



## UvA-DARE (Digital Academic Repository)

### Introduction: models of phonology in perception

Boersma, P.; Hamann, S.R.

**Publication date**

2009

**Document Version**

Submitted manuscript

**Published in**

Phonology in perception

[Link to publication](#)

**Citation for published version (APA):**

Boersma, P., & Hamann, S. R. (2009). Introduction: models of phonology in perception. In P. Boersma, & S. Hamann (Eds.), *Phonology in perception* (pp. 1-24). (Phonology and phonetics; No. 15). Mouton de Gruyter.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Introduction: models of phonology in perception

Paul Boersma and Silke Hamann, 8 April 2009

The aim of this book is to provide explicit discussions on how perception is connected to phonology. This includes discussions of how many representations a comprehensive view of phonology requires, and how these representations are mapped to each other in the processes of comprehension and production. Of the two directions of processing, this book centres on comprehension, the direction that has received relatively little attention from phonologists.

This introduction makes an attempt at providing a single common formalization for the various models that have been proposed in the literature, including those that are proposed by the authors in this volume.

The first step is to make explicit what representations and processes we are talking about. As for the notations of the representations, we use pipes for lexical or underlying forms (e.g. German |tag + əs| ‘day-GENITIVE’, with morpheme structure), square brackets for overt or phonetic forms (e.g. [t<sup>h</sup>a:gəs], with articulatory and/or auditory detail), and slashes for any non-underlying non-overt representations (e.g. /(tá.gəs)/, with foot and syllable structure). As for the notations of the processes, we use arrows (e.g. |tag + əs| → [t<sup>h</sup>a:gəs] for phonological-phonetic production, or [t<sup>h</sup>a:gəs] → |tag + əs| for phonological-phonetic comprehension).

The second step is to make explicit how the processes of comprehension and production work. For this introduction we assume that the listener’s comprehension process starts from an auditory phonetic representation and aims at arriving at an underlying lexical-phonological representation, and that the speaker’s production process starts from an underlying lexical-phonological representation and aims at arriving at an articulatory phonetic representation. All the grammar models we discuss in this introduction agree that these processes are at least partly guided by the grammar. The various models differ, though, in the number and kinds of representations that they consider, and in how they express the relationships between the representations.

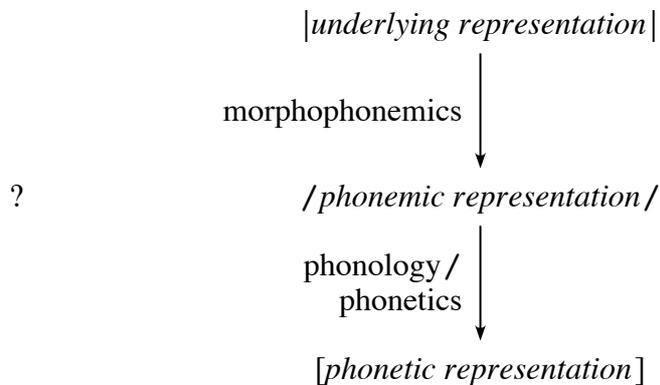
## 1. The structuralist grammar model

The oldest phonological grammar model in current use is the pre-generative structuralist model depicted in (1).

(1) Structuralist grammar model

COMPREHENSION

PRODUCTION



As in all the pictures in this introduction, we draw the comprehension model on the left and the production model on the right. However, the structuralists did not model the comprehension part, so we have a question mark on the left side of (1). On the production side we see three levels of representation, connected with two arrows, each of which represents a module of the grammar.

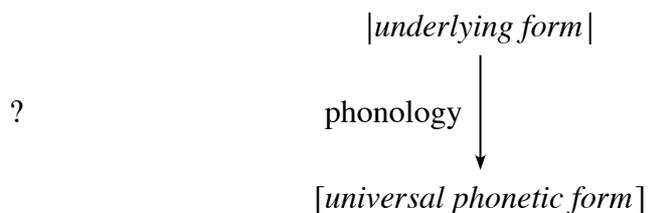
**2. Chomsky and Halle's grammar model**

In *The sound pattern of English*, Chomsky and Halle (1968) deliberately disposed of the intermediate form (the phonemic representation) of the structuralists, as in (2).

(2) Early generative phonology

COMPREHENSION

PRODUCTION

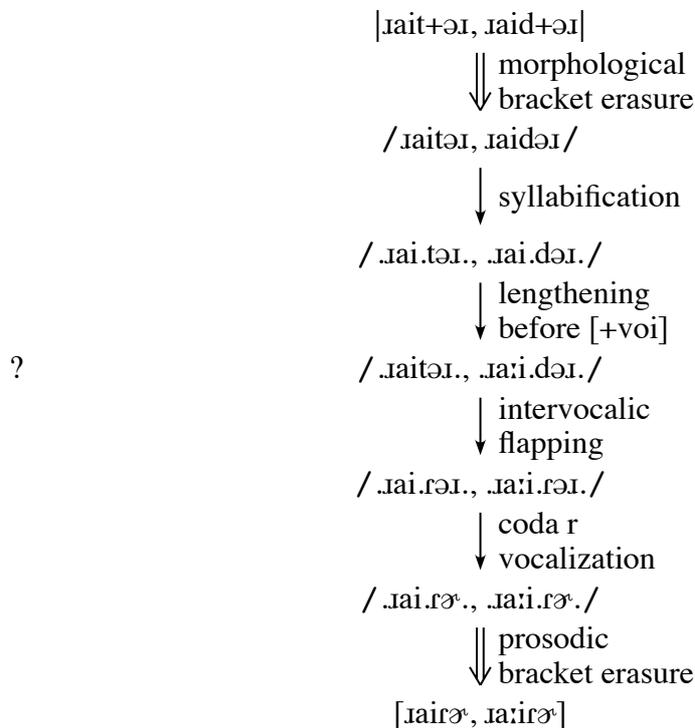


This production model can be looked upon in two ways. One way is to regard it as a single mapping, namely from underlying to surface form, and the other way is to regard it as a sequential series of mappings, starting with the lexical form, going through a multitude of intermediate forms, and ultimately resulting in the surface form. An example from English is shown in (3).

(3) Sequentially ordered rules with many unnamed intermediate representations

COMPREHENSION

PRODUCTION



On the right side of (3) we distinguish two kinds of mappings: universal ones, depicted with double arrows, and language-specific ones, depicted with single arrows. In this picture, the removal of the plusses in the first step can be regarded as universal if no phonological rules can ever be conditioned by such morphological boundaries. Likewise, the removal of the syllable boundaries in the last step can be regarded as universal if prosodic boundaries have to be universally absent from phonetic forms. The remaining four steps are ordered, language-specific processes and therefore depicted with single arrows. If one's views on what is universal and what is not are different from the ones just expressed, one would draw (3) differently.

The comprehension part in (3) is still a question mark. Comprehension could be implemented as the reverse application of rules until the listener ends up with a form that he/she has stored in the lexicon. In (3) this would indeed work for most of the steps. Starting from the auditory forms [ɹairə, ɹai:rə], the listener could perhaps undo “prosodic bracket erasure” by automatically inserting syllable boundaries (although there may not be any guarantee that syllable boundaries are universally audible). The next step, “coda r vocalization” can be undone by changing every instance of /rə/ into /əɪ/; this happens to work in this case because the only source of the retroflex vowel /rə/ in English is the sequence /əɪ/ (i.e., /rə/ is not the result of a neutralizing rule). The next step, intervocalic flapping, can be undone by changing every /r/ into /d/ after a lengthened vowel and into /t/ after a non-lengthened vowel; this happens to work in this case because the English flapping rule is not completely neutralizing: a cue to the underlying voicing of the plosive is still available in the duration of the preceding vowel. The next step, pre-voice lengthening, can be undone by changing every /ai:/ into /ai/. The next step, “syllabification”, can be trivially

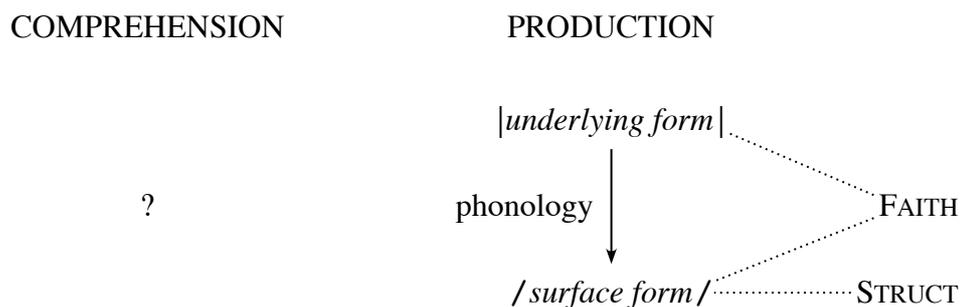
undone by removing all syllable boundaries. Only the final step, “morphological bracket erasure”, is thoroughly problematic, because it is incorrect to assume that a “+” can be inserted before every final /əɪ/ (consider the English monomorphemic words |maitəɪ| ‘mitre’ and |saidəɪ| ‘cider’); to resolve this, an English listener requires lexical information.

