Perception in Optimality Theory: The Frugality of the Base
Author(s): Paul Lassettre and Patricia Donegan
Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998), pp. 346-355

Perception in Optimality Theory: The Frugality of the Base

Paul Lassettre & Patricia Donegan
University of Hawai'i at Manoa

A complete theory of phonology must model perception as well as production. Children's accurate phonological perceptions and inaccurate phonological productions would seem to pose a problem for Optimality Theory (OT, Prince & Smolensky 1993), where we expect a single grammar for both, yet Smolensky (1996) gives an account of the asymmetry. In his model, structural well-formedness and faithfulness constraints evaluate surface representations, but only faithfulness constraints evaluate underlying representations. However, the OT account of perception seems to predict that adult phonological perception will be highly accurate – that adults will perceive the surface representation. Any adult who has ever studied an unfamiliar language knows that this is not true: distinctions that are perfectly obvious to speakers of one language can seem impossibly *subtle* to others. So, although there is much to recommend Smolensky's proposal, we believe it requires some modification. In our model of perception, not only the faithfulness constraints but also a certain class of structural well-formedness constraints evaluate underlying representations. The result is underlying representations that look very similar to traditional phonemic representations.

1. An OT model of perception.

Smolensky (1996) presents a model that accounts for the widely-noted discrepancy between linguistic production and perception in young children, who are able to hear (and remember) differences between sounds that they cannot pronounce distinctly. The model uses a single grammar for both production and perception, and it builds on previous work such as Prince & Smolensky (1993: 175-196) and Gnanadesikan (1995). It makes use of two representations, UR and SR. Production is modeled as mapping a given UR onto an unknown SR, and perception is modeled as mapping a given SR onto an unknown UR. The discrepancy between children's production and perception corresponds to a mapping that is more accurate from SR to UR than from UR to SR.

An OT grammar evaluates candidates for whichever representation is to be discovered. Two types of constraints are used: STRUCTURAL CONSTRAINTS (SCs) and FAITHFULNESS CONSTRAINTS (FCs). SCs evaluate the well-formedness of SRs. FCs evaluate URs and SRs with respect to each other; cases where UR and SR differ are less optimal than where UR and SR are alike. In the OT model, the evaluation function does not treat UR and SR equally: UR is evaluated only in terms of its faithfulness to SR, while SR is evaluated with respect to both its faithfulness to UR and its inherent well-formedness. In the early stages of acquisition, SCs are assumed to outrank FCs (Tesar & Smolensky 1996). The unequal

ranking of the two types of constraints and their inequality in evaluating the two representations account for the discrepancy between perception and production.

Tableau 1 shows the grammar of a child who pronounces *fish* as [fɪs] but rejects that pronunciation by adults; the same child pronounces *kiss* as [kɪs] and accepts that pronunciation by adults. This is our interpretation of Smolensky's model, using our own notation: candidates are shown as ordered pairs, in which the known member of the pair is fixed and the unknown varies freely (although we show only two variants). Faithfulness is UR-SR correspondence (McCarthy & Prince 1994, McCarthy 1995): MAX /feature/ is violated if the specified feature in UR lacks an identical correspondent in SR, and DEP [feature] is violated if the specified feature in SR lacks an identical correspondent in UR.

Perception: given [fiʃ], [kɪs]

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
a.</fɪs/, [fiʃ]>	*		*!	☞c.</kɪs/, [kɪs]>			
☞b.</fiʃ/, [fiʃ]>	*			d.</kɪʃ/, [kɪs]>		*!	

Production: given /fiʃ/, /kɪs/

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
☞e.</fiʃ/,[fɪs]>		*		☞g.</kɪs/,[kɪs]>			
f.</fiʃ/, [fiʃ]>	*!			h.</kɪs/, [kɪʃ]>	*!		*

Tableau 1. After Smolensky: accurate perception, inaccurate production. *C[+pal]: Consonants are non-palatal. MAX /+pal/: Palatality in UR corresponds to palatality in SR. DEP[+pal]: Palatality in SR corresponds to palatality in UR.

The production tableaux show cases where the URs /fiʃ/ and /kɪs/ are the given information. Candidates are ordered pairs of UR and SR in which the SR varies. The candidate set contains all possible SRs but just one UR. The highest ranked of the three constraints shown is a structural constraint, which evaluates the SR. The harmonic ordering of the candidate set is therefore affected by the well-formedness of the SR, and the winning candidates are [kɪs] and the unfaithful but well-formed [fɪs].

