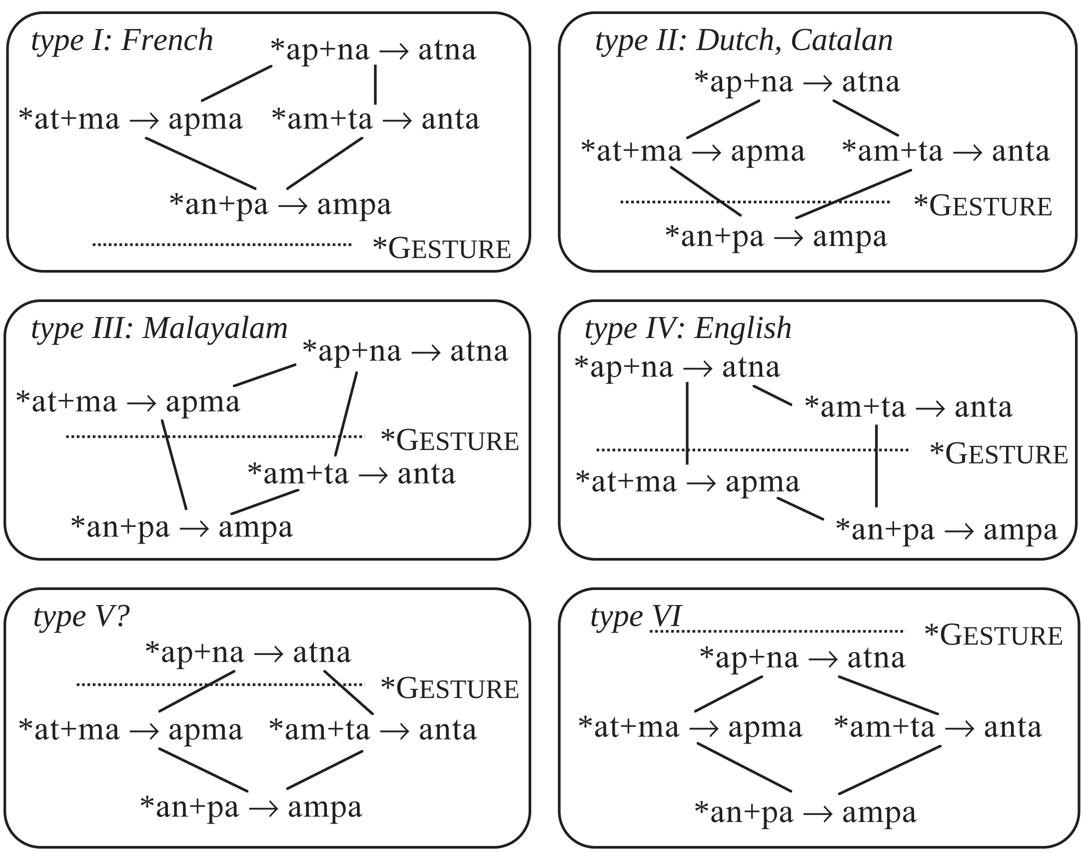
Shorthand: *am+ta → anta >> *an+pa → ampa

f. *REPLACE (place: *x*, *y* / *manner* / _V) >>
*REPLACE (place: *x*, *y* / *manner* / _C)

Do not implement a place value *x* as *y* in the output, for a segment with a certain *manner* (= plosive or nasal), in prevocalic (or preconsonantal) position. The fixed ranking is valid for all place values and both manners (plosive and nasal).

In (29b–e), the four shorthands ignore any dependence on the manner of the trigger. We see that the fixed rankings in (29a–e) generate exactly the six language types predicted by the implicational universals:

(30) *Maximally functional account of place assimilation*



The four fixed rankings give us exactly one language-specific freedom: the ranking of *REPLACE (place: cor, lab / plosive / _C) with respect to *REPLACE (place: lab, cor / nasal / _C). These two constraints differ in two of their arguments, which means that we cannot rank them in a universal way. The existence of the types III and IV reflects this freedom. In (30), the freedom is shown as a variable ranking of *at+ma → apma with respect to *am+ta → anta .

3.1.5 *Assessment*. In going through the proposals described in §3.1.1 through §3.1.4, we see that the degree of arbitrariness falls and the degree of direct explanatory functionality rises. The generative account (§3.1.1) explains nothing; Myers' account (§3.1.2) is phonetic without being explanatory; Jun's account (§3.1.3) is explanatory but puts arbitrary a-priori limits on the complexity of the formulation of constraints; Boersma's account (§3.1.4) is explanatory and splits up constraints every time a functional ranking can be formulated. The real question, of course, is which account reflects the reality of place assimilation. The generative account has to add an innate constraint to UG upon every new language fact (NASSIM, CASSIM, ASSIM, CNASSIM, not to mention the constraints for forward assimilation), only to arrive at a typology that would follow automatically from Jun's or Boersma's functional accounts. Taking this situation as a weakening of the generative standpoint (§1.2), there remain the two functional analyses. These make different empirical predictions: the restricted constraint set in (28) predicts that languages in which all nasals and all coronals assimilate, cannot exist, while the more elaborate constrant set in (30) predicts that it can.