Whereas in (3) a listener could reasonably successfully undo each production rule, such a comprehension strategy would not work in general. Especially cases where production rules lead to complete neutralization cannot be undone. Consider the case of the French words |ʒɔli| ‘nice’, |gʁoz| ‘fat’, and |pətiti| ‘small’. That these words must have these underlying forms is known from phrases like [ʒɔliami, gʁozami, pətitami] ‘nice friend, fat friend, small friend’; however, in isolation the three words are [ʒɔli, gʁo, pəti]. At some point in the derivation of production, therefore, there must be a rule that deletes final consonants, and this rule is completely neutralizing, i.e., the listener cannot know whether the “final consonant deletion” rule applied in comprehension would have to turn /.pə.ti./ into /.pə.tit./, /.pə.tiz./, or /.pə.ti./ . Because of such ambiguities, the listener cannot generally retrace the sequence of the production rules, so in the general case a Chomsky-and-Halle listener would have to consider a large number of underlying forms, compute a surface form from each, and thus decide which underlying form best matches both the incoming phonetic form and information from the lexicon and other modules. This ‘considering multiple forms’ is something that Optimality Theorists are comfortable with (at least in modelling production), and the next sections will show that indeed the parallel framework of Optimality Theory is better suited to handling comprehension than Chomsky and Halle’s sequential rule framework is.

### 3. McCarthy and Prince’s grammar model

Stepping into the field of Optimality-Theoretic proposals, we see that the simplest grammar model, namely that proposed by Prince and Smolensky (1993) and McCarthy and Prince (1995), has only two representations, namely underlying form and surface form.<sup>1</sup> As with the rule-based model of §2, this model is basically production-only, as is shown explicitly in (4).

#### (4) Two-level OT



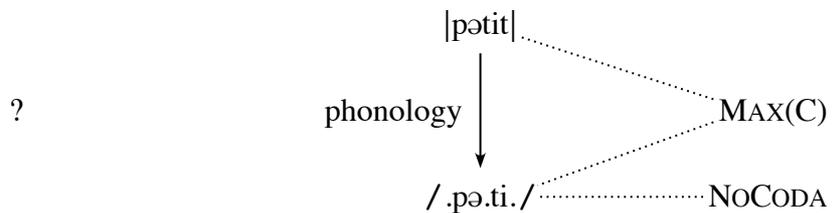
The surface form has been written here between slashes, because it may contain hidden material such as syllable boundaries and foot structure, i.e. it may not be a solely phonetic form. Example (5) applies the format of (4) to the French example discussed in §2.

<sup>1</sup> This formulation simplifies away from the fact that Prince and Smolensky’s output candidates were so-called “full structural descriptions” from which both the underlying form and the phonetic form could be derived mechanically, and from the fact that McCarthy and Prince made an additional representational distinction within the surface form, namely that between base and reduplicant. The model in (4) is more representative of the subsequent body of literature that was based on Prince and Smolensky’s and McCarthy and Prince’s proposals than of the two original proposals themselves.

(5) French coda deletion in two-level OT

COMPREHENSION

PRODUCTION



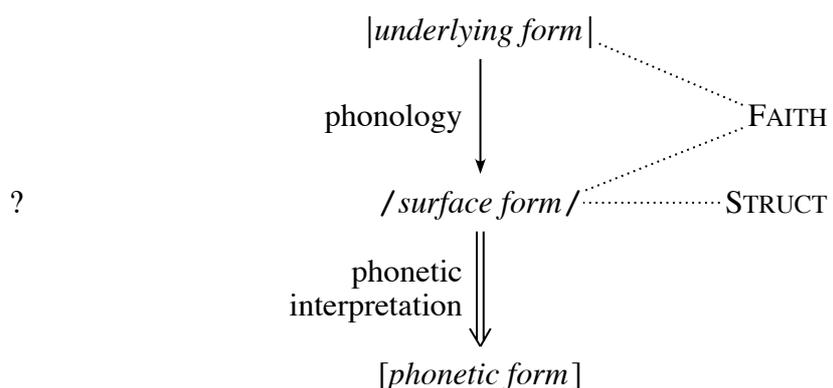
In (5), comprehension is still depicted by a question mark, because the early OT-ists did not consider it. The most naive way of implementing comprehension would be to copy from the serial-rule framework (§2) the idea of computing surface forms for a large number of underlying forms, then choosing all underlying forms that match the perceived surface form. This ‘analysis-by-synthesis’ idea was applied to Optimality Theory by Hale and Reiss (1998). However, this way of doing comprehension does not seem to entirely fit the spirit of OT, because the formulation of nearly all constraints proposed in the OT literature seems to allow them to be used in two directions, i.e. for comprehension as well as for production. We will investigate this point in the following sections. First, however, we describe a more elaborate production-only grammar with three representations.

When McCarthy and Prince are asked what the overt phonetic form looks like, they reply that the surface form contains both hidden material and overt phonetic detail. The phonetic form can then be computed in a universal way by removing from the surface form anything that is universally hidden, such as information on foot structure and on syllable boundaries. The grammar model would become:

(6) How to include phonetic implementation if one considers it universal

COMPREHENSION

PRODUCTION

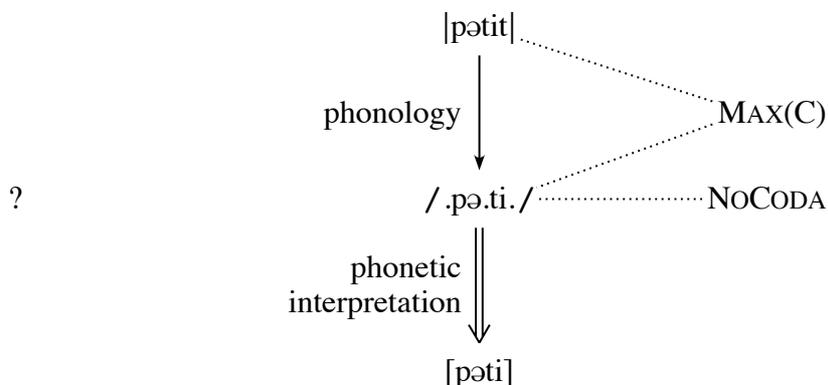


The double arrow here shows the idea that phonetic interpretation is universal. In our French example in (7), for instance, all that happens is that the syllable boundaries are removed.

(7) *Universal phonetic implementation in French*

COMPREHENSION

PRODUCTION

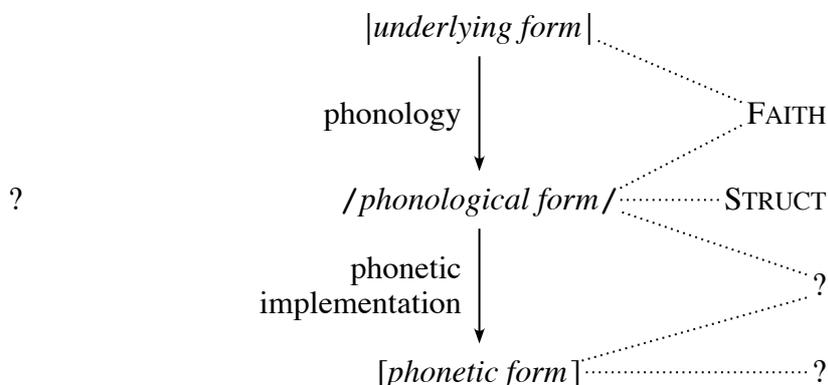


If one thinks that phonetic interpretation is a linguistically relevant language-specific module, one may want to call it ‘phonetic implementation’ and promote its arrow to a single one. This is what Hayes (1999) did in a footnote (p.250). He proposed that this module was Optimality-Theoretic, but did not supply any constraints or examples. This is shown in (8).