In the perception tableaux, the two SRs [fiʃ] and [kɪs] are given. As above, the candidates are ordered pairs, but now it is the SR that is constant. The candidate set includes all possible URs but just one SR. All SRs incur equal violations of the structural constraints, and therefore the well-formedness of the SR does not affect the harmonic ordering of the candidate set. FCs, although ranked below SCs, are the only constraints that affect the harmonic ordering and the winning candidate is the one where the UR and the SR are most alike.

The child's ability to hear the difference between [ʃ] and [s] is reflected in the model by the faithful mapping from [fiʃ] and [kɪs] to /fiʃ/ and /kɪs/. The same child's inability to pronounce /ʃ/ distinctly from /s/ is shown by the same grammar's mapping /fiʃ/ onto [fɪs].

2. Perception and richness of the base.

'Richness of the base' is the phrase by which Prince & Smolensky (1993: 191) refer to the basic assumption that the candidate set is unlimited. The implication is that systematic gaps and omissions in a speaker's outputs cannot be explained merely by gaps and omissions in the input, but must be explained by the phonology itself (in OT, this is accomplished by the evaluation function). Gaps and omissions in a language's lexicon reflect the surface patterns, not the other way around.[1] 'Richness of the base' is a reiteration of the claim that the 'locus of explanatory action' (Prince & Smolensky 1993:3) is, in OT, the ranked constraint hierarchy.

For example, Standard Hawaiian has no closed syllables. The explanation for this is not that Standard Hawaiian has no underlying strings that could be parsed into closed syllables. Rather, Standard Hawaiian phonology is such that every possible underlying string, including strings that contain such things as consonant clusters and final Cs, will be parsed into surface open syllables. The explanation lies in the constraint ranking, not in the inputs. Any lack of underlying consonant clusters and final consonants is a projection of the surface pattern.

As a further example, Korean has no surface contrast between [l] and [r]. The distribution of these liquids is predictable: [l] occurs syllable finally and [r] occurs elsewhere. The phonology of Korean disallows [l] and [r] in complementary surface structure environments. The lack of an underlying contrast between the two is due to their surface neutralization.

However, as Smolensky (1996) notes, the neutralization of Korean [l] and [r] is not just neutralization in production. Korean speakers also neutralize [l] and [r] in perception, tending to perceive English *bear* and *bail*, for example, as homophones.[2] The OT model with faithful mapping from SR to UR predicts that these sounds will be perceptually distinct for Korean speakers as [ʃ] and [s] are for the child acquiring English.

Smolensky proposes (in a footnote), that this effect is due to lexical retrieval: in addition to URs and SRs, there are underspecified LEXICAL REPRESENTATIONS (LRs). Features that are not used distinctively in a language are absent from the LR. In lexical retrieval, a UR is matched to an LR. Because of underspecification, some distinct URs may match the same LR. Failure to perceive a distinction is reflected in this model as matching a single LR. Smolensky suggests this is the case with [l] and [r] in Korean: Korean speakers hear *bear* and *bail* as homophones because Korean speakers cannot match them with distinct LRs (Tableau 2).

	Struct	Faith		Struct	Faith
a.</ber/, [ber]>	*		☞c.</ber/, [bel]>	*	*!
☞b.</bel/, [ber]>	*	*!	d.</bel/, [bel]>	*	

Lexical retrieval: /ber/ => lbeLl, /bel/ => lbeLl

Tableau 2. Korean: Faithful URs, homophony at lexical retrieval.

This model maintains the principle of 'richness of the base', but it does so indirectly. Perception of homophony (gaps and omissions in the perceptual pattern) is explained by gaps and omissions (underspecified features) in the LR. But gaps and omissions in the LR are ultimately explained in terms of the constraint hierarchy, by way of patterns in SRs.

3. Nasal vowels in American English.

A model of perception that depends on underspecified LRs does not account for all perceptual phenomena. The model described above predicts, for example, that the LRs of many American English dialects would have to include nasalized vowels. But speakers of such dialects do not perceive the large number of contrasts entailed by this analysis, or at least not more reliably than Korean speakers distinguish [r] and [l].