4 The gradual learning algorithm

The grammar model of §2.2 contains a way to change the ranking of existing constraints. As shown in Fig. 4, the learner can compare her perception of the adult output with her own output, as perceived by herself. The following tableau extends the simpler tableaus like (14) in such a way that it enables the comparison:

(31) *Learning*

[*model utterance*] /*model perc*/ |*spec*|	A	B
☞ [art_1] /$perc_1$/		←*
√ [art_2] /$perc_2$/	*!→	

In this tableau, we have two more representations, as compared with (9):

(32) *The five representations needed for learning*

a. The perceptual specification |*spec*|.
b. Articulatory candidates [art_i].
c. Perceptual results /$perc_i$/.
d. The *model utterance*: the phonetic events produced by the 'adult' speaker (or other models). It can be identified with its automatic result, the *acoustic input* (Fig. 4), which is available to the learner as the input to her perception

grammar. This is a continuous representation, stated in acoustic terms, e.g. with a microscopic transcription (Boersma 1998c: 30).

e. The *model perception*: the model utterance, as perceived by the learner. This is the output of the learner's perception grammar, and is called the *perceptual input* in Fig. 4. This is a discrete perceptual representation.

The learner will compare her own winning candidate, as perceived by herself, with her perception of the model utterance. In tableau (31), she will compare /*perc*$_1$/ with /*model perc*/. If the two do not match, the learner will take some action. In tableau (31), the two do not match: the learner's winner is marked by the pointing finger, but this symbol is provided with two asterisks (*☞*) in order to show that the learner must consider it to be the wrong candidate if she compares it to an adult form. The learner's response will be to decrease the probability that the incorrect winner will win again at the next evaluation. She does this by raising the ranking values of all the constraints that are violated in the incorrect winner (in this case, only constraint B), by a small amount along the continuous ranking scale. This is shown in tableau (31) by the arrow in the second column of the first row.

Tableau (31) also shows what happens if the correct candidate (the adult form) happens to be in the learner's candidate set. In this case, /*perc*$_2$/, which is the percepual result of the second candidate, is equal to /*model perc*/, the adult form as perceived by the learner. The learner will now notice that /*perc*$_2$/ is the correct candidate, and this is shown in Fig. 4 by the check mark (√). The learner's response will be to increase the probability that the adult form will win at the hext evaluation. She does this by lowering the ranking values of all the constraints that are violated in the adult form (in this case, only constraint A), by a small amount along the continuous ranking scale. This is shown in tableau (31) by the arrow in the first column of the second row. Constraints that are violated in the learner's winner as well as in the adult form, remain in place, since their raising and lowering cancel each other out.

This scheme of promoting the constraints violated in the learner's winner and demoting the constraints violated in the adult form can be shown to converge upon a grammar compatible with the surrounding language if the constraint set is correct (Boersma 1998c: 344). By contrast, learning algorithms based on the principles-and-parameters approach (Gibson & Wexler 1994, Pulleyblank & Turkel 1996) are not guaranteed to converge, even on simple three-parameter grammars (Boersma, Dekkers & Van de Weijer, to appear; Boersma 1998c: 310–320); the constraint-sorting algorithm by Tesar & Smolensky (1998) does converge, but fails on adult data that show variation (Boersma & Hayes, to appear) or a moderate percentage of errors (Boersma 1998c: 320), and does not show the realistic gradual learning curves characteristic of our gradual learning algorithm (Boersma, to appear a; Boersma 1998c: 284).

5 Acquisition of functional constraints and their rankings

I will now briefly describe the genesis and the ranking procedure of functional constraints for a learner of her first language. A fuller account is given in Boersma (to appear a; 1998c: 275–292). The initial state of the production grammar is *empty*, which means that it does not contain any substantive material (i.e. constraints). Conceptually, this can be seen as a top ranking of all thinkable articulatory constraints, and a bottom ranking of all thinkable faithfulness constraints (Fig. 5a). But no psychological reality can be claimed for these constraints, so they are labelled as *virtual* in Fig. 5a.

The initial state of the *perception grammar* is also empty, but the acquisition of perceptual categorization will move around some *WARP, *CATEG, and RECEIVE constraints (Boersma 1998c: ch. 8), so that the child learns to classify the acoustic input into a finite numbers of perceptual categories, in a way completely analogous to her general perceptual development. This empiricist view of the acquisition of phonetic categories is compatible with what is found in perception research with infants: the following general picture can be distilled from the overview by Jusczyk, Houston & Goodman (1998), who cannot be held responsible for my interpretation:

(33) *Infant speech perception*

a. Infants show innate categorization for privative acoustic features, as a result of the presence of non-speech-specific feature detectors. For example, even one-months-olds show categorical perception (good discrimination of between-category contrasts and poor discrimination of within-category contrasts) for the voicing contrast in English obstruents (Eimas, Siqueland, Jusczyk & Vigorito 1971). This may well depend on the presence versus the absence of aspiration noise, or on the presence versus the absence of higher spectral peaks without an F_1. Plus, *untrained* non-human mammals like chinchillas (Kuhl & Miller 1978) and macaques (Kuhl & Padden 1982) show the same categorization.

b. Infants show no innate categorization for uncontroversially continuous acoustic features, since no arbitrary prior categorization exists. For instance, even six-months-olds show no categorical perception for vowel height (Swoboda, Morse & Leavitt 1976). The static perceptual feature of vowel height corresponds to the position of the first formant, so that a changing vowel height has a continuously varying effect on the listener's basilar membrane.

c. Infants show ambiguous behaviour for not-very-continuous acoustic features without clear discontinuities. For example, the ba/da/ga place contrast depends on the *locus* of the second formant, which is continuous, but also on the *movement* of the second formant, which may have discrete properties like *steady* and *moving*. Thus, Eimas (1974) found categorical perception of the auditory [bæ] - [dæ] contrast in 2- to 3-months-olds, but a continuous discrimination of non-speech-stimuli with comparable second-formant transitions.