(8) *Hayes' footnote*

COMPREHENSION

PRODUCTION

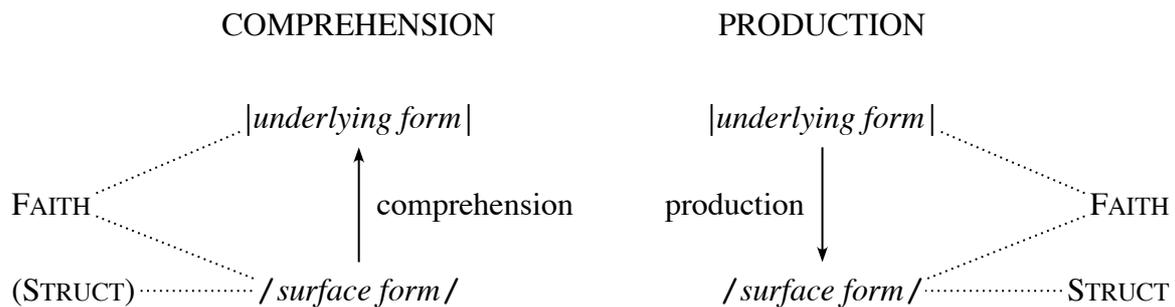


Since the topic of the book is perception, the following sections do away with the question mark on the left side, and consider bidirectional models.

**4. Smolensky's bidirectional grammar model (1996)**

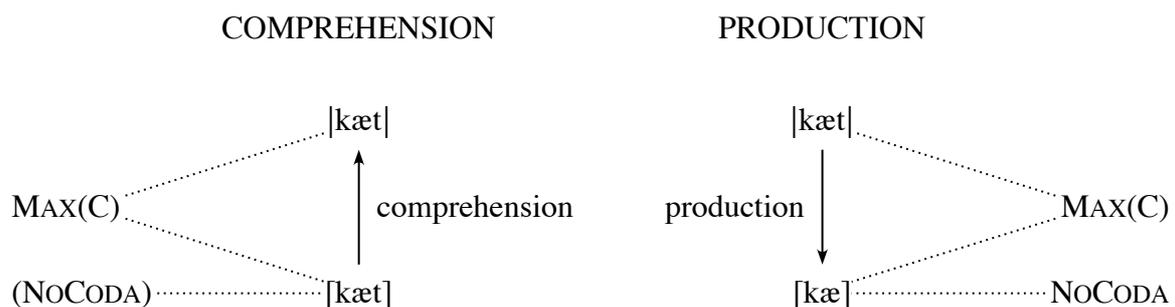
The simplest *bidirectional* grammar model, i.e. a grammar model that includes both production and comprehension, is that by Smolensky (1996), shown in (9).

(9) *Smolensky's (1996) grammar model*



Smolensky proposed that production and comprehension are handled by the same OT constraints with the same rankings, but that the STRUCT constraints do not really work in comprehension, since they evaluate the *input* to comprehension, which is the same for all candidates. Thus, comprehension for Smolensky (1996) is a faithfulness-only mapping, which usually renders it flawless. This allowed Smolensky to explain why young English children may pronounce *|kæt|* ‘cat’ as *[kæ]* themselves but would at the same time object if an adult also pronounces this word as *[kæ]*. The example is made explicit in (10).

(10) *Smolensky's (1996) grammar model*



In (10), the surface forms are between square brackets, in (9) between slashes. In a two-level phonology such as this, this choice is not usually so important: one could write slashes because the form may contain hidden material, and brackets because the form contains phonetic detail.

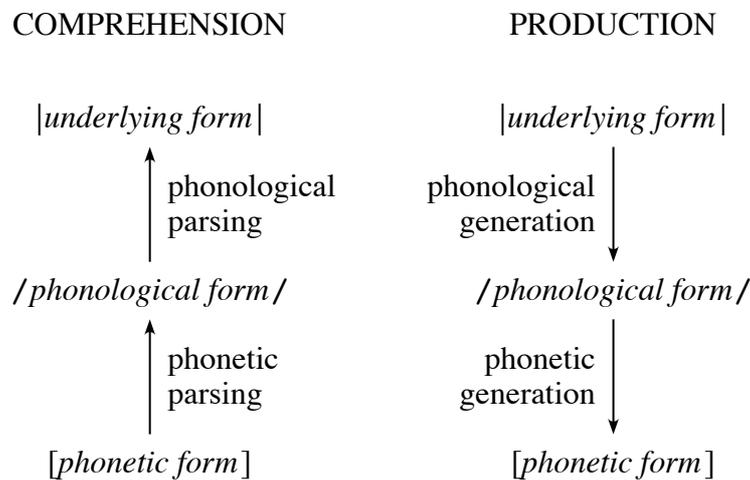
There are several ways to propose constraints in a model like (9). While Smolensky proposed that production and comprehension used the same constraints with the same rankings, one could also use the same constraints with different rankings (Kenstowicz 2001), or one could use partly different sets of constraints (Pater 1999; Broselow 2004; Yip 2006).

### 5. Bidirectional models with three representations

Four sources of evidence suggest that the bidirectional OT grammar model of (9) might be too simple, in the sense that we might need three representations rather than two. The first source of evidence comes from phonetics and psycholinguistics. Phoneticians could argue that the model in (9) does not really include *perception*. If perception is regarded as the mapping from a universal phonetically detailed form (*overt form*, perhaps) to a language-specific phonological structure (*surface form*, perhaps) without lexical access, an inclusion of perception into the grammar model could require three representations rather than two. Similarly, many psycholinguists, e.g. McQueen and Cutler (1997), argue that comprehension

is not a single module but consists of two sequentially ordered modules for prelexical perception and for word recognition, as in the model on the left side in (11). Analogous modularity has been proposed for production, e.g. Levelt (1989), and this is shown on the right side in (11) (also see (8)).

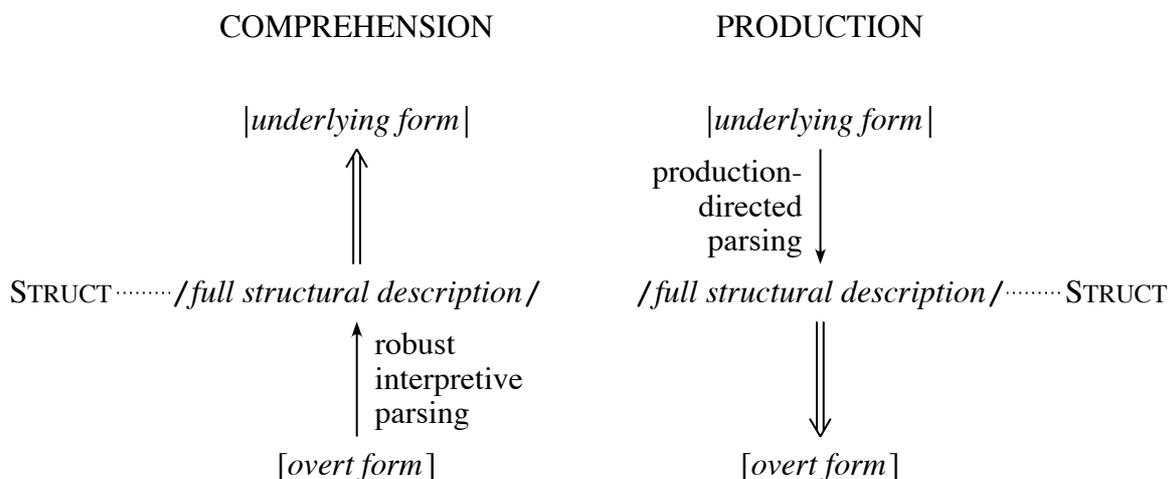
(11) *Two sequentially modular psycholinguistic models*



Grammar models with three representations have also been proposed on the basis of linguistic evidence. The listener is confronted with overt phonetic forms from which she has to construct an abstract surface structure in a language-specific way (Tesar 1997 et seq., Boersma 1998 et seq.), and in opposition to Smolensky's (1996) faithfulness-only comprehension model in (9), prelexical perception by young children is not flawless and therefore has to be modelled as a developing constraint system itself (Boersma 1998; Pater 2004).