American English speakers ordinarily pronounce *say, sate, sane*, and *saint* as [sei̯], [sei̯t], [sẽi̯n], and [sẽi̯t]. There does not exist any fifth lexical item ordinarily pronounced [sẽi̯], [sei̯n], [sei̯nt], or [sẽi̯nt] that is not homophonous with *say, sate, sane*, or *saint*. No such lexical item is possible. The phonology determines how nasalized vowels are distributed on the surface, by mapping the eight logically possible URs, /sei̯/, /sei̯t/, /sei̯n/, /sei̯nt/, /sẽi̯/, /sẽi̯t/, /sẽi̯n/, and /sẽi̯nt/ onto the four attested SRs, [sei̯], [sei̯t], [sẽi̯n], and [sẽi̯t].

Some SRs differ only by vowel nasalization, but are not neutralized in perception: *sate* [sei̯t] and *saint* [sẽi̯t], *debt* [dɛt] and *dent* [dɛ̃t], *lip* [lɪp] and *limp* [lɪ̃p], *duck* [dʌk] and *dunk* [dʌ̃k], and so on. URs are faithful to SRs, and LRs are not distinct from URs. Vowel nasalization is specified in LR or it is underspecified; there is no other option. If vowel nasalization were underspecified in LR, *debt* and *dent* and similar pairs would be perceived as homophones (Tableau 3). Therefore, vowel nasalization must be specified in LR to account for these minimal pairs.

	*V N [-nas]	*NT	DEP [+nas]	*$V_{[+nas]}$	*MAX /C/	MAX /+nas/
☞a.</dɛt/, [dɛt]>						
b.</dɛ̃t/, [dɛt]>						*!
c.</dɛt/, [dɛ̃t]>			*!	*		
☞d.</dɛ̃t/, [dɛ̃t]>				*		
e.</dɛ̃nt/, [dɛ̃t]>				*	*!	
f.</dɛnt/, [dɛ̃t]>				*	*!	

Lexical retrieval: /dɛt/ => $|dɛ_{[Ønas]}t|$ /dɛ̃t/ => $|dɛ_{[Ønas]}t|$

Tableau 3. English perception 1: [dɛt], [dɛ̃t] as homophones, with accurate URs, underspecified LRs. *V[-nas]N: Non-nasal vowels do not occur before nasals. *NT: Nasal consonants do not occur before homorganic voiceless stops. DEP [+nas]: Nasality in SR must be attributable to nasality in UR. *V[+nas]: Nasalized vowels do not occur. MAX /C/: A consonant in UR corresponds to a consonant in SR. MAX /+nas/: Nasality in UR corresponds to nasality in SR.

But if this OT model were correct, American English speakers would reliably distinguish the eight SRs, [seɻ], [seɻt], [seɻn], [seɻnt], [sẽɻ], [sẽɻt], [sẽɻn], and [sẽɻnt], whether or not they can produce them. We do not believe American English speakers can distinguish all eight any more reliably than Korean speakers can distinguish pairs like *bear* and *bail*. No scheme of lexical underspecification will condition a realistic pattern of perceptual mergers and distinctions.

The standard OT model predicts that adult perception will be highly accurate, more accurate than adult production. Smolensky's model, with its underspecified LRs, attempts to account for cases where this prediction is incorrect. But the underspecified LR model cannot account for more complicated examples of the same type, such as the American English perception of vowel nasalization, so it does not solve the problem.

4. Systematically unfaithful URs.

A modification to the way some structural constraints evaluate representations can resolve the nasalization dilemma. We recognize two classes of structural constraints, which can be distinguished formally and functionally. One class of constraints should evaluate URs as well as SRs, while the other should continue to evaluate only SRs.

One class we call SEGMENTAL WELL-FORMEDNESS (SEGWF) constraints, although we might equally well term them Fortitive or Enhancement or Paradigmatic constraints (cf. Stevens & Keyser 1989, Pulleyblank 1997). Formally, these constraints refer to the occurrence or simultaneous co-occurrence of phonological features. (1) shows our notation (cf. Lassettre 1995).[3]

(1) a. *[feature] b. * feature_i feature_k

Empirical evidence for SEGWF constraints comes from context-free phonological phenomena, including crosslinguistic patterning of phoneme inventories, unconditioned historical sound changes, and stylistic variation observed in various kinds of hyperarticulated speech. SEGWF constraints have a functional basis in perception, though they may have articulatory motivation, too.