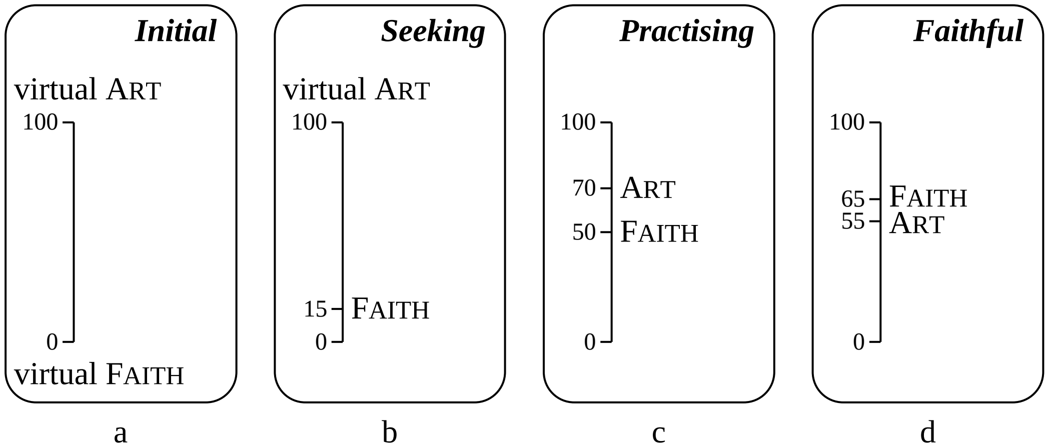


Fig. 5
The first four stages of the production grammar in the acquisition of perception and production.

Thus, infant categorical perception is more likely to occur for dynamic than for static acoustic features, and even more for privative acoustic features. This is exactly the situation we would expect if human categorization capabilities refer to general innate properties of mammalian audition. It means that phonological feature values along continuous dimensions such as vowel height must probably be learned during contact with the ambient language, and that phonological features do not belong to UG. See Boersma (to appear c) for a more complete discussion.

Apart from the categorization of perceptual features, the child will also learn to group together frequent simultaneous and sequential combinations of features, thus establishing more abstract categories. For instance, the utterance [apa] is microscopically [[a˺_pa]]. In a language where intervocalic plosives are common (i.e. practically every language), the sequence of the labial transition [[˺]] and the labial release burst [[p]] will be very common, and these two acoustic cues will be heard together as a single labial sound, despite the intervening silence [[_]]. Likewise, in languages with common homorganic nasal+plosive clusters, the cluster in [ampa] will be perceived as a single labial entity.[1]

As soon as the child recognizes a perceptual feature value and knows at least one lexical concept that uses it (learned by innate cognitive capacity 3d), she will want to use this same feature value in her own utterances, possibly as a result of Hurford's (1989) functional innateness of the reciprocity of the Saussurean sign. In constraint terms, this means that a number of faithfulness constraints will enter the production grammar from the bottom. These include several constraints for single features (TRANSMIT, *REPLACE, *DELETE, *INSERT), for simultaneous relations (TRANSMITPATH, *REPLACEPATH, *DELETEPATH, *INSERTPATH), and for sequential relations (PRECEDENCE, LEFT, RIGHT, *INTERVENE, *OVERLAP), reflecting the innate cognitive capabilities mentioned in (3a)

[1] The perception of sequences as single or double entities is handled in the perception grammar by a constraint that Boersma (1998c: 241) dubbed OCP. This constraint often causes a faithfulness violation in the production grammar when two morphemes are concatenated (Boersma 1998b). For example, the two morphemes |am| and |pa| have separate specifications for labial. After simple concatenation, the result [ampa] may be perceived with a single labial feature value if OCP (place: labial / side | silence | burst) is ranked high, thus violating TRANSMIT (place) once, which may force the emergence of an otherwise less harmonic candidate that satisfies TRANSMIT by virtue of epenthesis or dissimilation.

and (3b). The remaining cognitive ability, generalization and extrapolation (3c), will be discussed in §7.

As an example, consider the acquisition of the perception of sibilant noise by an English child. She hears the utterance [siː], which she can since recently perceive as /siː/. Noticing that adults use this perceptual form to signify the concept 'see', she will store it in her lexicon as the sign 'see'-|siː|. When trying to pronounce it herself, she will fail and perhaps produce [tiː] as the best match:

(34) *Seeking* /siː/

[siː] /siː/=|siː|	0 *DELETE (sibilant)
☞ [tiː] /tiː/	←*

The learner will notice the discrepancy between her perception of her own utterance and the form she has stored. As an automatic result of the innate learning algorithm (§4), she will promote the violated constraint *DELETE (sibilant) by a small amount along the continuous ranking scale, as shown by the arrow in (34). If the initial ranking of the constraint was 0, as in (34), the ranking will raise to, perhaps 0.1. After 150 of these learning pairs (adult form plus learner's form), the constraint will be at a height of 15, as shown in Fig. 5b.

Tableau (34) does not yet contain the adult candidate, because there is no articulation of which the child knows that it produces /siː/. But she plays with her speaking instruments and will eventually learn that a certain tongue-grooving gesture will implement the desired effect. In the production grammar, this shows up as a gestural constraint entering from above, e.g. at a ranking of 100.