A simple computational bidirectional grammar model with three representations was devised by Tesar (1997, 1998, 1999, 2000) and Tesar and Smolensky (1998, 2000). It can be abbreviated as in (12).

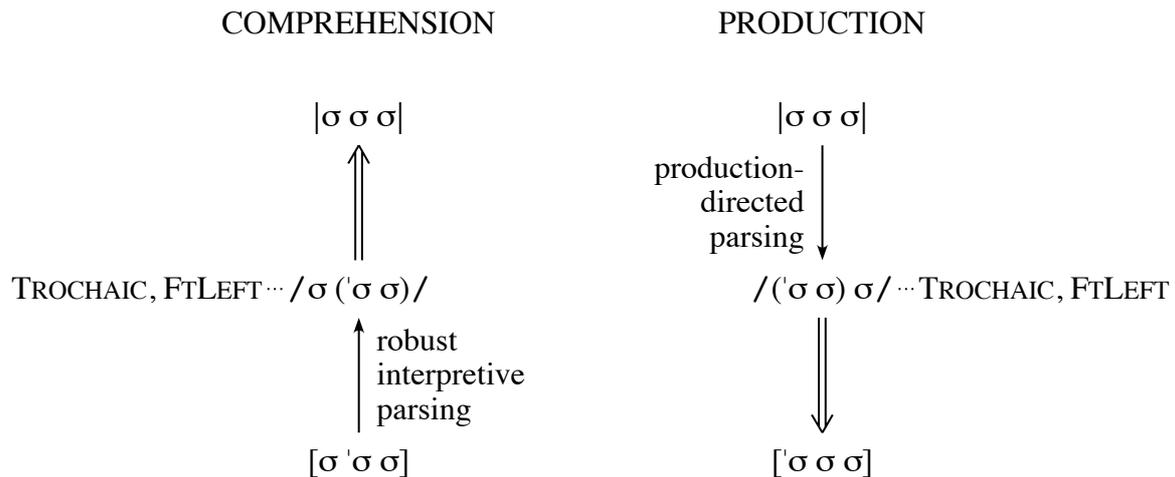
(12) *Tesar and Smolensky's grammar model*



The single arrows here indicate the two processes that are handled by the grammar. The 'robust interpretive parsing' module can be said to perform perception, but since the full

structural description *contains* the underlying form (as in Prince and Smolensky’s notion of ‘containment’), this interpretive parsing may involve lexical access as well. The ‘production-directed parsing’ module computes the full structural description in production, using the same constraints and rankings as in interpretive parsing. Since the full structural description contains enough information to derive both the underlying form and the overt form without looking at the grammar, the remaining two mappings (which we can equate with recognition and phonetic implementation) are depicted with double arrows. Here is an example:

(13) *Tesar and Smolensky’s metrical phonology learner*



This picture shows the reason for the two double arrows: Tesar and Smolensky’s mapping from full structural description to underlying form is universal (it removes parentheses and stress marks), as is their mapping from this surface form to the overt form (it removes parentheses).

All the OT bidirectional models with three representations have been informed by considerations of *learning*. Example (13), for instance, shows what happens if the ranking is TROCHAIC >> { IAMBIC, FTLEFT } >> FTRIGHT: when confronted with the second-syllable-stressed overt form  $[\sigma \sigma \sigma]$ , the listener will parse (perceive) it as / $\sigma$  ( $\sigma \sigma$ )/, because that form satisfies the constraints better than its sole competitor /( $\sigma \sigma$ )  $\sigma$ / would do; however, when the listener computes what she herself would have said given the underlying form  $|\sigma \sigma \sigma|$ , the result is /( $\sigma \sigma$ )  $\sigma$ /. This discrepancy between the surface forms in comprehension and production is taken by Tesar and Smolensky as evidence that the listener’s grammar is incorrect; as a result, the listener (*learner*) may change the ranking of her constraints.

In an OT model *without* containment, the recognition (lexical access) module will be language-specific, so that the comprehension model will get two single arrows, thus returning to Cutler’s psycholinguistic model on the left in (11). Models that implement this were devised by Boersma (1998, 2001), Pater (2004), and Boersma (2007). We will discuss all three in the following.

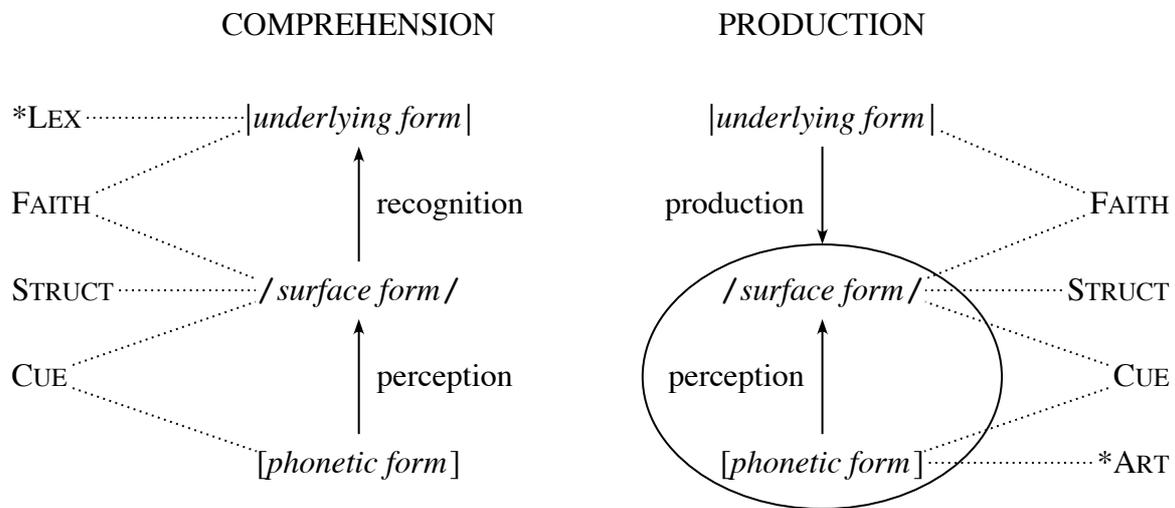
The model by Boersma (1998, 2000, 2001, 2003) has all the representations in (11), i.e. three representations in comprehension as well as in production,<sup>2</sup> although it lacks a phonetic

---

<sup>2</sup> The actual model also divides up the phonetic form into an articulatory form and an auditory form.

implementation module in production; instead, phonetic implementation is implemented as a *perceptual control loop*. An attempt at a visualization is shown in (14).

(14) Boersma's (1998) bidirectional grammar model with control loop



On the comprehension side in (14), we see that in contrast with Tesar and Smolensky's model, the prelexical perception module involves *cue constraints* (Boersma 1998, 2000, 2007; Escudero and Boersma 2003, 2004; Escudero 2005) whose ranking expresses the details of phonological categorization, i.e. the language-specific mapping from auditory cues to phonological elements. In general, then, perception is seen as an interaction between structural restrictions (\*STRUCT) and a sort of auditory-to-phonological faithfulness (\*CUE); this interaction has been discussed by Boersma (1998, 2000), R.Hayes (2001ab), and Pater (2004). In the present book, the cue constraints are discussed in three chapters: the chapter by **Boersma** discusses how cue constraints interact with structural constraints in perception; the chapter by **Hamann** discusses how various rankings of cue constraints help explain sound change in first-language acquisition; and the chapter by **Escudero** discusses how the ranking of cue constraints changes during second-language acquisition.

