Nasal harmony in functional phonology

Paul Boersma, University of Amsterdam, http://www.fon.hum.uva.nl/paul/

draft, January 1999

1 Three grammar models

Twentieth-century phonology has brought us two main competing grammar models. The first could be called ***structuralist*** and has a discrete phonological level intermediate between two sequentially ordered modules, which can be associated with ***phonology*** and ***phonetics***:

(1) *Structuralist model of phonology*

$|$underlying$|$ → /phonemic/ → [phonetic]

The second grammar model could be called ***generative*** and collapses the phonological and phonetic modules into one, rejecting the discrete phonological surface level:

(2) *Generative model of phonology*

$|$underlying$|$ → [phonetic]

But I will argue that these two models actually share the major assumption of hybrid cognitive phonological representations. A third grammar model, which can be called ***functional***, does not share this assumption, since it distinguishes between articulatory and perceptual representations. The model has an articulatory phonetic level between two sequentially orderd modules, which can be associated with ***production*** and ***perception***:

(3) *Functional model of phonology (in structuralist and generative terms)*

$|$underlying$|$ → [phonetic] → /phonemic/

I will show that this model only combines the virtues of the structuralist and generative models, and I will claim that it comes nearer to the truth in accounting for phonological phenomena.

I chose the terminology in (3) for maximal comparability with the structuralist and generative models. However, there exist more appropriate terms that make explicit the distinction between articulation and perception:

(4) *Functional model of the production grammar (in functional terms)*

$|$perceptual specification$|$ → [articulatory implementation] → /perceptual output/

1.1 Modular grammars include a discrete surface representation

1.1.1 Structuralist modularity

The later structuralists (Bloomfield 1933, Hockett 1965) maintained an intermediate ***autonomous phonemic*** level, which contained all contrastive surface elements (***phonemes***), and no non-distinctive variants (***allophones***):

(5) *Structuralist grammar model*

|morphophonemic|
↓ phonology
/autonomous phonemic/
↓ phonetic implementation
[phonetic]
→ sound

One of the main arguments for the phonemic form was that there had to be a surface representation relating to the perception of “same” and “different” (Bloomfield 1933: 128; Hockett 1965: 194).

1.1.2 An intuitive grammar model

This is perhaps supported by a majority nowadays. E.g. Gussenhoven (1998), despite noting that phonological and phonetic constraints *can* appear in the same OT grammar, proposes:

(6) *Mainstream stratification*

|output of the lexicon| (underlying form)
↓ postlexical phonology
/discrete phonological surface structure/
↓ phonetic implementation
[continuous articulatory utterance]
→ sound

The main arguments for discrete phonological surface structure are: feeling, belief, intuition.

Stratifications like (6), sometimes with a different division of labour, are also often based on a kind of opacity that used to be expressed by counterbleeding rule ordering, i.e. Ernestus (this conference): word-final neutralization of obstruent voice in Dutch is phonological (probably lexical), whereas optional intervocalic obstruent voicing is phonetic.

1.1.3 Hale & Reiss (1998): no grammar of performance

Hale & Reiss (1998) make a point of distinguishing between a phonological grammar, which is used for production as well as comprehension, and a performance system, which is sensitive to external influences, fatigue, drunkenness, or local anaesthesia:

(7) *Grammar and body according to Hale & Reiss*

$$\begin{tikzcd}
\text{|underlying representation|} \arrow[d, "\text{phonology}"] \\
\text{/output of grammar/} \arrow[d, "\text{performance system and body}"] \\
\text{[output of body]} \\
\rightarrow \text{sound}
\end{tikzcd}$$

Their main argument is that the grammar cannot control the output of the body, because grammar is about mental states. In the words of Reiss (Optimality List, April 27, 1998), “If I don’t start flossing, all my teeth may fall out — my pronunciation will change, but my phonology won’t.” The following is how a sudden loss of teeth would be accounted for in Hale & Reiss’s model:

(8) *A sudden teeth loss according to Hale & Reiss*

$$\begin{tikzcd}
\textit{Before loss of teeth:} & \textit{After loss of teeth:} \\
\text{|s|} \arrow[d, "\text{grammar}"] & \text{|s|} \arrow[d, "\text{grammar}"] \\
\text{/s/} \arrow[d, "\text{body}"] & \text{/s/} \arrow[d, "\text{body}"] \\
\text{[s]} & [\theta] \\
\rightarrow \text{sound} & \rightarrow \text{sound}
\end{tikzcd}$$

Hale & Reiss note that a monostratal model would show a sudden grammar change:

(9) *Hale & Reiss’s paradox*

$$\begin{tikzcd}
\textit{Before loss of teeth:} & \textit{After loss of teeth:} \\
\text{|s|} \arrow[d, "\text{grammar}"] & \text{|s|} \arrow[d, "\text{grammar}"] \\
\text{[s]} & [\theta] \\
\rightarrow \text{sound} & \rightarrow \text{sound}
\end{tikzcd}$$

Hale & Reiss argue that this state of affairs is incompatible with the idea that the grammar is about mental states, which should not change suddenly as a result of a performance problem.

1.2 Monostratal grammars combine postlexical phonology and phonetics

There are at least three reasons to handle the modules of postlexical phonology and phonetic implementation in a single module, since three criteria for dividing them can be shown to fail.

1.2.1 Optionality and pragmatical conditioning

Phonetic implementation rules tend to be optional and pragmatically conditioned. But the same goes for many postlexical phonological rules:

(10) *“Phonemic” optionality and conditioning*
  a. Normal Dutch: |aːn+pɑsə| → [aːmpɑsə] (80%) or [aːnpɑsə] (20%).
  b. Clear Dutch: |aːn+pɑsə| → [aːmpɑsə] (20%) or [aːnpɑsə] (80%).

1.2.2 Language-dependency

Phonology must be a part of the grammar, because it is language-dependent. But phonetic implementation is often language-dependent, too:

(11) *“Allophonic” variation specific to Dutch*
  a. Mid back vowel is high before nasals: [ɦʊndəɦɔk] ‘dog kennel’.
  b. Mid back vowel is low before non-nasals: same example.

which means that it must constitute a part of the grammar.