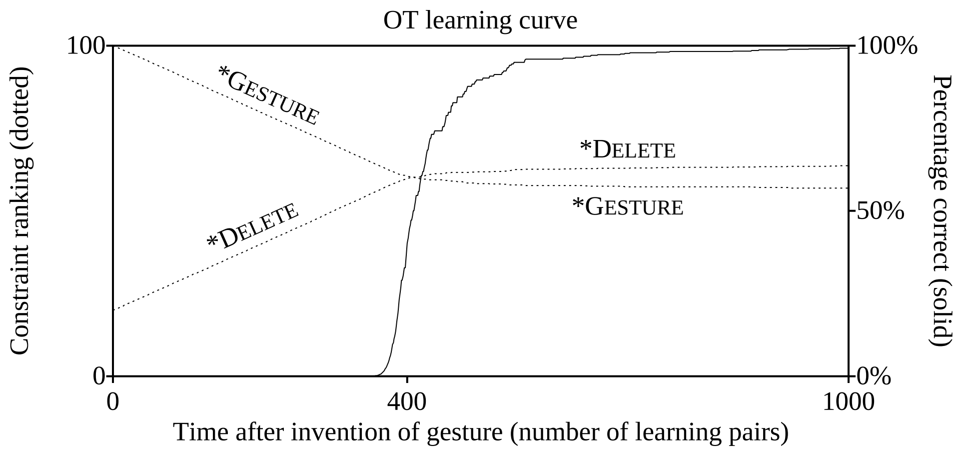
(35) *Practising* /siː/

[siː] /siː/=|siː|	100 *GESTURE (groove)	20 *DELETE (sibilant)
☞ [tiː] /tiː/		←*
√ [siː] /siː/	*!→	

Initially, the child does not manage to produce the desired result. But now that she knows how to do it, her sensorimotor learning will show up as a gradually falling gestural constraint, still with a rising faithfulness constraint. In perhaps 300 steps, she will arrive from the rankings in tableau (35) to the situation in Fig. 5c. At that point, pragmatic reranking may already sometimes reverse the order of the constraints. For instance, the child may now already faithfully imitate adult speech, since faithfulness constraints will be higher if the sole purpose is imitation, and articulatory constraints may be lower.

As the ranking values of the gestural and faithfulness constraints approach each other, there will emerge a non-zero probability that the learner produces the correct form, because some noise is added at evaluation time. This gives a realistic LEARNING CURVE (after Boersma 1998c: 284):

(36)



Learning will proceed until the faithfulness constraint has risen above the gestural constraint, maintaining a safety margin depending on the evaluation noise. The result will be a faithful outcome in almost all cases, and learning will come to an end:

(37) *Faithful* /siː/

[siː] /siː/=|siː|	65 *DELETE (sibilant)	55 *GESTURE (groove)
[tiː] /tiː/	*!	
√ ☞ [siː] /siː/		*

This situation is depicted in Fig. 5d. All faithfulness outranks all articulatory effort.[2] However, this pattern causes overly faithful outputs, and these are indeed found. Dutch four-year-old children, for instance, show a lack of place assimilation, degemination, and post-obstruent fricative devoicing at the sentence level. Spanish-speaking children (Hernández-Chávez, Vogel & Clumeck 1975), too, proceed from unassimilated to assimilated NC clusters across word boundaries. The overly faithful rendering of place for nasals is shown in this grammar:

(38) *Overly faithful nasals as a learning stage*

*DELETE (place / _V) *DELETE (place: labial) *DELETE (place / plosive)

*DELETE (place / _C) *DELETE (place: coronal) *DELETE (place / nasal)

*GESTURE

This figure and the following tableau use a Jun-like separated constraint set (§3.1.3) in order to show how the learning algorithm handles multiple constraint violations.

Originally, the child's underlying form equalled her perception of the adult form. This is shown by an equal-sign in (34), (35), and (37). But from a certain point in time, the child starts to notice that adults violate faithfulness in sentence phonology:

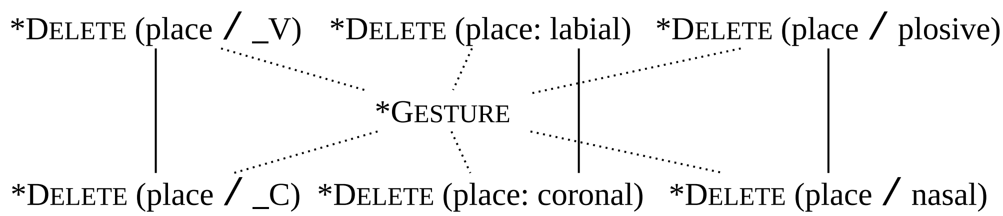
[2] According to Hale & Reiss (1998), this must be the actual starting point of acquisition, because correct perception would entail high-ranked faithfulness constraints. However, their standpoint testifies of a failure to distinguish between articulatory and perceptual representations (Boersma, to appear d).