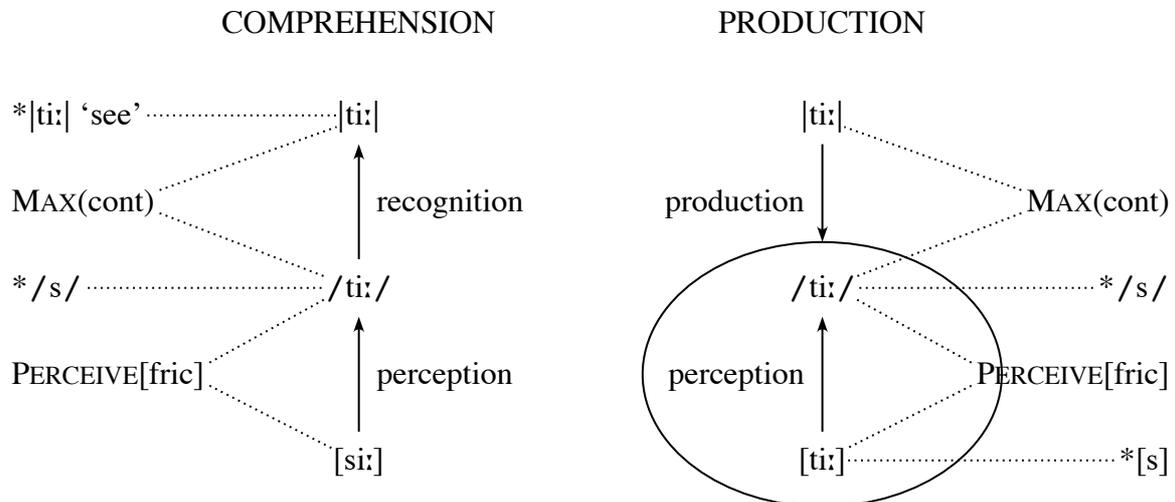
Another difference between the model in (14) and Tesar and Smolensky's model is the inclusion of *lexical constraints* (\*LEX) in the recognition process. According to Boersma (2001), these constraints use the lexicon to help to resolve ambiguities, a property that improves the bidirectional use of faithfulness constraints.<sup>3</sup> Escudero (2005) used these constraints for modelling the loss of phonological categories in second-language acquisition, and Apoussidou (2007) used them for modelling the acquisition of abstract underlying forms.

The third difference is that (14) includes phonetic detail in production. It does so by regarding phonetic implementation as maximally listener-oriented, namely as the speaker's view of how the listener will perceive the speaker's pronunciation. As a result, structural constraints (i.e. constraints against phonological structures and phonotactics) evaluate the output of perception only, both in comprehension and in production. To show that this model handles some perception-precedes-production effects in acquisition, we consider an example from Boersma (1998) in two pictures. In (15) we see a stage in which the learner does not yet

<sup>3</sup> The model by Smolensky (1996), which we discussed in §4, failed to resolve lexical ambiguities. This failure was the reason for Hale and Reiss (1998) to propose the production-only use of faithfulness constraints that we discussed in §3.

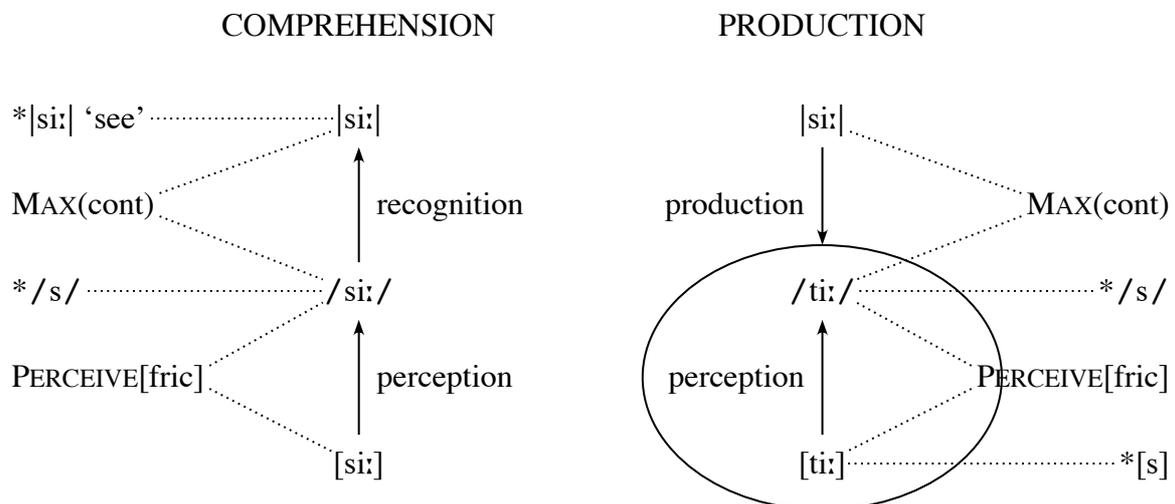
perceive the English /s/-/t/ contrast, so that the English word ‘see’, produced by an adult as [si:], is perceived as /ti:/, hence stored as |ti:|, hence pronounced as [ti:]. The ranking in the perception grammar is \*/s/ >> PERCEIVE[fric].

(15) *Boersma’s example of non-adultlike perception and production*



In (16) we see the next stage, in which PERCEIVE[fric] outranks \*/s/, so that ‘see’ is perceived and stored correctly. However, the articulatory-effort constraint against [s] (e.g. against the complicated tongue-grooving gesture) still outranks MAX(cont), so that |si:] is still pronounced [ti:].

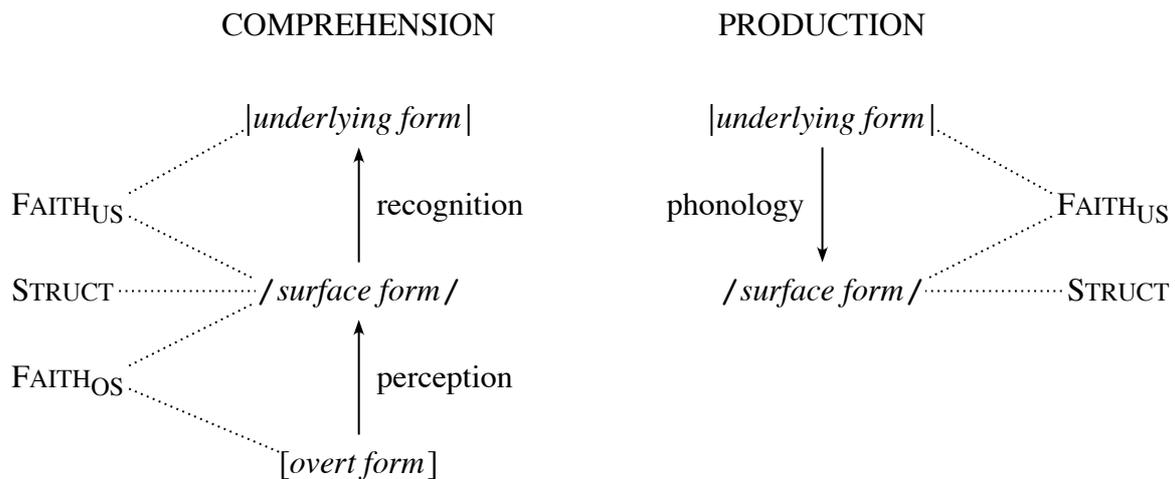
(16) *Boersma’s example of adultlike perception and non-adultlike production*



Since the two surface forms, /si:/ in perception and /ti:/ in production, are different, the learner will be able to take action and change the ranking of MAX(cont) and \*/s/, analogously to what happened in Tesar and Smolensky’s case of (13).