1.2.3 Halle’s argument

Some processes do not fit in the artificial division between postlexical phonology and phonetic implementation. The original argument is due to Halle (1959), and I will apply it to an example in Dutch. This language has a lexical voicing contrast between /t/ and /d/, but not between /k/ and /g/, since /g/ is not in the set of underlying segments in Dutch. Now, Dutch has a rule of voicing assimilation, which causes |wɪt+buk| ‘white + book’ to be pronounced as [wɪdbuk], and |zɑk+duk| ‘pocket + towel = handkerchief’ to be pronounced as [zɑgduk]. Since the autonomous phonemic representation was supposed to show only the phonemic oppositions, the two utterances would be represented at this intermediate level as /wɪdbuk/ (showing the phonemic voicing) and /zɑkduk/ (not showing the allophonic voicing):

(12) *Halle’s argument applied to Dutch*

morphophonemic	autonomous phonemic	phonetic
wɪt + buk	wɪdbuk	wɪdbuk
zɑk + duk	zɑkduk	zɑgduk

Thus, an obviously single voicing rule shows up in the phonological component (creating /wɪdbuk/) and in the phonetic component (creating [zɑgduk]). This artificial separation of the duties of the two components led the early generativists (Halle 1959: 22, Chomsky 1964: 88, Postal 1968: 3–314) to reject the autonomous level of representation, thus leaving only two levels:

(13) *Generative monostratal phonology*

|systematic phonemic|

↓ phonology & phonetic implementation

[phonetic]
→ sound

Thus, Chomsky & Halle (1968), long considered the bible of generative phonology, presented a single stratum of sequentially ordered phonological rules leading from a very abstract underlying representation of English lexical forms to a very concrete surface representation in terms of universal phonetic feature values. The stratum includes, for instance, the nasalization of vowels before nasal consonants, regardless of whether the dialect in question makes any distinction in that respect between the vowels in English *freeness* and *keenness* or *see now* and *seen now*, which would otherwise be a criterion for the inclusion of nasalization in the phonological or phonetic module.

1.3 Functional grammars distinguish articulation and perception

Since the arguments of §1.2 render every proposed separation between postlexical phonology and phonetic implementation arbitrary, i.e. based on an arbitrary definition by the linguist, they seem to belong together in a single module of the grammar. So the intermediate level must be dismissed, as the early generativists did.

But this does not mean that the arguments of §1.1 must be rejected as well! The paradox is the result of a fallacy shared by the structuralists, the early generativists, and the people who intuit a discrete phonological surface structure:

(14) *Assumption shared by supporters of phonemics and opponents of modularity:*
If there is a discrete phonologial surface structure,
then it must be located between the underlying form and the phonetic surface form.

But the real arguments in §1.1 and §1.2 were not logically contradictory:

(15) *Desired properties of a grammar model*
a. There should be a discrete phonological surface structure.
b. There should be no intermediate level between phonology and phonetics.

The problem of regarding (15a) and (15b) as contradictory, is caused by a failure to distinguish between articulatory and perceptual representations in phonology.

In a functional theory of phonology, the paradox vanishes. The perceptual surface representation must necessarily be derived from the articulatory phonetic form and its automatic acoustic result:

(16) *Modularity in functional phonology*

|output of the lexicon| (= perceptual specification = underlying form)

↓ postlexical phonology & phonetic implementation

[continuous articulatory utterance]
→ sound

↓ perception grammar

/discrete phonological surface structure/ (= perceptual output)

The hypothesis is that postlexical phonology and phonetic implementation are evaluated in parallel within a single Optimality-Theoretic production grammar. But we do have a potentially discrete surface representation, namely the perceptual output, though it is not intermediate between the underlying form and articulatory phonetic form.

Reiss's argument, which would have serious consequences for the description of phonological acquisition if true, vanishes:

(17) *Functional view of teeth loss*

Before loss of teeth:	*After loss of teeth:*	*After relearning:*
|s|	|s|	|s|
↓ production	↓ production	↓ production
[groove]	[groove]	[lateral groove]
↓ perception	↓ perception	↓ perception
/s/	/θ/	/s/
→ sound	→ sound	→ sound

We see that immediately after teeth loss, the grammar produces the same articulatory candidate and the mental state has not changed. As a long-term strategy, the speaker may try to compensate her loss and search for a more complicated gesture that will faithfully implement the sibilancy of the specified |s|. Only after a more gradual relearning phase will she master a new lateral grooving gesture, as shown on the right-hand side of (17).

The following two observations summarize the main properties of the production grammar according to a functional model of phonology:

(18) *Production grammar*

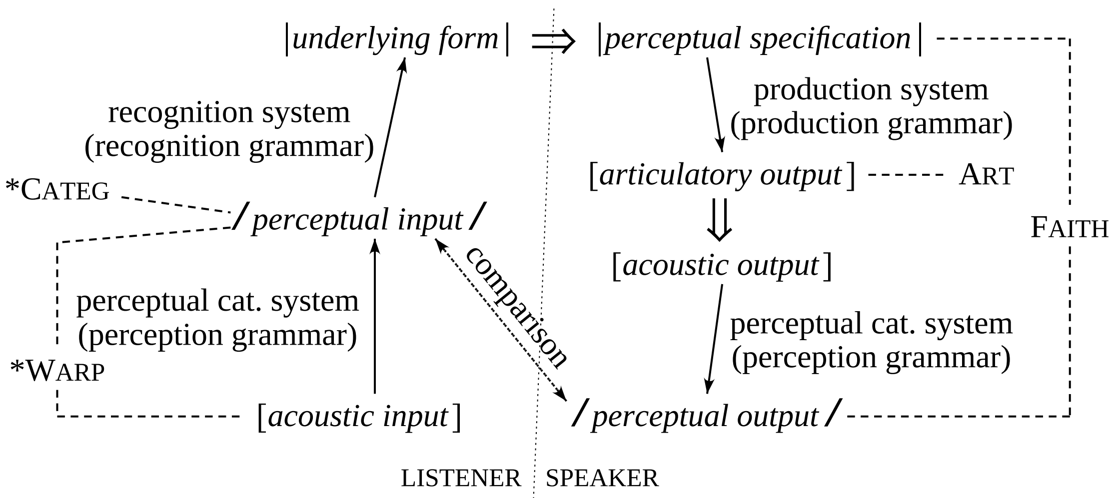
a. The grammar determines the articulatory output.

b. Production is ultimately perception-oriented, like all human behaviour (Powers 1973).

2 The complete grammar model

Using terminology more appropriate for discussing speech production and perception, the complete grammar model as defended in Boersma (1998) is:

(19) *Grammar model of functional phonology*



A module new to phonological theory is the perception grammar. It turns raw physical acoustic forms, which are the automatic results of the articulations of the speaker or others, into more discrete perceptual representations.

2.1 The production grammar and its local rankings

The production grammar is modelled as an Optimality-Theoretic grammar consisting of acquired articulatory and faithfulness constraints. Its input is the perceptual specification; its candidates are articulatory implementations paired with perceptual results:

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

|*spec*|	A	B
☞ [art_1] /$perc_1$/		*
[art_2] /$perc_2$/	*!	

Note the intertwining of the production and perception grammar: the faithfulness constraints, which compare the output of the perception grammar with the underlying specification, are elements of the production grammar.