(39) *Detecting an adult faithfulness violation*

[ampa] /ampa/ |an#pa|	*DELETE (place / nas)	*DELETE (place: cor)	*DELETE (place / _C)	*GESTURE (blade)
☞ [anpa] /anpa/				←*
√ [ampa] /ampa/	*!→	*→	*→	

Now that the learner's winning perception /anpa/ differs from her perception of the adult form /ampa/, the learner will promote the constraint violated in her incorrect winner, and demote the three constraints violated in the candidate that would have produced the correct perception. Eventually, the child will acquire the adult grammar:

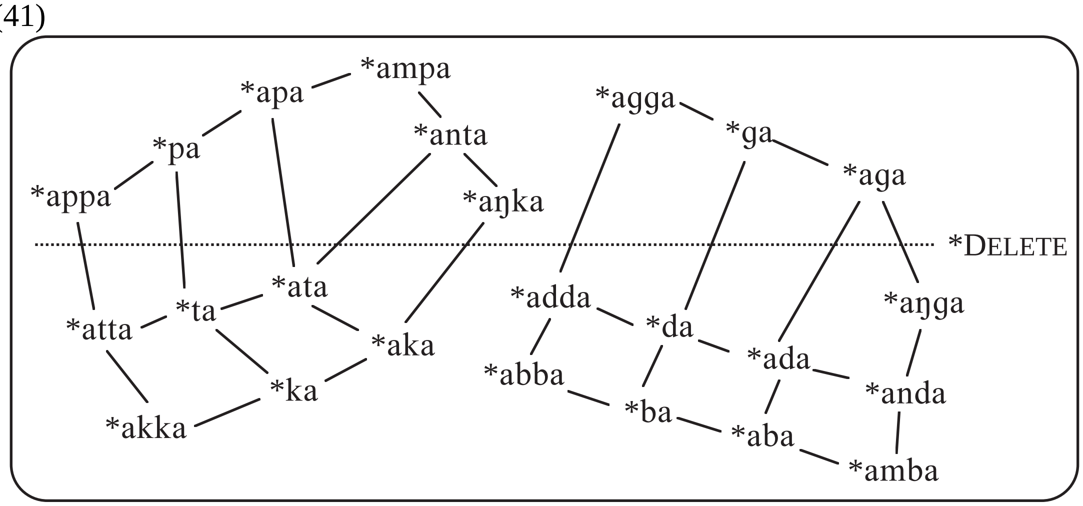
(40) *Finally, the adult stage*



6 Preserving universal rankings of functional constraints

During the acquisition process, constraints are continually promoted and demoted by small amounts. This brings up the question of how the universal rankings identified in §3 are maintained. The solution is simple. If a constraint threatens to be promoted above a constraint that is supposed to be universally ranked above it, it will push that other constraint along with it up the ranking scale; likewise, falling constraints will push down their universally subordinate family members. As an example, consider the following grammar for the implementation of a voicing contrast in obstruents:

(41)



The constraint names are shorthands for members of the implementational *[+voice] and *[–voice] families. For instance, *ampa stands for *[–voice / plosive, labial / m _ a], which is again a perceptually-oriented shorthand for an articulatory constraint like *GESTURE (much glottis widening plus wall stiffness). There are no fixed rankings between the [+voice] and [–voice] families, because these use different articulations. *Within* the families, however, we have some functional rankings, shown by the solid lines: plosives are more difficult to voice if they are further back, if they are longer, or if they are initial, and plosives are more easy to voice if they follow a nasal (Pater 1996). Grammar (41) generates a language without [g], with [bb] but no [pp], and without NC̥ clusters. During the acquisition of this grammar, there may have been an occasion on which *adda got demoted while *abba was not ranked much lower than *adda:

(42) *Demotion of* *adda *past* *abba*?*

[adːa] /adːa/ |adːa|	90+0.6 *agga	70.28–0.4 *adda	70.25–0.8 *abba	40+0.4 *DELETE (+voice)
√ [adːa] /adːa/		*→		
☞ [atːa] /atːa/				←*

Demoting *adda to a ranking value of 70.18 would mean that it fell below *abba. As a consequence, *abba will also be demoted to 70.18.

The family reranking advocated here reflects some common properties of motor learning or even simple acoustics. For instance, if running at least 20 km/h will cost me an effort *E*, running at least 10 km/h will never cost me more than *E*, even if the effort *E* of running at least 20 km/h decreases as a result of training at that high speed. In this example, there is even a logical implication, comparable to the fact that if my glottis and pharynx configuration is such that a coronal geminate stop surfaces as voiced, a labial geminate stop will certainly surface as voiced, too, with the same glottis and pharynx configuration.

A prediction of this is that if a language has the stop system { p, t, k, d, g }, the lack of |b| is an accidental gap, and speakers will have no trouble faithfully borrowing words with /b/ from other languages.

7 The acquisition of arbitrary constraints and their rankings

The theory of functional phonology maintains that non-arbitrary phonology is "phonetic", i.e., that theories of autosegmental phonology and feature geometry should be replaced with a theory that expresses functional principles directly as articulatory and faithfulness constraints in a production grammar and categorization constraints in a perception grammar. Thus, all universal phonology is phonetic.

But not all phonology is phonetic. Many morphological alternations must be regarded as arbitrary in any theory of phonology, whether nativist or empiricist. For instance, the /k/ – /s/ alternation in /ɪlˈɛktrɪk/ 'electric' – /ɪlɛktrˈɪsɪtɪ/ 'electricity' does not appeal to any articulatory or perceptual relationship between the sounds, unless seen from an abstract analysis that invokes all the techniques of internal reconstruction. According to the empiricist view, these processes cannot be universal, and all of them must be learned. Still, even arbitrary morphological alternations often refer to natural classes of segments and features. In order to account for this, we must assume that learners will hypothesize generalizations across groups of segments or features related by articulatory or perceptual similarities. This is a natural phenomenon, not restricted to speech.

I will show how the learning of partly arbitrary morphology with phonological generalizations takes place for the Gradual Learning Algorithm. Our example will be the formation of the past tense in English, a subject that has received a lot of attention in observational research and the modelling of acquisition (Berko 1958; Ervin 1964; Brown 1973; Kuczaj 1977, 1978; Bybee & Slobin 1982; Rumelhart & McClelland 1986; Pinker & Prince 1988).