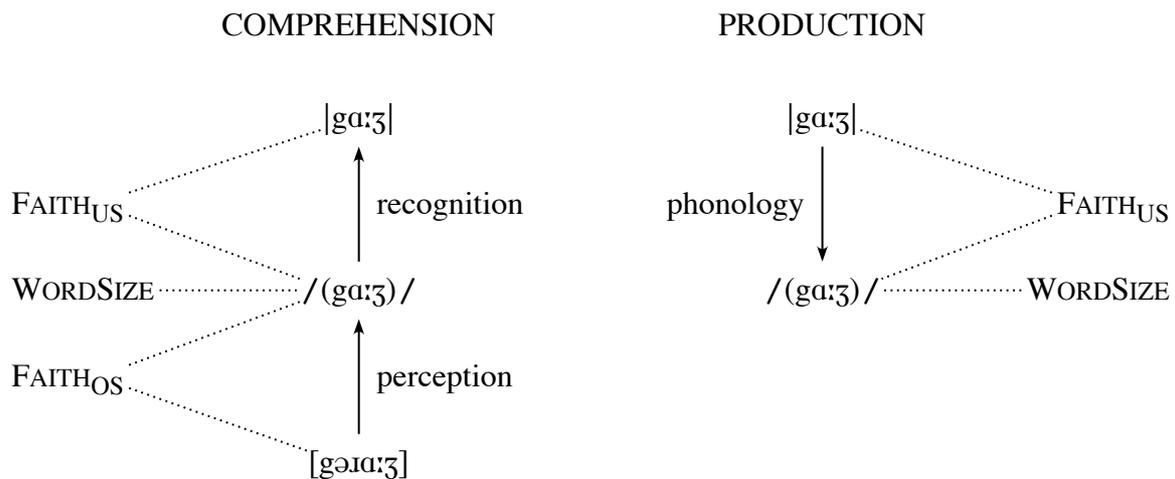
Another example of perception-precedes-production is given by Pater (2004), who has the slightly simpler model shown in (17): when compared with (14), this model is identical with respect to comprehension but does not include phonetic detail in production.

(17) *Pater's partial grammar model*



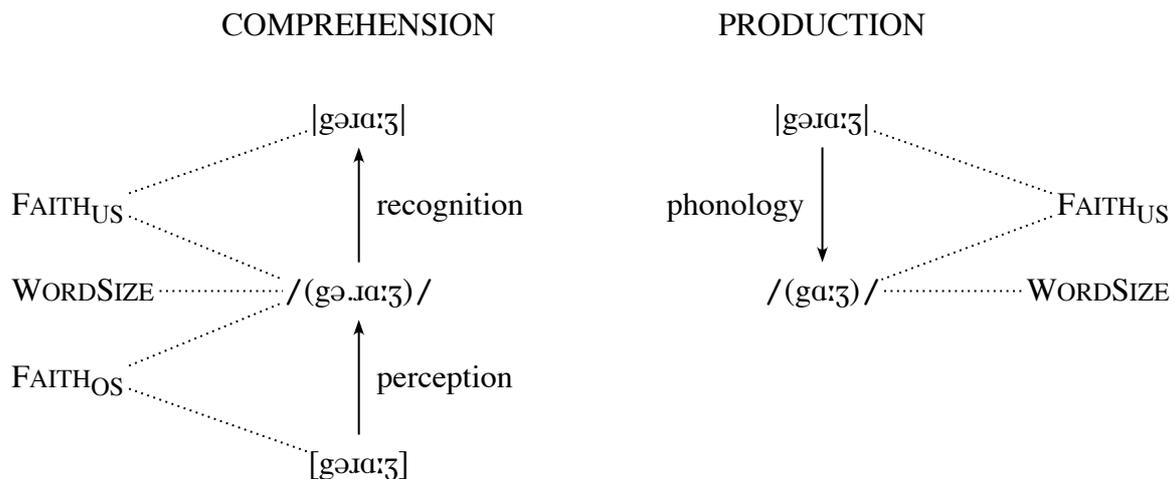
Pater (2004) uses this model for explaining that children's perception is not always flawless. If the ranking is STRUCT >> { FAITH<sub>US</sub>, FAITH<sub>OS</sub> }, perception (and therefore production) is non-adultlike, as in the following example of learning to perceive the English word *garage*:

(18) *Pater's example of non-adultlike perception and production*



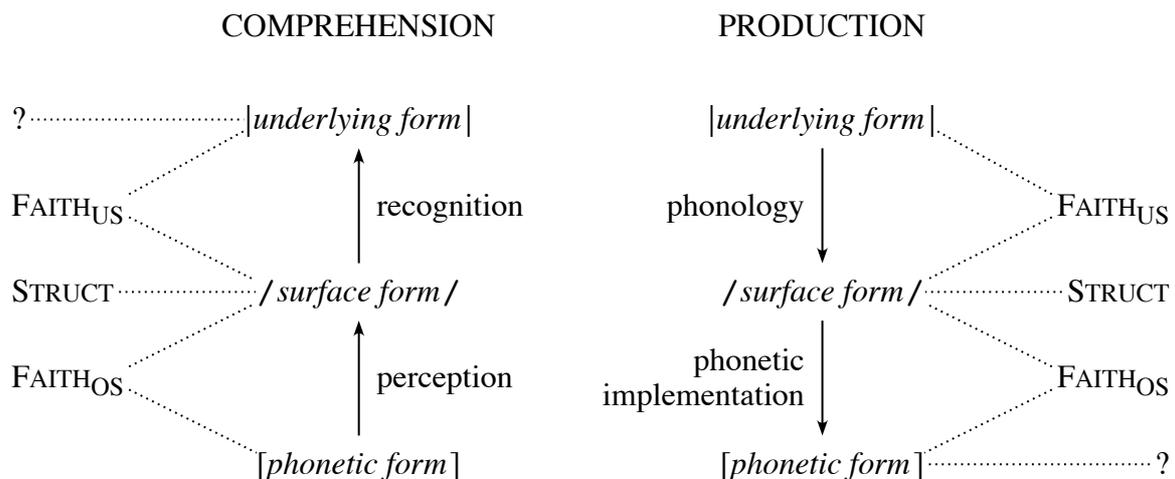
In this example, WORDSIZE is a constraint against polysyllabic words. In the early stage shown here, it limits the output of perception (as in Tesar and Smolensky's model). If in a later stage FAITH<sub>OS</sub> has risen above WORDSIZE, both perception and lexical storage will be fine, but production is still not adultlike:

(19) *Pater's example of adultlike perception and non-adultlike production*



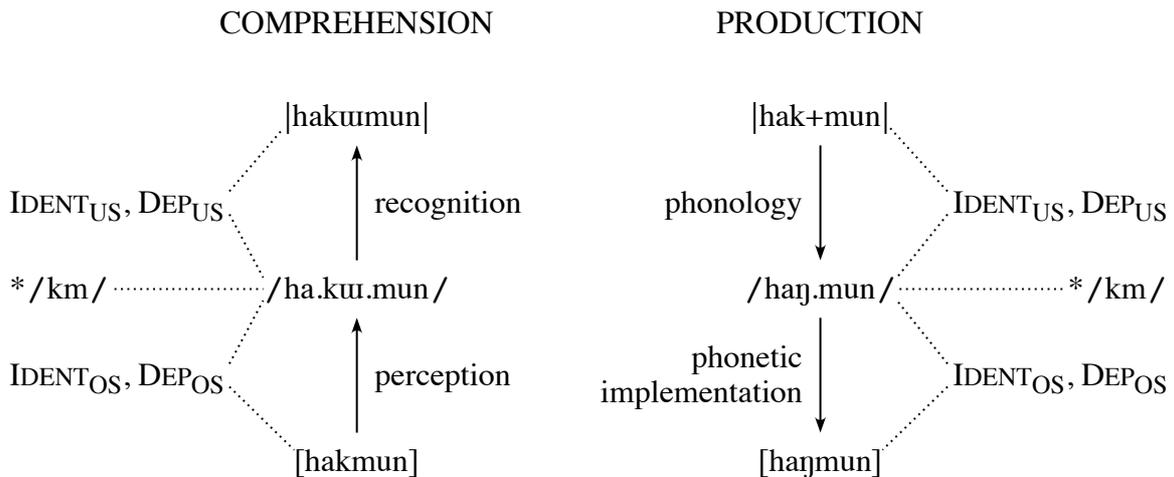
It is straightforward to extend such a model with a phonetic implementation module, bringing it closer to Hayes' footnote and Levelt's modularity. In the most general case, OT constraints would be able to evaluate the results of recognition and phonetic implementation. We would end up with something like (20).