The production grammar contains a number of ***articulatory constraints*** (depicted as ART in the figure), which evaluate aspects of the innate functional principle of ***minimization of effort***, and work directly on each articulatory output candidate (Boersma 1998: ch. 7). Such a constraint enters the production grammar from above as soon as the learner discovers the relation between an articulatory gesture and its perceptual result. The most typical example is:

(21) *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 or sequential gestures.

The ***local-ranking principle*** (Boersma 1998) restricts the typologically possible languages by assuming that pairs of constraints can be ranked in a universal manner if they differ in a single argument or condition, and that they can be ranked in a language-specific manner otherwise. Thus, articulatory constraints can be locally ranked according to articulatory effort, e.g. (21) is ranked higher if the *distance*, *duration*, *precision*, or *velocity* is greater, and everything else stays equal. Otherwise, the rankings are largely language-specific: a *global* measure of articulatory effort (e.g. Boersma 1998: eq. 7.4) can only account for cross-linguistic statistical tendencies.

The production grammar also contains a number of ***faithfulness constraints*** (depicted as FAITH in the figure), which evaluate aspects of the innate functional principle of ***minimization of confusion*** indirectly in an evaluation of aspects of the similarity between the perceptual result of each candidate and the underlying perceptual specification (Boersma 1998: ch. 9). Such a constraint enters the production grammar from below as soon as the learner's perception grammar has supplied her with a perceptual category. The most typical example is:

(22) *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.

Faithfulness constraints can be locally ranked according to perceptual confusion, e.g. (22) is ranked higher if *value*$_1$ and *value*$_2$ are further apart or if the condition or the environment contribute to a smaller amount of confusion, and everything else stays equal. Otherwise, the rankings are largely language-specific: a *global* measures of perceptual confusion (e.g. Boersma 1998: eq. 4.24) can only account for cross-linguistic statistical tendencies.

As an example, consider the case of nasal place assimilation, in which any nasal consonant assimilates to any following consonant:

(23) *Nasal place assimilation*

|an+pa|	*REPLACE (place / _ V)	*REPLACE (pl / plosive / _ C)	*GESTURE	*REPLACE (pl / nasal / _ C)
[anpa] /anpa/			*!	
☞ [ampa] /ampa/				*
[anta] /anta/	*!			

At the same time, plosives do not assimilate:

(24) *Plosives immune to nasal place assimilation*

|at+ma|	*REPLACE (place / _ V)	*REPLACE (pl / plosive / _ C)	*GESTURE	*REPLACE (pl / nasal / _ C)
☞ [atma] /atma/			*	
[apma] /apma/		*!		
[atna] /atna/	*!			*

A functional theory of phonology cannot accept the usual practice (e.g. Padgett 1995) of attributing this phenomenon directly to a nasality-specific structural constraint like NASSIM that outranks a general faithfulness constraint like PARSE (place). The structural (gestural) constraint that is satisfied by place assimilation must be ranked equally high for nasals and plosives because the effort of the tongue-blade gesture in [anpa] and [atma] does not depend on whether the velum is high or low. Instead, the difference between nasals and plosives must be in the ranking of the faithfulness constraints: since the plosives /p/ and /t/ are easier to distinguish than the nasals /m/ and /n/, the constraint against implementing |t| as something perceived as /p/ must be universally ranked higher than the constraint against implementing |n| as something perceived as /m/.

2.2 The perception grammar and its local and global rankings

In (19), the perception grammar performs four functions:

(25) *Functions of the perception grammar*

a. To produce a representation from which we can evaluate *faithfulness*.
b. To produce a representation that we can compare with our perception of the utterances of others, in order to *learn* to speak in the same way as others do.
c. To produce a perceptual representation (***perceptual input***) of the acoustic form of the speech spoken by others (***acoustic input***), in order to compare this with our own speech as we perceive it (perceptual output), in order to learn to speak in the same way as others do.
d. To produce a perceptual representation of the speech of another person, as an input to the ***recognition system*** that will ultimately lead to comprehension.

The perception grammar is implemented as an Optimality-Theoretic grammar consisting of acquired categorization constraints. Its input is a continuous acoustic signal; its candidates are discrete perceptual representations:

(26) *Evaluation of perceptual candidates*

[acoustics]	A	B
☞ /$perc_1$/		*
/$perc_2$/	*!	

The perception grammar contains a number of constraints that evaluate the similarity between the acoustic form and the perceptual form (correspondence, faithfulness), or limit the number of categories. A typical example is:

(27) *WARP (*feature*: *ac, perc* / *condition*):
Do not perceive an acoustic value (*ac*) on a perceptual tier (*feature*) as a different value (*perc*), under a certain *condition*.

This constraint can be locally ranked according to similarity: it is ranked higher if the perceptual distance between *ac* and *perc* is greater. The least violated *WARP constraint thus usually determines the resulting perceptual category.

Other constraints in the perception grammar control the abstraction of simultaneous and sequential co-occurrence. For instance, [m] may be perceived as the single percept [labial nasal], not just as the separate [labial] and [nasal] or the path [labial & nasal], if it is a frequent articulation in the language. Likewise, a pair of constraint families in the perception grammar together determine the abstraction of sequential acoustic cues into a single percept:

(28) OBLIGATORYCONTOURPRINCIPLE (*f*: *x*; cue_1 | *m* | cue_2)
A sequence of acoustic cues cue_1 and cue_2 with intervening material *m* is heard as a single value *x* on the perceptual tier *f*.

(29) LINECROSSINGCONSTRAINT (*f*: *x*; cue_1 | *m* | cue_2)
A sequence of acoustic cues cue_1 and cue_2 with intervening material *m* is not heard as a single value *x* on the perceptual tier *f*.

I'll abbreviate these constraints as OCP and LCC. Both originated in generative phonology as inviolable constraints on representations. In functional phonology, they are violable constraints on the perceptual representation that is derived from an acoustic signal by the perception grammar. The existence of these constraints is a result of general properties of human perception: if we see an object partly obscured by other objects, we can still sometimes perceive the various visible parts together as a single object.

Some universal rankings of OCP and LCC can be derived from the local-ranking principle:

(30) *Local rankings of OCP*
a. Higher if the sequential combination of cue_1 and cue_2 is more common.
b. Lower if there is more intervening material.

(31) *Local rankings of LCC*
a. Lower if the sequential combination of cue_1 and cue_2 is more common.
b. Higher if there is more intervening material.