The other class we call SEQUENTIAL WELL-FORMEDNESS (SEQWF) constraints, although they might be called Lenitive constraints. Formally, these constraints refer to sequences of features or feature combinations (2). Empirical evidence for these constraints comes from context-sensitive phonological phenomena, such as assimilation and reduction, and 'conditioned' historical sound changes. SEQWF constraints, as their name implies, have a functional basis in the articulation of sequences, typically reducing the number or magnitude of differences between gestures that follow one another.[4]

(2) *[feature_i] [feature_k]

SEGWF constraints evaluate URs, while SEQWF constraints evaluate SRs only. The result is that in perception URs are not fully faithful to given SRs, although the mappings from SR to UR and from UR to SR are still asymmetrical. Perception of contrast is modeled as mapping of distinct SRs onto distinct URs. Perception of homophony, however, is modeled as mapping of distinct SRs directly onto the same UR.

Tableau 4 shows how this resolves the English vowel nasalization problem. *$V_{[-nas]}N$ and *NT are SEQWF constraints and do not evaluate in perception. The only SEGWF is *$V_{[+nas]}$, but its application in perception is sufficient to change the optimal UR for [dɛ̃t] from /dɛ̃t/ to /dɛnt/, given the ranking DEP [+nas], *$V_{[+nas]}$ >> MAX /C/. Thus [dɛt] and [dɛ̃t] are associated with distinct URs, without allowing underlying /ṽ/ in English. This constraint ranking will lead to all surface occurrences of stressed [ṽ] being perceived as /VN/ by English speakers.[5]

	*V N [-nas]	*NT	DEP [+nas]	*$V_{[+nas]}$	*MAX /C/	MAX /+nas/
☞a.</dɛt/, [dɛt]>						
b.</dɛ̃t/, [dɛt]>						*!
☞ c.</dɛt/, [dɛ̃t]>			*!	*		
d.</dɛ̃t/, [dɛ̃t]>				**!		

Tableau 4. English perception 2: [dɛt] as /dɛt/, [dɛ̃t] as /dɛnt/, with unfaithful URs.

In fact, not all surface occurrences of [ṽ] have to be perceived as /VN/. Because of another sequential constraint, *$NV_{[-nas]}$, which dominates *$V_{[+nas]}$, unstressed syllables of the form VNT and VT are homophonous after a nasal consonant, as in *approximate* (adj.) and *approximant*, both of which are usually [əˈprɑksɪ̃mɪ̃t]. Tableau 5 suggests that the nasalized vowel here will be perceived, other things being equal (i.e., without certain morphological or orthographic knowledge – cf. Stampe 1987) as a nonnasalized vowel without a following nasal consonant. Its nasality can be attributed to the preceding nasal. *$V_{[+nas]}$, Max /C/, and Max /+nas/ are unviolated. (See Tableau 5.)

	*V N [-nas]	*N V [-nas]	*NT	Dep [+nas]	*V [+nas]	Max /C/	Max /+nas/
☞<-/mət/, -[mɪ̃t]>							
<-/mənt/, -[mɪ̃t]>						*	

Tableau 5: English perception of unstressed -[mɪ̃t].

Perception of the final syllable as including a postvocalic nasal (as /VNT/) requires that Max /C/ be violated. This is of course possible, since Max /C/ may be ranked below morphological or orthographic constraints. Spelling errors like *dominate* (adj.) for *dominant* and *pregnate* or *pregnit* for *pregnant* (but, we

predict, not *litigate* for *litigant*) give evidence that this (mis)perception does occur. Thus, the revised model is capable of handling production and perception of nasalized vowels in various environments in English simply.

The Korean case is resolved in the same way as the English one. *Bear* [ber] and *bail* [bel] are mapped directly onto the same UR (Tableau 6). As in Smolensky's model, perception of homophony is mapping onto the same representation. Unlike that model, however, this model derives that mapping directly from an optimality- theoretic constraint hierarchy evaluating a candidate set. Korean phonology neutralizes liquids in both perception and production.

	Structure	SEQWF	SEGWF
a.</ber/, [ber]>	*	**!	
☞b.</bel/, [ber]>	*	*	
c.</ber/, [bel]>	*	*!	
☞d.</bel/, [bel]>	*		

Tableau 6. Korean perception: homophony at UR.

5. On children's perceptions.

Returning to the question with which we started, that of children's perceptions, we note that the revised model allows the possibility that children learning English may at first be unable to perceive the difference between [ʃ] and [s]. If structural constraints outrank faithfulness constraints initially, children will need to rerank constraints in order to perceive the distinction. Now, there is little if any evidence of systematic misperceptions (like perceptual merger of [ʃ] and [s]) in children. It seems, rather, that children are able to perceive virtually all of the phonetic distinctions used in languages from a very early age (Eimas et al. 1971, Trehub 1976, etc.). This ability changes somewhat during the babbling period, but at the onset of speech children still appear to be able to perceive all the phonemic distinctions of the ambient language (Werker & Pegg 1992, cf. Smith 1973.).