(20) *The Serial Bidirectional Three-Representation Model*



Since as far as we know nobody has ever proposed this model we will simply call it the Serial Bidirectional Three-Representation Model. It can be regarded as Tesar and Smolensky without containment, or as an OT implementation of Cutler and Levelt, or as Hayes' footnote made bidirectional. As an example of the usefulness of this model, consider the Korean example employed by Kabak and Idsardi (2007) (somewhat simplified here) of an underlying [hak + mun] that is produced as /haŋ.mun/, and a phonetic [hakmun] that is perceived by Korean listeners as /ha.ku.mun/. This example seems to be handled rather well by the Standard 3-Rep Model, as shown in (21).

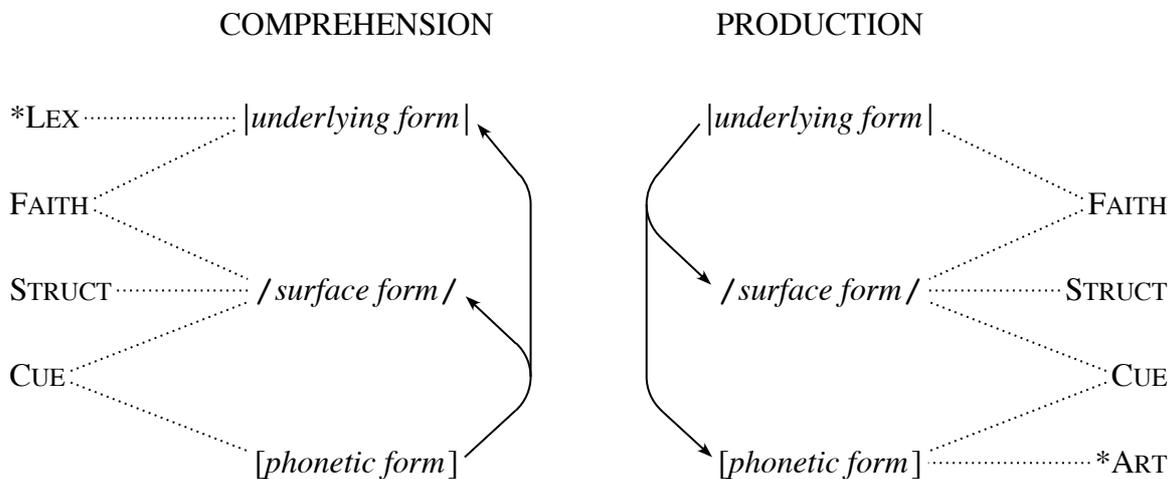
(21) *The Serial Bidirectional Three-rep Model for two different repairs in Korean*



Both in perception and in production, the structural constraint  $*/km/$  prevents the occurrence of the phonological surface form  $*/hak.mun/$ , but in different ways. In production this is possibly caused by the ranking  $\{ DEP_{US}(V), */km/ \} \gg IDENT_{US}(nas)$ , in perception possibly by the ranking  $\{ IDENT_{OS}(nas), */km/ \} \gg DEP_{OS}(V)$ .

While the model in (20) handles each direction of processing as a sequence of two modules, where the output of one module is the input to the other module, it is also possible, and even more in the spirit of Optimality Theory, to perform the two mappings in parallel. Such a Parallel Bidirectional Three-Representation model is shown in (22).

(22) *The Parallel Bidirectional Three-Representation Model*



In this model the structure of the grammar, namely the representations and the constraints, is identical to that in the bidirectional grammar model with control loop in (14). However, the two models differ in the way the processes of comprehension and production are defined on it. For comprehension, the prelexical perception stage and the word recognition stage are evaluated in parallel, so that structural and cue constraints can interact with lexical and faithfulness constraints; as Boersma (this volume) shows, this can straightforwardly account for several observed phenomena, such as lexical influences on phoneme identification (Ganong 1980) and phonemic restoration (Samuel 1981). For production, the abstract phonology and the concrete phonetics are evaluated in parallel, so that articulatory and cue

constraints can interact with structural and faithfulness constraints; this can account for phenomena such as phonetic enhancement (as has been shown by Boersma and Hamann 2008), licensing by cue (as has been shown by Boersma 2008), and incomplete neutralization.

## 6. Models with more than three representations

Some models have been proposed that include more than the three representations discussed in §5. Boersma (1998) divided the phonetic representation into an articulatory and an auditory form, where the articulatory constraints evaluate the articulatory form, the cue constraints evaluate the relation between the auditory form and the phonological surface form, and the articulatory and auditory forms are related by sensorimotor constraints (Boersma 2006). Likewise, Apoussidou (2007) divided the underlying representation into a phonological underlying form and the meaning of the morpheme, where lexical constraints evaluate the relation between these two representations; Apoussidou showed that this split enables us to model the acquisition of abstract lexical representations. In this volume, the morphemic meaning representations of ‘Focus’ and ‘Topic’ are used by **Féry, Keyser, Hörnig, Weskott and Kliegl** in their OT modelling of intonation contours; these authors also discuss a phonetic form and a surface phonological form, which is directly connected to underlying meaning by association constraints without the intervention of any phonological underlying form.

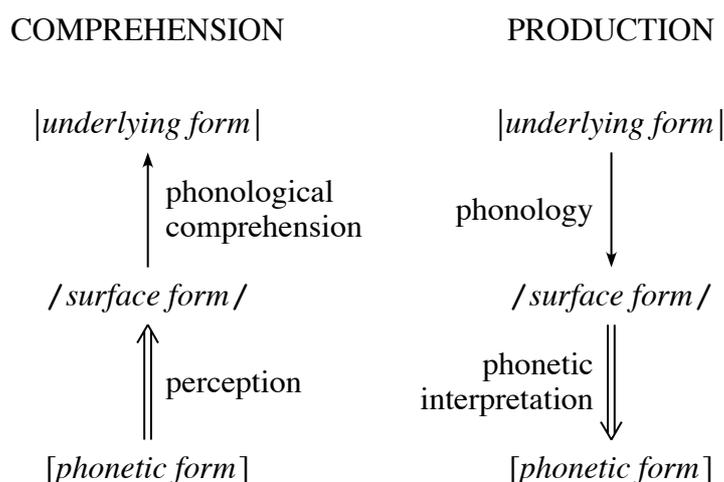
Once the number of representations can be raised to four or five like this, it becomes imaginable that a realistic OT model of language processing may contain a large number of representations, and that using a single phonetic form, or a single lexical form, or a single phonological form, are all just simplifications. In fact, one can imagine that all the phonological and phonetic forms are connected to a multitude of semantic and syntactic representations, probably via the morphemic meaning mentioned above. An attempt to model this is provided by González (2006), who considers six levels of representation when modelling language mixing in bilinguals.

By providing models for language processing, multi-level OT joins the playing ground of psycholinguistic research. While some observed psycholinguistic phenomena can be replicated by OT models (as mentioned before), some others still pose challenges. An example of this are the *lexical neighbourhood density* phenomena discussed in this volume by **Ussishkin and Wedel**. At first sight, these phenomena seem to require an explanation that involves connections between elements *within* a level of representation, whereas the multi-level OT models discussed above only proposed connections *between* levels. A future solution may be based on realizing that even in these OT models relations within a level automatically exist as a result of bidirectional connections with an adjacent level.

## 7. Perception as extralinguistic

If one regards perception and phonetic implementation as universal extralinguistic processes, the grammar model could be the one in (23).

(23) Grammar model with extralinguistic perception and phonetic interpretation



This model is compatible with McCarthy and Prince's two-representation model in (6), in which everything to do with overt phonetic forms is outside phonology proper. It seems to be the model *assumed* (not necessarily *defended*) by most phonologists, including people who study the interaction 'between' perception and phonology, such as Steriade (1995, 2001) and Hume and Johnson (2001), and many researchers working on loanword phonology (see e.g. LaCharité and Paradis 2005 and Uffmann 2006). Interestingly, the model in (23) seems to be assumed by two opposing schools in loanword research, namely those that think that loanword adaptation is in *phonology* (in the strict sense, i.e. the 'phonology' in the production part of the picture) and those who think that loanword adaptation is in *perception* (for both schools, the 'phonological comprehension' module usually simply stores the perceived surface form in the lexicon unaltered).