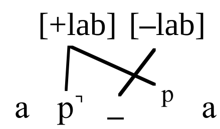
For instance, an underlying |apa|, which means "low vowel followed by labial plosive followed by low vowel", will be implemented in most languages by the articulation [apa], which means "low tongue, open jaw, adducted vocal folds, contracting lungs, and a lip closing & opening gesture". The acoustic result will be [[a p˺ _ ᵖ a]], which means "high F1, labial closure transition, silence, labial release burst, high F1". Now suppose the labial closure transition, the silence, and the labial release burst tend to co-occur very often in this language (which is the usual case in languages where intervocalic single consonants are common). Then the perception grammar will abstract away from this acoustic detail and perceive these three acoustic cues as a single “labial plosive”, which makes the perceptual output faithful to the specification. We can formalize this as:

(32) *A near-universal example of abstraction*

acoustics: [[a p˺ _ ᵖ a]]	OCP (place: labial; transition │ silence │ burst)	LCC (place: labial; transition │ silence │ burst)
/ap˺_ᵖa/ (labial tier: –|+|–|+|–)	*!	
☞ /apa/ (labial tier: –|+|–)		*

To see that the winning candidate violates a line-crossing constraint, consider the perceptual representation, which shows that a non-labial silence intervenes between the two labial cues:

(33) *Universally strong OCP dominates universally weak LCC*

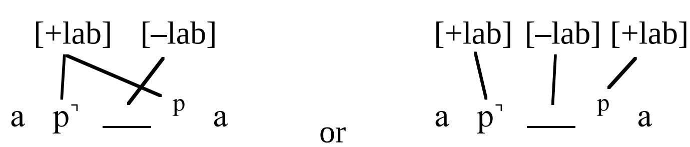


The abstraction further shows in the existence of the production constraints *DELETE (labial), *DELETE (plosive), *DELETE (labial plosive), instead of *DELETE (silence) and *DELETE (labial burst).

The fact that the sequential abstraction of the labial cues in [apa] is a probably universal perception process, shows that the domination of this particlar OCP over this particular LCC is also universal. In a functional theory of phonology, this must be attributed to a large difference between the global ranking measures of these two constraints, since they cannot be related by the local-ranking principle. Languages will choose their own ranking values for these constraints, thus giving a statistical tendency, but in this case the skewing will be so severe that the current number of languages on earth that exhibit perceptual separation of the two labial cues is likely to be zero.

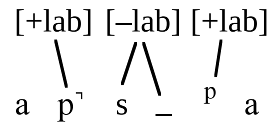
While [ap˺_ᵖa] contains minimal intervening material (just a short silence), sequences with more material between the labial cues will be less likely to be perceived with single labiality. Articulations that produce a longer silence will be heard either as two short labial plosives, or as a single long one, mainly depending on whether the language “has” geminates or not (Boersma 1998: 426):

(34) *Factorial typology of OCP and LCC of universally intermediate ranking*



A comparable typology could be given for articulations like [mp]: they could be perceived as two separate labial segment, or as a single labial NC cluster, mainly depending on whether homorganic NC is a common tautomorphemic cluster in the language (Boersma 1998: 242). An articulation like [apspa], with an intervening sibilant noise (plus a silence), will very likely be perceived with two separate labial values on the place tier, separated by a coronal value, in almost all languages (Boersma 1998: 242):

(35) *Universally strong LCC dominates universally weak OCP*



2.3 The recognition grammar

The ***recognition grammar*** translates other people's speech, as perceived by the listener, into an underlying lexical form expressed in perceptual features. This equals Saussure's (1916) ***acoustic image*** and is connected to meaning in the Saussurean sign. Since the recognition grammar handles lexical access, it can probably be excellently described by means of an Optimality-Theoretic grammar consisting of lexical-access constraints whose rankings depend on the semantic context, and faithfulness constraints. For example, the perceptual form /mɪst/ could be recognized by an English speaker as |mɪst| 'mist' or as |mɪs+d| 'miss+ed', but in the semantic context of 'train', the latter candidate tends to win:

(36) *Recognition of an ambiguous word (disregarding syntactic context)*

/mɪst/	FAITH (nasal)	FAITH (labial)	*LEX (|mɪst| / 'train')	*LEX (|mɪs+d| / 'train')	*LEX (|tɪkɪt| / 'train')	FAITH (voice)
|mɪst| 'mist'			*!			
☞ |mɪs+d| 'missed'				*		*
|tɪkɪt| 'ticket'	*!	*			*	

We see that even if 'ticket' has the closest semantic match (the lowest anti-access constraint), it loses to 'missed' because of the rather poor faithfulness of the image |tɪkɪt| to the perception /mɪst/.

This view of a grammar of comprehension is different from the view shared by the otherwise quite opposing standpoints of Smolensky (1996) and Hale & Reiss (1998), according to which a single grammar should be able to account for production as well as comprehension. As Hale & Reiss note, however, Smolensky's (1996) algorithm for comprehension fails in the case of phonological neutralization: in his model, only the faithfulness constraints (the same as in the production grammar) are relevant, so the listener would always comprehend |mɪst| 'mist', never |mɪs+d| 'missed'. Hale & Reiss's (1998) alternative has already been discussed in §1.1.3.

3 Nasal harmony, type A

Piggott (1992) distinguishes two kinds of nasal-harmony systems, which he calls type A and type B. In type-A nasal harmony, nasality spread from a nasal segment until the spreading is blocked by a segment that is apparently not compatible with nasality. In Malay, for example, nasality may spread rightward though [j] but not through [k]:

(37) *Nasal spreading in Malay (from the initial consonant to the right)*
[mãj̃ãn] ‘stalk’
[mãkan] ‘eat’

Thus, /j/ is a ***target*** for nasalization (it’s ***nasalizable***), whereas /k/ is a ***blocker*** (it’s ***opaque***). The following typology summarizes the possible targets in type-A languages:

(38) *Nasalizable segments (Piggott 1992)*

glides	*liquids*	*fricatives*	*plosives*	*language example*
–	–	–	–	Sundanese
+	–	–	–	Malay, Warao
+	+	–	–	Ijo, Urhobo
+	+	+	–	Applecross Gaelic

This typology corresponds to the following implicational universals:

(39) *Universals of nasal spreading*
a. If glides can be nasalized, so can vowels and laryngeals.
b. If liquids can be nasalized, so can glides.
c. If fricatives can be nasalized, so can liquids.
d. Plosives cannot be nasalized.

3.1 Functional analysis of type-A nasal harmony

Functionally, the generalization is straightforward. Suppose first that the constraint that is honoured by spreading nasality to the right is *MOVE (velum), i.e. a constraint that aims at postponing the raising gesture of the velum, as an indirect way to minimize the number of raising and lowering gestures of the velum. This gestural definition immediately accounts for the Malay type:

(40) *Glides undergo nasal spreading in Malay*

|maja|	*REPLACE (nas: –, + / liquid)	*MOVE	*REPLACE (nas: –, + / glide)
[mãja] /mãja/		*!	
☞ [mãj̃ã] /mãj̃ã/			*

We see that glides undergo nasal spreading, because the gestural constraint outranks the constraint against implementing the segment |j|, which is specified as non-nasal, as something that is perceived as being nasal.