When confronted with the pair /sɪŋ/ ~ /sæŋ/ ‘sing - sang’, the learner may produce many generalizations on the basis on the phonological similarities and differences between the two:

(43) *Constraints generated by* /sɪŋ/ - /sæŋ/:

a. ‘sing’-sɪŋ ~ sæŋ
‘sing’ /sɪŋ/ + PAST → /sæŋ/ (or ‘sang’ /sæŋ/ + PRESENT → /sɪŋ/)
A non-phonological constraint: lexical suppletion.

b. ɪ ~ æ / _ ŋ]$_\sigma$
Endings in /ɪŋ/ correspond to endings in /æŋ/.
Is also satisfied by /ɹɪŋ/ ~ /ɹæŋ/ ‘ring - rang’.

c. ɪ ~ æ / _ ŋ
Verbs with /ɪŋ/ correspond to verbs with /æŋ/.
Is also satisfied by /dɹɪŋk/ ~ /dɹæŋk/ ‘drink - drank’.

d. ɪ ~ æ / _ nas
The correspondence works before any nasal. This generalization is possible if [nasal] is a perceptual category for the learner.
Is also satisfied by /swɪm/ ~ /swæm/ ‘swim - swam’.

e. ɪ ~ æ / _ vel
The correspondence works before any velar. This generalization is possible if [velar place] is a perceptual category for the learner.
Is satisfied by those verbs that already satisfy the more specific (43c).

f. ɪ ~ æ
The correspondence works for all verbs with /ɪ/.
Is also satisfied by /sɪt/ ~ /sæt/ ‘sit - sat’.

g. low F1 ~ high F1
Is also satisfied by /fiːd/ ~ /fɛd/ ‘feed’ and /tʃuːz/ ~ /tʃouz/ ‘choose’.

h. high F2 ~ low F2
Is satisfied by /weik/ ~ /wouk/ ‘wake’, and by more than 85 percent of all other English strong verbs with a vowel alternation.

i. And so forth.

Once a constraint is in the grammar, it will not be re-produced, but it may be reranked as the result of a mismatch between the learner’s own form and the adult’s. For instance, suppose that eight constraints are ranked as in (44) and the learner hears /bɹɔːt/ as a past tense of ‘bring’, knowing that its present tense is /bɹɪŋ/. With the grammar in (44), she will produce /bɹæŋ/ herself. The learner will note that this form differs from the adult’s, and she will consequently lower the ranking values of all constraints violated in /bɹɔːt/, and raise those of all constraints violated in /bɹæŋ/:

(44) *Learning* /bɹɔːt/ ‘brought’

|‘bring’-bɹɪŋ + PAST| Adult: /bɹɔːt/	ɪ ~ æ /_ŋ	ɪ ~ æ	lower F2	ɪ ~ ʌ	ɪ ~ æ /_vel	‘bring’-bɹɪŋ ~ bɹɔːt	ɪ ~ æ /_nas	V ~ V+(ɪ)d
√ bɹɔːt	*!→	*→	*→	*			*→	*
☞ bɹæŋ				*		←*		*
bɹɪŋd	*!	*	*	*	*	*	*	
bɹʌŋ	*!	*	*		*	*	*	*

This action will gradually increase the probability of the learner producing /bɹɔːt/.

We can see how this learning process contributes to generalization. Suppose that the rather specific constraints ɪ ~ æ / _ ŋ]$_\sigma$ and ɪ ~ æ / _ ŋk are ranked a bit above the more general constraint ɪ ~ æ / _ ŋ , and that the learner makes mistakes for the past tenses of ‘sing’-/sɪŋ/ and ‘drink’-/dɹɪŋk/, perhaps by a high ranking of the constraint for regular verbs V ~ V+(ɪ)d , so that the forms end up as /sɪŋd/ ‘sang’ and /dɹɪŋkt/ ‘drank’. In the case of /sɪŋd/, the constraints ɪ ~ æ / _ ŋ]$_\sigma$ and ɪ ~ æ / _ ŋ will rise, and in the case of /dɹɪŋkt/, the constraints ɪ ~ æ / _ ŋk and ɪ ~ æ / _ ŋ will rise. We see that the more general constraint rises twice as often as each of the more specific constraints, so that there is a fair chance that although it started out lower, it will rise above the other two, thus making them superfluous with respect to the determination of future winning candidates.

Restricting ourselves to verbs that contain /ɪŋ/, we see that there are several possible final adult grammars:

(45) *Adult* /ɪŋ/*-verb typology*

a. ‘sink’-sɪŋk ~ sæŋk >> V ~ V+(ɪ)d
Strong /ɪŋ/ verbs like ‘sink’-/sɪŋk/ are lexical exceptions to the general *-ed* pattern for regular verbs. We can detect this situation by asking the adult to produce the past tense of the nonsense verb /plɪŋ/. She will answer /plɪŋd/.

b. ‘wink’-wɪŋk ~ wɪŋkt >> ɪ ~ æ >> V ~ V+(ɪ)d
Regular /ɪŋ/ verbs like ‘wink’-/wɪŋk/ are lexical exceptions to the more general ɪ ~ æ pattern for /ɪŋ/ verbs. When asking the adult to produce the past tense of /plɪŋ/, she will say /plæŋ/.

c. ‘wink’-wɪŋk ~ wɪŋkt >> ‘sink’-sɪŋk ~ sæŋk >> ɪ ~ ʌ >> V ~ V+(ɪ)d
If the ɪ ~ ʌ constraint is high (on the basis of ‘sting’-stɪŋ ~ stʌŋ), both the regular ‘wink’-wɪŋk and the strong ‘sink’-sɪŋk will be lexical exceptions. As a past tense of /plɪŋ/, the adult will propose /plʌŋ/.