Loanword adaptation does not necessarily have to be perception *or* phonology. In fact, the models in (12), (14), (17), (20), and (22) predict that both perception and production are capable of imposing restrictions, and that structural (i.e. phonological) constraints, such as language-specific phonotactics, already determine the perceived surface form. This is in line with the observation by Polivanov (1931) and Dupoux, Kakehi, Hirose, Pallier and Mehler (1999) that Japanese listeners perceive [tak] and [ebzo] as /ta.ku/ and /e.bu.zo/, following Japanese-specific phonotactics (see Boersma this volume for details). In the present book, **Broselow** defends the viewpoint that loanword adaptation is both phonology *and* perception, with a language-specific perception grammar.<sup>4</sup> The contribution by **Balas** also deals with loanword adaptation and puts forward the view that the influence of language-specific perception can be better captured in Natural Phonology than in a two-level representational OT approach.

## 8. Conclusion

This book brings together for the first time a number of contributions from different areas of phonology that model the phonological comprehension process by explicit linguistic means. The book contains chapters that provide overviews, case studies, applications to acquisition

---

<sup>4</sup> The same point was made by Kenstowicz (2001), Broselow (2004), and Yip (2006). However, they worked within a model with only two levels of representation (§4), so that they had to propose different constraints for comprehension and production. Boersma and Hamann (to appear) work with the three levels of representation discussed in this Introduction, and therefore with the same constraints in both directions of processing.

and change, experimental evidence, and remaining challenges. The book ends with a commentary by **McClelland**, who views the book from the standpoint of cognitive science.

## References

- Apoussidou, Diana 2007 *The Learnability of Metrical Phonology*. Ph.D. dissertation, University of Amsterdam.
- Boersma, Paul 1998 *Functional phonology*. Ph.D. dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.
- Boersma, Paul 2000 The OCP in the perception grammar. *Rutgers Optimality Archive* 435.
- Boersma, Paul 2001 Phonology-semantics interaction in OT, and its acquisition. In: Robert Kirchner, Wolf Wikeley and Joe Pater (eds.), *Papers in Experimental and Theoretical Linguistics*. Volume 6, 24–35. Edmonton: University of Alberta.
- Boersma, Paul 2003 Overt forms and the control of comprehension. In: Jennifer Spenader, Anders Eriksson and Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 47–56. Department of Linguistics, Stockholm University.
- Boersma, Paul 2006 Prototypicality judgments as inverted perception. In: Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky and Ralf Vogel (eds.), *Gradience in Grammar*, 167–184. Oxford: Oxford University Press.
- Boersma, Paul 2007 Some listener-oriented accounts of hache aspiré in French. *Lingua* 117: 1989–2054.
- Boersma, Paul 2008 Emergent ranking of faithfulness explains markedness and licensing by cue. In: *Rutgers Optimality Archive* 954.
- Boersma, Paul, and Silke Hamann 2008 The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.
- Boersma, Paul, and Silke Hamann to appear Loanword adaptation as first-language phonological perception. To appear in: Andrea Calabrese and W. Leo Wetzels, *Loanword phonology*. Amsterdam: John Benjamins. [*Rutgers Optimality Archive* 975]
- Broselow, Ellen 2004 Language contact phonology: richness of the stimulus, poverty of the base. *North-Eastern Linguistic Society* 34: 1–22.
- Chomsky, Noam, and Morris Halle 1968 *The Sound Pattern of English*. New York: Harper and Row.
- Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier and Jacques Mehler 1999 Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568–1578.
- Escudero, Paola 2005 *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. Ph.D. dissertation, University Utrecht.
- Escudero, Paola, and Paul Boersma 2003 Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In: Sudha Arunachalam, Elsi Kaiser and Alexander Williams (eds.), *Proceedings of the 25th Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics* 8: 71–85.
- Escudero, Paola, and Paul Boersma 2004 Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551–585.
- Ganong, William F. 1980 Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6: 110–125.
- González, Maria Angélica 2006 *Inhibition and Activation in Language Mixing by Bilinguals*. M.A. thesis, University of Amsterdam.

- Hale, Mark, and Charles Reiss 1998 Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29: 656–683.
- Hayes, Bruce 1999 Phonetically-driven phonology: the role of Optimality Theory and Inductive Grounding. In: Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer and Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics*, Volume I: *General papers*, 243–285. Amsterdam: John Benjamins.
- Hayes, Rachel 2001a *An Optimality-Theoretic Account of Novel Phonetic Category Formation in Second Language Learning*. Manuscript, University of Arizona.
- Hayes, Rachel 2001b The perception of novel phoneme contrasts in a second language: A development study of native speakers of English learning Japanese singleton and geminate consonant contrasts. In: Rachel Hayes, William D. Lewis, Erin L. O’Byrne and Tania S. Zamuner (eds.), *Language in Cognitive Science*, (University of Arizona Coyote Working Papers 12.) 28–41.
- Hume, Elizabeth, and Keith Johnson 2001 A model of the interplay of speech perception and phonology. In: Elizabeth Hume and Keith Johnson (eds.), *The Role of Speech Perception in Phonology*, 3–26. New York: Academic Press.
- Kabak, Barış, and William Idsardi 2007 Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50: 23–52.
- Kenstowicz, Michael 2001 The role of perception in loanword phonology. A review of *Les emprunts linguistiques d’origine européenne en Fon* by Flavien Gbéto. To appear in *Linguistique Africaine*.
- LaCharité, Darlene, and Carole Paradis 2005 Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36: 223–258.
- Levelt, Willem 1989 *Speaking: From Intention to Articulation*. Cambridge, Mass: MIT Press.
- McCarthy, John, and Alan Prince 1995 Faithfulness and reduplicative identity. In: Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, 249–384. (University of Massachusetts Occasional Papers 18.) Amherst, Mass.: Graduate Linguistic Student Association.
- McQueen, James M., and Anne Cutler 1997 Cognitive processes in speech perception. In: William J. Hardcastle and John Laver (eds.), *The Handbook of Phonetic Sciences*, 566–585. Oxford: Blackwell.
- Pater, Joe 1999 From phonological typology to the development of receptive and productive phonological competence: Applications of minimal violation. *Rutgers Optimality Archive* 296.
- Pater, Joe 2004 Bridging the gap between perception and production with minimally violable constraints. In: René Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 219–244. Cambridge: Cambridge University Press.
- Prince, Alan, and Paul Smolensky 1993 Optimality Theory: constraint interaction in generative grammar. Technical Report TR-2, Rutgers University Center for Cognitive Science. [published in 2004 by Blackwell, Malden]
- Polivanov, Evgenij Dmitrievič 1931 La perception des sons d’une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4: 79–96.
- Samuel, Arthur G. 1981 Phonemic restoration: insights from a new methodology. *Journal of Experimental Psychology: General* 110: 474–494.
- Smolensky, Paul 1996 On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720–731.
- Steriade, Donca 1995 *Positional neutralization*. Unfinished manuscript, UCLA.
- Steriade, Donca 2001 Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume and Keith Johnson (eds.), *The Role of Speech Perception in Phonology*, 219–250. New York: Academic Press.

- Tesar, Bruce 1997 An iterative strategy for learning metrical stress in Optimality Theory. In: Elizabeth Hughes, Mary Hughes and Annabel Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 615–626. Somerville, Mass.: Cascadia.
- Tesar, Bruce 1998 An iterative strategy for language learning. *Lingua* 104: 131–145.
- Tesar, Bruce 1999 Robust interpretive parsing in metrical stress theory. In: Kimary Shahin, Susan Blake and Eun-Sook Kim (eds.), *Proceedings of the 17th West Coast Conference on Formal Linguistics*, 625–639. Stanford, Calif.: CSLI.
- Tesar, Bruce 2000 On the roles of optimality and strict domination in language learning. In: Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*, 592–620. New York: Oxford University Press.
- Tesar, Bruce, and Paul Smolensky 1998 Learnability in Optimality Theory. *Linguistic Inquiry* 29: 229–268.
- Tesar, Bruce, and Paul Smolensky 2000 *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.
- Uffmann, Christian 2006 Epenthetic vowel quality in loanwords: empirical and formal issues. *Lingua* 116: 1079–1111.
- Yip, Moira 2006 The symbiosis between perception and grammar in loanword phonology. *Lingua* 116: 950–975.