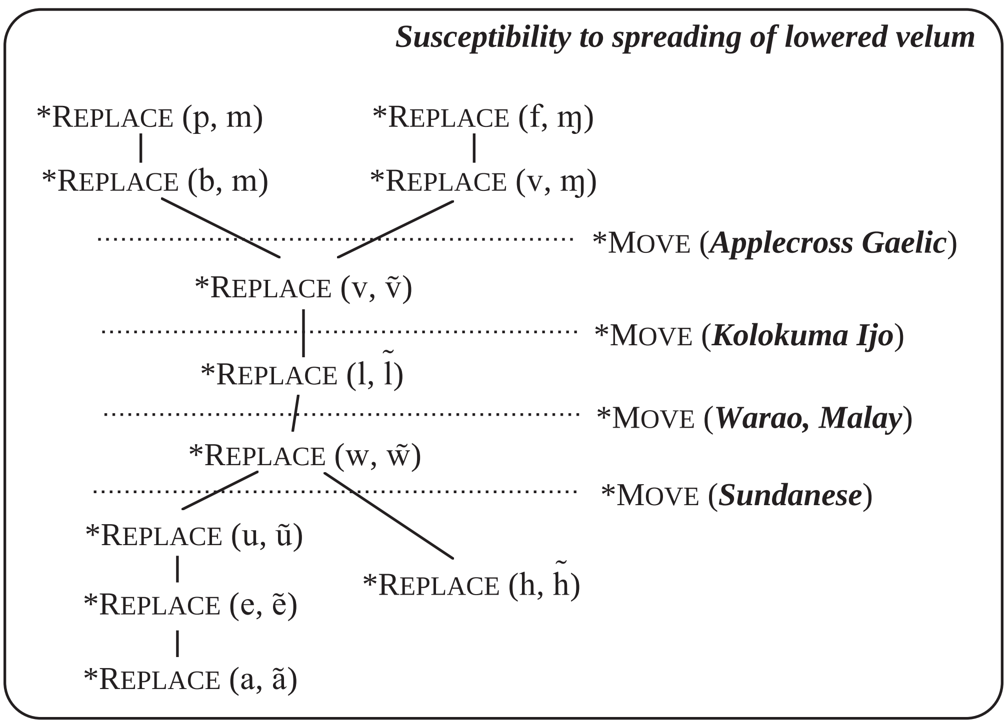
The two *REPLACE constraints in (40) are ranked according to the local-ranking principle, since the closer an oral constriction is, the more its perceptual result is modified by adding a nasal side branch (Cohn 1993). Thus, Malay liquids block nasal spreading:

(41) *Liquids block nasal spreading in Malay*

|mara|	*REPLACE (nas: –, + / liquid)	*MOVE	*REPLACE (nas: –, + / glide)
☞ [mãra] /mãra/		*	
[mãr̃ã] /mãr̃ã/	*!		

The whole universal confusion-based hierarchy of nasalizability is shown as the *REPLACE constraints connected with solid lines in (42):

(42) *Susceptibility to nasal spreading*



The cases of the obstruents need some comment. If the velum is lowered during a gesture that would otherwise produce a labial plosive, a nasal stop will automatically result. So honouring *MOVE would violate some probably highly ranked faithfulness constraints against deletion of the perception of plosiveness and insertion of the perception of sonorancy:

(43) *Plosives block nasal spreading in Malay*

|maka|	*REPLACE (k, ŋ)	*MOVE
☞ [high velum etc.] /mãka/		*
[low velum etc.] /mãŋã/	*!	

The fact that these faithfulness (correspondence) constraints are ranked so high, can be attributed to the strong perceptual repercussions of their violation, according to any reasonable global measure of perceptual distance. For instance, *REPLACE (k, ŋ) can be seen as a shorthand for the conjunction of *DELETE (plosive) & *INSERT (sonorant).[1] For fricatives, a tableau analogous to (43) can be drawn for most languages:

(44) *Fricatives block nasal spreading in most languages*

|masa|	*REPLACE (s, n)	*MOVE
☞ [high velum etc.] /mãsa/		*
[low velum etc.] /mãnã/	*!	

Since we have no way of telling whether the deletion of frication is worse than the deletion of plosiveness, the top pairs of constraints are not universally ranked with respect ot each other. Note that the candidate set in (44) is restricted in such a way that a nasalized fricative cannot occur as a perceptual output. This is because speakers of most languages would not know how to produce such a sound, so that the relevant articulation does not show up as a candidate at all. In Applecross Gaelic, people are reported to be able to produce it, so their tableau is like:

(45) *Nasalized fricatives reported for Applecross Gaelic*

|masa|	*REPLACE (s, n)	*MOVE	*REPLACE (s, s̃)	*GESTURE (special trick)
[high velum etc.] /mãsa/		*!		
[low velum etc.] /mãŋã/	*!			
☞ [special trick] /mãs̃ã/			*	*

So (42) shows a four-way typology, based on the ranking of *MOVE with respect to the fixed hierarchy. The typology also seems to predict languages that show nasalization of plosives and/or fricatives (as nasal stops), though the number of these languages may be very low because of considerations of global measures of distinctivity, which may force the language-specific rankings of *REPLACE (p, m) and *MOVE (velum) to be drawn from distributions (“windows”) that hardly overlap. Nevertheless, final plosives in Sanskrit become nasal when followed by a nasal consonant (though in nasal harmony, the spreading is also from vowels), and the dialect of Bemmel in the Netherlands shows occasional nasalization of plosives across vowels (|ɔ̆k+ӑk+min| ‘even + if I + me’ → [ɔ̆ŋӑŋmin]),[2] though this seems by no means systematic.