Which of these three possibilities will emerge for any particular learner, is a question of the initial ranking of the constraints, the number of lexicalized forms at the time the learner starts to produce generalizations, and the order and relative frequency in which the data are presented to her.

A special case is the form /wɛnt/ 'went', which for adults is the past tense of 'go'-/gou/. Since the forms are hard to relate phonologically, it may well be that for a beginner in English, the form /wɛnt/ is just the past tense of the *concept* 'go', instead of the past tense of the *sign* 'go'-/gou/. So we must expect that /wɛnt/ and /goud/ can co-exist with a learner, even if the relevant constraints are at very different heights.

The simulation explains:

(46) *Phenomena accounted for*:
  a. The usual ability of English speakers to tell that the past tense of *wug* is *wugged*.
  b. The variation attested with English speakers in the production of the past of /plɪŋ/ as /plɪŋd/, /plæŋ/, or /plʌŋ/.
  c. The overgeneralization of English-learning children in the forms *bringed* and *brang*.
  d. The 'errors' of saying *stang* instead of *stung*, or *sung* instead of *sang*.
  e. The common variation in the past tenses of *sink* (*sank* or *sunk*) and *stink*.

A question that remains is whether there is any evidence for how arbitrary constraints enter the grammar. While this was clear in the functional case (faithfulness from the bottom, articulatory constraints from the top), it is harder to find a universal criterion for arbitrary constraints. One would think that more general constraints should start out with a higher ranking than more specific constraints, so that if a more general constraint holds, it will continue determining the outcome, and the grammar would be minimized in the sense that specific constraints play no role if their case can be handled by more general constraints. Some more analysis of the course of phonological development is needed to find out whether this holds.

8 Conclusion

General innate properties of perception, motor behaviour, cognition, and constraint-based learning, together with a human-specific capacity to use arbitrary signs, are enough to explain functional and arbitrary phonology and its development. Any innateness of phonological features and their values is superfluous. Indeed, the arbitrariness expected if they were innate would hamper the local optimization of the efficiency of human communication.

REFERENCES

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded phonology*. Cambridge: MIT Press.
Bergem, Dick R. van (1993). Acoustic vowel reduction as a function of sentence accent, word stress and word class. *Speech Communication* **12**. 1–23.
Berko, Jean (1958). The child's learning of English morphology. *Word* **14**. 150–177.
Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences* **21**, University of Amsterdam. 43–58. Web-readable at http://fonsg3.hum.uva.nl. [equals ch. 15 of Boersma 1998c]

Boersma, Paul (1998a). Spreading in Functional Phonology. *Proceedings of the Institute of Phonetic Sciences* **22**, University of Amsterdam. xx-xx [ROA **280**; equals ch. 19 of Boersma 1998c]

Boersma, Paul (1998b). The OCP in Functional Phonology. *Proceedings of the Institute of Phonetic Sciences* **22**, University of Amsterdam. xx-xx [ROA **283**; equals ch. 18 of Boersma 1998c]

Boersma, Paul (1998c). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. PhD dissertation, University of Amsterdam. LOT International Series **11**. The Hague: Holland Academic Graphics. [http://fonsg3.hum.uva.nl/paul/diss/]

Boersma, Paul (to appear a). Learning a grammar in Functional Phonology. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.) *Optimality Theory: Phonology, syntax, and acquisition*. Oxford University Press. [equals most of ch. 14 of Boersma 1998c]

Boersma, Paul (to appear b). *Nasal harmony in functional phonology*. Ms. University of Amsterdam. Presented at HILP 4, January 1999, Leiden.

Boersma, Paul (to appear c). Acquisition in functional phonology. In R. Kirchner, D. Steriade & B. Hayes (eds.) *Phonetically-driven phonology*.

Boersma, Paul (to appear d). In defense of a grammar of performance.

Boersma, Paul, Joost Dekkers & Jeroen van de Weijer (to appear). An introduction to Optimality Theory: Phonology, syntax, and acquisition. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.) *Optimality Theory: Phonology, syntax, and acquisition*. Oxford University Press.

Boersma, Paul & Bruce Hayes (to appear). Empirical tests of the Gradual Learning Algorithm.

Brown, Roger (1973). *A first language*. Cambridge: Harvard University Press.

Bybee, Joan L. & Dan I. Slobin (1982). Rules and schemas in the development and use of the English past tense. *Lg* **58**. 265–289.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

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

Eimas, Peter D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics* **16**. 513–521.

Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk & James Vigorito (1971). Speech perception in infants. *Science* **171**. 303–306.

Ervin, Susan (1964). Imitation and structural change in chidlren’s language. In E. Lenneberg (ed.) *New directions in the study of language*. Cambridge: MIT Press.

Flemming, Edward (1995). *Auditory representations in phonology*. PhD dissertation, University of California, Los Angeles.

Fletcher, Harvey & W.A. Munson (1933). Loudness, its definition, measurement, and calculation. *JASA* **5**. 82–108.

Gibson, Edward & Kenneth Wexler (1994). Triggers. *LI* **25**. 407–454.