Yet it is not necessary to rule out the possibility of perceptual mergers in the early period of acquisition, especially since data on linguistic perception at this period is notoriously hard to come by. In Tableau 6, examples a. - d. show a possible (though unattested) stage of acquisition where the dominant SEGWF constraint $*C_{[+pal]}$ determines perception. Examples e. - h. in the same tableau show a later stage, where the faithfulness constraint, DEP[+pal], determines perception; note that at this point the faithfulness constraint MAX /+pal/ is still outranked by $*C_{[+pal]}$, so $*C_{[+pal]}$ determines the child's productions.

The model presented here is a model of linguistic perception, but we assume that both representations, SR as well as UR, are percepts. The surface representation has already undergone significant cognitive processing, and it is not a direct representation of the acoustic wave. Rather, it is a representation of what is heard. It may be that the youngest children studied perceive all possible

phonetic distinctions because they treat linguistic stimuli simply as sounds, so that there would be at this point no difference between 'speech perception' and non-speech auditory perception. This would mean that the youngest infants respond simply to differences in SR.

Earlier Perception: given [fiʃ], [kɪs]

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
a.</fɪs/, [fiʃ]>	*		*!	☞c.</kɪs/, [kɪs]>			
☞b.</fiʃ/, [fiʃ]>	**!			d.</kɪʃ/, [kɪs]>	*!	*	

Later Perception: given [fiʃ], [kɪs]

	DEP [+pal]	*C [+pal]	MAX /+pal/		DEP [+pal]	*C [+pal]	MAX /+pal/
☞e.</fɪs/, [fiʃ]>	*!	*		☞g.</kɪs/, [kɪs]>			
f.</fiʃ/, [fiʃ]>		**		h.</kɪʃ/, [kɪs]>		*!	*

Tableau 7. Earlier and later linguistic perception.

Only when the experience of using their own vocal tracts causes the constraints on their phonetic abilities to become part of their mental processing of speech would 'speech perception' in children begin to be different from non-speech auditory perception. [6] When the child begins to perceive in terms of a smaller set of categories, limited by SEGWF constraints, which are presumed to characterize the speaker's intentions – then we can say that the child begins to perceive in terms of URs. This change does not involve a loss of auditory sensitivity or ability, of course: adults may be enabled to hear in terms of SRs under certain conditions (Werker & Tees 1984, Best et al. 1988). Under such conditions, Korean speakers may become able to judge that *bear* and *bail* do not sound exactly alike, even though they may remain unable to apprehend the difference linguistically.

6. Variability across languages.

Gaps, omissions, and other patterns in URs are not mere projections of surface patterns. Differences among languages with respect to URs, as well as SRs, can be attributed to different constraint rankings. For example, consider the constraints:

*VN	Nonnasal vowels do not occur before nasal consonants (SEQWF).
$*V_{[+nas]}$	Nasalized vowels do not occur (SEGWF).
MAX /nas/	Nasality in UR corresponds to nasality in SR.
DEP [nas]	Nasality in SR corresponds to nasality in UR.

These constraints, ranked differently, account for vowel nasalization in different languages. In English: *VN >> $*V_{[+nas]}$ >> DEP[nas]. As a result, nasalized vowels occur before nasal consonants in SRs, but the SEGWF constraint, $*V_{[+nas]}$, ensures that vowels in UR are all nonnasal. In French, the faithfulness

constraints MAX/nas/ and DEP[nas] >> *VN and *$V_{[+nas]}$. With the faithfulness constraints most highly ranked, both nasalized and nonnasalized vowels occur in URs, and they appear as nasalized and nonnasalized vowels, respectively, in SRs. In Hindi, *VN >> DEP[nas] >> *$V_{[+nas]}$, MAX /nas/. As a result of this ranking, both nasalized and nonnasalized vowels appear in URs, and some nonnasalized UR vowels are nasalized in SRs.

7. Conclusion.

The input to perception must be just as unlimited as the input to production: 'richness of the base' is bidirectional. Gaps and omissions in perception must be explained by the phonological system itself, as the result of the interaction of faithfulness and perceptually motivated structural constraints.