As an alternative functional analysis of type-A nasal harmony, we may propose that the nasal spreading is caused by a faithfulness constraint, namely maximizing the duration of the perception of nasality, and that gestural constraints are not involved at all. This will only work if discontinuous sequences such as [mãkã] are ruled out by a strong constraint against the insertion

[1] By virtue of enabling additive ranking, a phonological theory that allows conjunction of faithfulness constraints or of gestural constraints (but no mixes) is more directly functional than a theory without this possibility.
[2] This example occurred to me before I went into linguistics, so I guess it is reliable.

of nasality,[3] which again only works if [mãkã] is perceived with two separate instances of nasality, which is probably the default perception because of the intervening [k], which causes the line-crossing constraint to outrank the OCP for nasality in most languages:

(46) *Faithfulness-only account of Malay*

|maka|	*INSERT (nasal)	*REPLACE (k, ŋ)	MAXDURATION (nasal)	*REPLACE (j, j̃)
[maka] /maka/			***!	
☞ [mãka] /mãka/			**	
[mãŋã] /mãŋã/		*!		
[mãkã] /mãkã/	*!		*	

Whether the account in terms of minimizing velum gestures and the account in terms of maximizing nasality are empirically equivalent, remains to be seen. The conclusion, however, must be that any account of type-A nasal harmony expressed into functionally rankable directly functional constraints, is observationally, descriptively, and explanatorily adequate, since it accounts for the data, predicts the typology, and needs no assumptions and principles except those rooted in general properties of motor behaviour and perception. This means that **phonology is perfect**, in the sense of being indistinguishable from what would be functionally optimal. There could be three explanations for this remarkable fact:

(47) *How phonology could have become perfect*
  a. The grammar directly expresses innate functional principles (defended here).
  b. Humans have been exposed to such a large selection pressure that the substantive details of the innate Universal Grammar have become perfect during the course of evolution (challenged here).
  c. A super-engineer gave us language in a single, inspired stroke (ignored here).

Note that the empiricist standpoint (47a) does not claim that there is no arbitrary phonology. Instead, it claims that **there are no arbitrary universals** (Boersma, to appear):

(48) *Empirical claims of functional phonology*
  a. All universal phonology is directly functional.
  b. All arbitrary phonology is language-specific.

This claim is strong: it challenges all generative theories of autosegmental phonology and feature geometry. I will presently show that with this standpoint, we need fewer assumptions and principles than with any of the generative accounts.

[3] Not just against the insertion of a *path* (nasality linked to a segment by an association line), as in previous constraints, but against the insertion of a separate positive value (stretch) on the naslity tier.

3.2 Walker's (1998) approach to type A

Walker (1998) proposed a family of SPREAD constraints, with an explicit definition in terms of the number of nasal association lines. This hybrid formulation (i.e. a formulation in terms of features that do not distinguish between articulation and perception) is equivalent to the gestural constraint *MOVE as well as to the faithfulness constraint MAXDURATION, at least if the inviolability of line crossing for /mãkã/ is assumed, as it is in most generative theories of autosegmental phonology. Walker expresses the nasalizability hierarchy with feature-cooccurrence constraints:

(49) *Walker's hierarchy of nasalizability*
*NASOBSSTOP >> *NASFRICATIVE >> *NASLIQUID >>
>> *NASGLIDE >> *NASVOWEL >> *NASSONSTOP

These are constraints in the style of the ***grounding conditions*** of Archangeli & Pulleyblank (1994). Such a constraint is thought to have become an innate element of Universal Grammar during the course of evolution, as a result of the selection pressure associated with the interaction between functional principles. In a functional theory of phonology, which expresses function directly, these indirectly functional constraints should be superfluous, like NASSIM.

And indeed, some of Waker's constraints have no correlate in a functional account. The structural constraint *NASOBSSTOP can be written as the cooccurrence filter *[+nas, –cont, –son] (Walker 1998: 36). Walker needs this constraint in order to rule out the unpronounceable nasalized labial plosives, which would otherwise be the winner:

(50) *The need for superfluous cooccurrence constraints*

maka	IDENT-IO (±sonorant)	*NASOBSSTOP	SPREAD (nasal)	*NASVOWEL
maka			***!	
☞ mãka			**	*
mãk̃ã		*!		**
mãŋã	*!			**

In the functional account of (43), a candidate perceived as /mãk̃ã/ can never occur, simply because no articulation can produce it. This means that if we distinguish between articulation and perception in the production grammar, several phonetically impossible combinations do not have to be stated as inviolable constraints in the grammar, as they have to in a grammar with hybrid representations (note, by the way, that such constraints can never contribute to a factorial typology).

The second problem is that the structural constraints do not generate the typology by themselves. To rule out [mãŋã], Walker still needs a faithfulness constraint, equivalent to *REPLACE (k, ŋ). Thus, Walker's approach needs structural as well as faithfulness constraints, whereas the functional approach only needs the faithfulness constraints.

3.3 Sonority hierarchy and type A

The hierarchies in (42) and (49) are reminiscent of the ***sonority hierarchy***, and indeed the sonority scale has come up in at least one account of nasalizability (Gnanadesikan 1995). This is natural from a generative point of view, since the sonority hierarchy is a very good candidate for an innate phonological device, since nearly all languages will then use it for syllabification. However, the position of /h/ in the hierarchy is problematic, as Gnanadesikan notes. In the nasalizability hierachy, faithfulness for /h/ is ranked on a height comparable to that of vowels, because nasalizing this sound will not strongly change its main perceptual features (noise and spectrum). In hierarchies for syllabification, on the other hand, /h/ will pattern with the other fricatives /f/ and /s/ in its preference for the syllable margin, which is again only natural since the sound is voiceless (Boersma 1998: 455). Gnanadesikan gives an example of a two-year-old child, who pronounces |bilˈou| 'below' as [fibˈou], copying the initial obstruent to replace the sonorant onset of the stressed syllable, but pronounces |bihˈajnd| 'behind' as [fihˈajn], *not* copying the initial obstruent to replace the apparently non-sonorant onset of the stressed syllable.

We must conclude that the plausibly innate device of the sonority hierarchy has an exception in the direction of immediate functionality, and is not an arbitrary universal. This is a strong argument against substantive innateness in phonology (Boersma, to appear), especially if the exception does not play a role in many languages, since this would leave only a small number of generations to have selected the presumably innate exception.

3.4 Piggott's (1992) account of type A

Piggott's (1992) account for the nasal-spreading typology (38) proposes some problem-specific innate principles for UG:

(51) *Piggott's principles of nasal harmony (simplified)*

a. The class of blockers must constitute a natural class with the nasal consonants.

Nasals are ***stops***, so one of those classes must be the class of stops: /m/, /n/, /p/, /t/, which accounts for the blockers in Applecross Gaelic.

Nasals are also ***consonantal***, so depending on whether glides are consonantal, we have the classes /m/, /n/, /p/, /t/, /f/, /s/, /l/, /r/ (Warao) and /m/, /n/, /p/, /t/, /f/, /s/, /l/, /r/, /j/, /w/ (Sundanese).

And nasals are ***sonorant***, so we would expect the class /m/, /n/, /l/, /r/, /j/, /w/, i.e. a language in which obstruents are targets, but sonorants block!

b. The class of blockers must not be limited to sonorants.

This ad-hoc exception rules out the third possibility in (51a).

c. There is a natural class called ***non-approximant consonants***.