Gnanadesikan, Amalia (1995). *Markedness and faithfulness constraints in child phonology*. Ms, University of Massachusetts. [Rutgers Optimality Archive **67**]

Guerssel, Mohand & Jean Lowenstamm (1996). Ablaut in Classical Arabic Measure I active verbal forms. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.), *Studies in Afroasiatic grammar*. The Hague: Holland Academic Graphics. 123–134.

Hale, Mark & Charles Reiss (1998). Formal and empirical arguments concerning phonological acquisition. *LI* **29**. 656–683.

Hayes, Bruce (to appear). Phonetically driven phonology: the role of Optimality Theory and inductive grounding. *Proceedings of the 1996 Milwaukee Conference on Formalism and Functionalism in Linguistics*. [ROA-158]

Hernández-Chávez, Eduardo, Irene Vogel & Harold Clumeck (1975). Rules, constraints and the simplicity criterion: An analysis based on the acquisition of nasals in Chicano Spanish. In Charles A. Ferguson, Larry M. Hyman & John J. Ohala (eds.) *Nasálfest*. Stanford University. 231–248.

Hirt, Hermann (1939). *Die Hauptprobleme der indogermanischen Sprachwissenschaft*. Edited by Helmut Arntz. Halle/Saale: Max Niemeyer.

Hurford (1989). Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua* **77:2**. 187–222.

Jun, Jongho (1995). Place assimilation as the result of conflicting perceptual and articulatory constraints. *WCCFL* **14**. 221–237.

Jusczyk, Peter W., Derek Houston & Mara Goodman (1998). Speech perception during the first year. In Alan Slater (ed.) *Perceptual development: Visual, auditory, and speech perception in infancy*. Hove: Psychology Press. 357–388.

Kuczaj, Stan A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* **16**. 589–600.

Kuczaj, Stan A. (1978). Children’s judgments of grammatical and ungrammatical irregular past tense verbs. *Child Development* **49**. 319–326.

Kuhl, Patricia K. & J.D. Miller (1978). Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America* **63**: 905–17.

Kuhl, Patricia K. & D.M. Padden (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics* **32**. 542–550.

Lehmann, Winfred P. (1993): *Theoretical bases of Indo-European linguistics*. London: Routledge.

Leoni, F.A., F. Cutugno & R. Savy (1995). The vowel system of Italian connected speech. *Proceedings of the 13th International Congress of Phonetic Sciences*. Vol. **4**. 396–399.

McCarthy, John J. & Prince, Alan (1993). *Prosodic Morphology I: Constraint interaction and satisfaction*. Ms. University of Massachesetts, Amherst, and Rutgers University, New Brunswick.

Mohanan, K.P. (1993). Fields of attraction in phonology. In John A. Goldsmith (ed.) *The last phonological rule: Reflections on constraints and derivations*. University of Chicago Press. 24–69.

Myers, Scott (1997). Expressing phonetic naturalness in phonology. In Iggy Roca (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 125–152.

Padgett, Jaye (1995). Partial class behavior and nasal place assimilation. *Proceedings of the Arizona Phonology Conference: Workshop on Features in Optimality Theory*. Coyote Working Papers, Univ. of Arizona, Tucson. [Rutgers Optimality Archive **113**, http://ruccs.rutgers.edu/roa.html]

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

Pater, Joe (1996). Austronesian nasal substitution and other NC̥ effects. *Proceedings of the Utrecht Prosodic Morphology Workshop*.

Pinker, Steven (1994). *The language instinct*. London: Penguin.

Pinker, Steven & Alan Prince (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* **28**. 73–193.

Powers, William T. (1973). *Behavior: The control of perception*. Chicago: Aldine.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar*. Rutgers University Center for Cognitive Science Technical Report **2**.

Pulleyblank, Douglas (1993). Vowel harmony and Optimality Theory. *Actas do Workshop sobre Fonologia*. University of Coimbra. 1–18.

Pulleyblank, Douglas & William Turkel (1996). Optimality Theory and learning algorithms: the representation of recurrent featural asymmetries. In J. Durand & B. Laks (eds.) *Current trends in phonology: Models and methods*. University of Salford.

Recasens, Daniel (1991). *Fonètica descriptiva del català* [*Biblioteca Filològica* **XXI**]. Barcelona: Institut d'Estudis Catalans.

Rumelhart, David E. & James L. McClelland (1986). On learning the past tense of English verbs. In J.L. McClelland, D.E. Rumelhart and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol 2: *Psychological and biological models*. Cambridge: MIT Press. 216–271.

Saussure, Ferdinand de (1916). *Cours de linguistique générale*. Edited by Charles Bally & Albert Sechehaye in collaboration with Albert Riedlinger. Paris: Payot & Cie. [2nd edition, 1922]

Sluijter, Agaath (1995). *Phonetic correlates of stress and accent*. PhD dissertation, University of Leiden.

Smith, Neilson V. (1973). *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.

Steriade, Donca (1995): *Positional neutralization*. Two chapters of an unfinished manuscript, Department of Linguistics, UCLA.

Stevens, Kenneth N. (1971). Airflow and turbulence noise for fricative and stop consonants: static considerations. *JASA* **50**. 1180–1192.

Swoboda, P., P.A. Morse & L.A. Leavitt (1976). Continuous vowel discrimination in normal and at-risk infants. *Child Development* **47**. 459–465.

Tesar, Bruce & Paul Smolensky (1998). Learnability in Optimality Theory. *LI* **29**. 229–268.