Notes.

[1] The phrase is new, but the concept is not (e.g., Andersen 1973, Stampe 1987). In a phonological theory which attempts to encompass phonological acquisition and change, things could hardly be otherwise.

[2] The Korean difficulty applies also to *bear* and *bell*, an alternative minimal pair.

[3] They may be expressed as constraints on licensing or feature geometry.

[4] Our definition of the simultaneous-sequential distinction is an approximation. We have not addressed prosodically defined positional limitations of occurrence, like *[-son +voi] in final position (German) or *[-cont +asp] in 'minor' (unstressed) syllables (Khmer). Nor have we addressed the question of diphthongization (of vowels or consonants). Diphthongization addresses simultaneity: it seems to avoid the simultaneity of a combination – not one of the given features. But diphthongization, being dissimilative, is not sequence-optimizing; it optimizes individual phonetic properties at the expense of the integrity of the segment.

[5] Children's spontaneous spellings of words like *ant* and *dent* as AT, DAT (Read 1975: 54ff) might be thought to suggest that children at first interpret the nasalized-vowel-plus-voiceless stop [ṽT] as such, /ṽT/. Spellings like AD, AGRE for *and*, *angry* which are not pronounced with [ṽD], suggest that this is not the only possible explanation of such spellings. (See Read for possibilities.)

[6] Whether or not this requires the kind of abstraction from phonetic abilities (viewed as construction of phonological constraints) implied in Hayes' (1996) proposal of 'inductive grounding' is not at issue here (cf. Donegan 1995, where constraints are more directly phonetic).

References

Andersen, Henning.1973. Abductive and deductive change. Language 49:765-793.

Best, C.T., G.W. McRoberts, & N. M. Sithole. 1988. The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. Journal of Experimental Psychology: Human Perception and Performance 14.345-60.

Donegan, Patricia. 1995. The innateness of phonemic perception. WECOL 7, ed. by Vida Samiian, 59-69. Fresno, CA: Western Conference on Linguistics.

Eimas, Peter D., E.R. Siqueland, P. Jusczyk, & J. Vigorito. 1971. Speech perception in infants. Science 171.303-306.

Gnanadesikan, Amalia. 1995. Markedness and faithfulness constraints in child phonology. ROA. University of Massachusetts, Amherst.

Hayes, Bruce. 1996. Phonetically driven phonology: The role of Optimality Theory and inductive grounding. ROA. UCLA.

Lassettre, Paul. 1995. Process effects in non-processual phonology. University of Hawai'i Working Papers in Linguistics 27.37-54.

McCarthy, John. 1995. Extensions of faithfulness: Rotuman revisited. ROA. University of Massachusetts, Amherst.

McCarthy, John, & Alan Prince. 1994. Faithfulness and reduplicative identity. ROA. University of Massachusetts, Amherst, and Rutgers University.

Prince, Alan, & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. ms. Rutgers University & University of Colorado at Boulder.

Pulleyblank, Douglas. 1997. Optimality Theory and features. Optimality Theory: An Overview, ed. by D.Archangeli & D. T. Langendoen, 59-101. Malden MA: Blackwell.

Read, Charles. 1975. Children's Categorization of Speech Sounds in English. Urbana: National Council of Teachers of English.

Smith, Neilson V. 1973. The Acquisition of Phonology. Cambridge: Cambridge University Press.

Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. Linguistic Inquiry 27:720-731.

Stampe, David. 1987. On phonological representations. Phonologica 1984, ed. by W. Dressler, H.C. Luschutzky, O.E. Pfeiffer, and J.R. Rennison, 287-300. London: Cambridge University Press.

Stevens, Kenneth, and Samuel J. Keyser. 1989. Primary features and their enhancement in consonants. Language 65.81-106.

Tesar, Bruce, and Paul Smolensky. 1996. Learnability in Optimality Theory (short version). ROA. Johns Hopkins University.

Trehub, S.E. 1976. The discrimination of foreign speech contrasts by infants and adults. Child Development 47:466-472.

Werker, J.F. & J.E. Pegg. 1992. Infant speech perception and phonological acquisition. Phonological Development: Models, Research, Implications, ed. by C.A. Ferguson, L. Menn, & C. Stoel-Gammon, 285-311. Timonium, MD: York Press.

Werker, Janet F. & R.C. Tees. 1984. Phonetic and phonemic factors in adult cross-language speech perception. Journal of the Acoustical Society of America 75:1866-1878.