This ad-hoc class consists of /m/, /n/, /p/, /t/, /f/, /s/, accounting for Ijo.

While (51a) sounds like a general principle that could find application in other areas, the UG principles (51b) and (51c) are obviously specific to the problem of type-A nasal harmony. Since these principles are of advantage to only a very small minority of languages, they are very unlikely to have had any chance of emerging by a rich enough selection during the course of evolution (a few hundred generations).

4 *Nasal harmony, type B*

Type-B nasal-harmony languages (Piggott 1992) have word-level specifications for [+nasal] or for [–nasal]. Every [–nasal] segment has a [+nasal] counterpart:

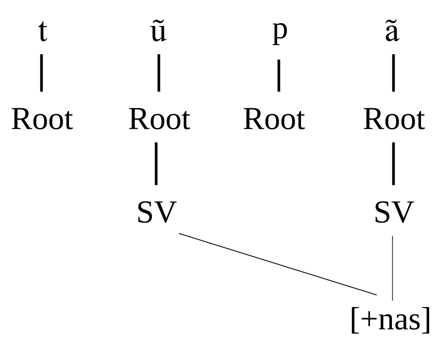
(52) *Type-B nasality contrasts (Barasano)*

[–nasal]	[+nasal]
a, u	ã, ũ
w, j	w̃, j̃
l, r	l̃, r̃
ᵐb	m
s	s
t, k	t, k

4.1 Transparency of plosives

One of the conspicuous properties of type-B nasal-harmony languages is the transparency of plosives to nasal spreading. So, in Guaraní we have [tupa] ‘bed’ and [tũpã] ‘god’, but no *[tupã] or *[tũpa] (Piggott 1992). In Piggott’s analysis, nasality is spread from right to left across all segments that have a Spontaneous Voicing (SV) node, i.e. all sonorants:

(53) *Piggott’s spreading along the Spontaneous Voicing tier*



Since the two vowels are adjacent on the SV tier, Piggott’s analysis has the desirable property of ***locality*** in spreading processes. As a detail, we may note that the always recalcitrant segment /h/ has no SV node, so that Piggott’s analysis would predict that it is transparent (not nasalizable).

Piggott & Van der Hulst (1997) reanalyse the process as spreading on the syllable level: a nasalized vowel, being the head of its syllable, makes nasalization a property of the syllable, and this then spreads to adjacent syllables, nasalizing all the sonorants in every affected syllable. This move allows them to account for more facts than Piggott (1992), such as the fact that all sonorants in syllables with nasalized vowels are nasalized themselves, and the similarity with vowel-harmony processes. Again, however, the locality requirement has informed the search for a higher structure in which nasalization is continuous.

Walker (1998) also explicitly wants to honour the locality requirement, invoking the line-crossing constraint as an inviolable well-formedness condition on (hybrid) phonological representations. Her analysis, though stated in terms of the ***sympathy*** device introduced by McCarthy (1998), is equivalent to the following derivation:

(54) *Walker's derivation of transparency*

$$\begin{array}{l} \text{tupa} \\ \big\downarrow \ \text{nasalization} \\ \text{t̃ũp̃ã} \\ \big\downarrow \ \text{denasalization of plosives} \\ \text{tũpã} \end{array}$$

Both Piggott's and Walker's theories work. However, a theory that distinguishes between articulation and perception in phonology must maintain that feature geometries are illusions evoked in the linguist who advocates hybrid representations, and that these illusions will evaporate if the correct distinctions are made (Boersma 1998: 22, 442). I will show that if we separate articulation and perception, we do not need Piggott's feature geometry or Walker's counterbleeding serial derivation.

Because of the word-level specification, which succesfully nasalizes at least all vowels, sequences of two nasal vowels with an intervening obstruent are very common in these languages. For this reason, the OCP for nasality may well outrank the line-crossing constraint in these exceptional cases:

(55) *Perceiving nasality across a plosive*

acoustics: [tũpã]	OCP (nas: +; Ṽ | plosive | Ṽ)	LCC (nas: +; Ṽ | plosive | Ṽ)
perception: [–nas] [+nas] [–nas] [+nas] / t ũ p ã	*!	
☞ perception: [–nas] [+nas] [–nas] / t ũ p ã		*

Thus, the two nasalized vowels are perceived with a single value [+nasal] on the perceptual nasality tier, despite the intervening plosive. The articulation, on the other hand, cannot be regarded as continuous, since the velum has to go up and down for the labial plosive. This combination of discontinuous articulation and perceptual unity can de pictured as:

(56) *Asymmetry between articulation and perception*

[–nas] [+nas] [–nas]

t ũ p ã

velum: up down up down

The universality of the locality condition is an illusion brought about by the ubiquity of articulation-based spreading, which is never discontinuous (because that would violate its very purpose of gesture reduction), and by the rarity of a high OCP across salient intervening material. However, a high rate of cooccurrence of nasality in adjacent syllables, as in these languages, will shift the balance in the direction of perceptual unity. At least, if speech perception is like other kinds of perception.

4.2 Why all sonorants are nasalizable in type-B languages

Another conspicuous property of type-B nasal-harmony languages as that they all nasalize their glides and liquids. A hierarchy of nasalizability, similar to the one for type-A languages, does not appear to exist. This may be due to several causes.

First, type-A languages seem to be like most languages in that they have to deal with constraints against the replacement of the trill |r| with a nasalized trill /r̃/. If the constraint refers to the difference between the two, and the difference is in the nasality, we must conclude that the underlying form contains a [–nasal] specification, or, if you like privative features, at least the surface form violates *INSERT PATH (+nasal).

This will be different in type-B languages, which make a point of applying nasality to the word level. If nasality is suprasegmental, segments are less likely to be specified for nasality themselves. So the perceptual specification of the segment will not contain any specification for [–nasal]. The only relevant specification for |r| is |trill|, and both /r/ and /r̃/ honour *DELETE (trill). So the relevant specifications for all segments are:

(57) *Type-B nasality contrasts (Barasano)*

[–nasal]	[+nasal]	
a, u	ã, ũ	[low/high vowel, +son]
w, j	w̃, j̃	[back/front glide, +son]
l, r	l̃, r̃	[lat/trill, +son]
ᵐb	m	[stop, +son] (following Piggott 1992)
s	s	[fricative, –son]
t, k	t, k	[plosive, –son]

This means that the surface forms handle the faithfulness constraints in the following way:

(58) *Faithfulness handling*

a. All the specified features surface faithfully in oral as well as in nasal words.
b. The sonorant stops violate *INSERT (half a nasal) in oral words.
c. The obstruents violate *DELETEPATH (nasal).

Here are the tableaus for all segments:

(59) *Nasalizing a liquid (or vowel or glide)*

|ara + *nasal*|	*DELETE (*any feature*)	*DELETEPATH (half a nasal)	*INSERT (half a nasal)	*GESTURE (velum)
☞ [ã r̃ ã] /ã r̃ ã/				
[ã r ã] /ã r ã/		*!*		**

(60) *Nasalizing a sonorant stop*

|a[+son,stop]a + *nasal*|	*DELETE (*any feature*)	*DELETEPATH (half a nasal)	*INSERT (half a nasal)	*GESTURE (velum)
☞ [ãmã] /ãmã/				
[ã^{m}bã] /ã^{m}bã/		*!		**
[ãbã] /ãbã/	*!	**		**

(61) *Oralizing a sonorant stop*

|a[+son,stop]a|	*DELETE (*any feature*)	*DELETEPATH (half a nasal)	*INSERT (half a nasal)	*GESTURE (velum)
[ama] /ama/			**!	**
☞ [a^{m}ba] /a^{m}ba/			*	**
[aba] /aba/	*!			

(62) *Nasalizing a plosive (or fricative)*

|a[–son,plos]a + *nasal*|	*DELETE (*any feature*)	*DELETEPATH (half a nasal)	*INSERT (half a nasal)	*GESTURE (velum)
[apa] /apa/		**		
☞ [ama] /ama/	*!*			**

The ranking can be summarized as *DELETE (segmental feature) >> FAITH (nasal) >> *GESTURE (velum). The first ranking seems quite natural, because nasality will always be realized on the vowels, so that it is not very important to have it realized on the consonants as well. The second ranking expresses the idea that velar gestures do not play any role in the phonology. In this respect, type-B languages differ completely from type-A languages.

There is a possible problem for the current analysis. What if type-B languages are like type-A languages in that they do have underlying oral segments? In that case, we must still account for the nasalization of [ãr̃ã], since this violates *INSERTPATH (nasal). I propose that the form [ãrã] is ruled out by the higher-ranked constraint *GESTURE (velum: up & down / fast). So we are left with having to explain the universal high ranking of this constraint (though it is violated in [ãpã]). First, this gestural constraint will be ranked higher than *MOVE (velum), since its violation *always* involves *two* velum movements, where honouring *MOVE *sometimes* involves *one* velum movement. Secondly, *GESTURE is parametrized by the condition "/ fast", because a normal single velum raising gesture followed by a normal single velum opening gesture will give rise to gestural overlap, which will cause the velum to stay more or less closed, thus violating *INSERTPATH (nasal) after all.

Having established the relatively high ranking of the fast double gestural constraint, we have to show why its ranking above *INSERTPATH (nasal) can be universal (it must be ranked below *DELETE (–sonorant), though), i.e. why the typology predicted by the local-ranking principle does not work. As an example, consider the ***horse***.

There is a probably near-universal, though definitely not innate, ranking for the judgements for the perceptual distinction between a horse and some other animals:

(63) *The horse-distance hierarchy, compared to the apple-pear difference*

horse - ant	
horse - duck	
horse - dolphin	
horse - rhino	
horse - cow	↑
horse - deer	apple-pear
horse - donkey	↓

Now ask people to rank the difference between an apple and a pear along this scale. The average judgement for the apple-pear difference lies somewhere near the horse-deer difference, though deviations from the average will occur, even below the horse-donkey difference or above the horse-cow difference, but definitely not extending toward the horse-duck difference. The local-ranking principle as formulated in Boersma (1998: 17) would predict that any ranking of apple-pear along scale (63) should be possible, though large deviations from the global measure of distinguishability will be uncommon, some even to the extent that their number in any finite group of informants is zero. Likewise, the number of languages of the world may be too small for there to be even a single language that shows the ranking A B, if the globally determined average ranking of B is much higher than that of A, even if A and B are incommensurable (e.g. A is a gestural, and B a faithfulness constraint).

Thus, the ranking *GESTURE (velum: up & down / fast) >> *INSERTPATH (nasal) may well be near-universal. The reader is advised to try to produce a non-nasal liquid between two nasalized vowels.

Conclusion

Generative accounts of nasal harmony have to take recourse to ad-hoc natural classes, exceptions to exceptions, grammaticization of constraints against unproducable perceptual output, functional exceptions to innate hierarchies, feature geometry, and derivation. If all these things were really needed, UG would be full of substantive phonological detail. However, the functional approach to phonology can account for the facts of nasal harmony without assuming anything but general properties of human motor behaviour and perception. This is compatible with the view that the phonological part of the innate language device does not contain much more than: the cognitive abilities of categorization, abstraction, wild generalization, and extrapolation; the storage, retrieval, and access of arbitrary symbols; a stochastic constraint grammar; a gradual learning algorithm; laziness; the desire to understand others; and the desire to make oneself understood.

References

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded phonology*. Cambridge: MIT Press.

Bloomfield, Leonard (1933). *Language*. British edition, 1935. London: George Allen & Unwin.

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

Boersma, Paul (to appear). *Typology and acquisition in functional and arbitrary phonology.* Third Utrecht Phonology Workshop, June 1998.

Chomsky, Noam (1964). *Current issues in linguistic theory*. The Hague: Mouton.

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

Cohn, Abigail (1993). The status of nasalized continuants. In Marie K. Huffman & Rena Krakow (eds.), *Nasals, nasalization, and the velum. Phonetics and phonology*, Vol. 5. San Diego: Academic Press. 329–367.

Gnanadesikan, Amalia (1995). *Markedness and faithfulness constraints in child phonology*. Manuscript, University of Massachusetts. [Rutgers Optimality Archive **67**, http://ruccs.rutgers. edu/roa.html]

Gussenhoven, Carlos (1998). *Can phonological constraints and phonetic constraints appear in the same grammar?* Handout CNRS-LOT Phonology Workshop, Nijmegen, October 28, 1998.

Halle, Morris (1959). *The sound pattern of Russian*. The Hague: Mouton.

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

Hockett, Charles. F. (1965). Sound change. *Language* **41**: 185–205.

McCarthy, John J. (1998). *Sympathy and phonological opacity*. Ms. Universaity of Massachusetts, Amherst. [Rutgers Optimality Archive, http://ruccs.rutgers. edu/roa.html]

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]

Piggott, Glyne (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* **10**: 33-78.

Piggott, Glyne & Harry van der Hulst (1997). Locality and the nature of nasal harmony. *Lingua* **103**: 85–112.

Postal, Paul M. (1968). *Aspects of phonological theory*. New York: Harper and Row.

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**.

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

Walker, Rachel (1998). *Nasalization, neutral segments, and opacity effects*. PhD dissertation, University of California, Santa Cruz.