It has been recognized since the discovery of phonemic representation that adults perceive speech in terms of the phonemic distinctions of their own language. Baudouin’s definition of the phoneme as ‘the mental image of a speech sound’ (1895), Sapir’s discussion of his Nootka interpreter’s transcription of ‘the intention of the actual rumble of speech’ (1921, 56), and Swadesh’s observation that adults perceive all speech in terms of the phonemes of their native language (1934) emphasized the status of the phoneme as a perceptual phenomenon. Evidence for the phonemic representation of speech sounds includes the widespread use and learnability of alphabetic writing systems; the traditional arrangements of syllabic writing systems (where symbols are arranged by phoneme groups); the characteristics of rhyme and alliteration patterns in oral and written verse; folk naming of correlative phoneme sets (like the ‘broad’ and ‘slender’ consonant groups of Irish); and differential learning of L2 sounds that can and cannot readily be identified with an L1 phoneme (as described in Wode 1992). Experimental investigation has confirmed that adult speakers perceive speech sounds categorically (Liberman et al. 1967, Lisker and Abramson 1970), and that they have considerable difficulty in discriminating between sounds which are phonetically different if those sounds do not represent a phonemic distinction in their native language (Goto 1971, Trehub 1976, Strange and Jenkins 1978, etc.).

The perceptions of very young infants are quite different from those of adults. It is widely recognized that infants as young as one month of age show something like categorical perception of speech sounds (e.g. Eimas et al. 1971, 1987). It has further been shown that young infants can perceive most of the phonetic distinctions used in any language (Eimas 1975, Streeter 1976, Lasky et al. 1975, Trehub 1976, Aslin et al. 1981, Werker et al. 1981, Werker and Tees 1984, etc.), even distinctions that are not used or perceived by the adult speakers around them. For example, Kikuyu-learning infants aged 2-3 months have been shown to be able to perceive the voiced-voiceless distinction that is used in English but not in Kikuyu (Streeter 1976). And English-learning infants at 6-8 months could discriminate both the Hindi /a/ - /a/ and /hp/ - /p/ distinctions almost as well as Hindi-speaking adults could, although most of the English speaking adults who were tested could not perform either of these discrimination tasks (Werker et al. 1981).

From examples like this, we must conclude that very young infants start out being able to perceive all of the usable phonetic distinctions – the universal set of distinctions used in the world’s languages – and end up as adults with seemingly more limited perceptual capabilities. The decline in sensitivity to phonetic differences is specific to linguistic perception; it does not involve a loss of general auditory capabilities. Adult English speakers, who ordinarily fail to discriminate the Hindi retroflex-dental contrast because they process the stimuli phonemically, can perceive the difference in certain circumstances (Werker and Lalonde 1988). If the inter-stimulus interval is very short (less than 500 msec.), for example, adult English speakers give evidence of being able to discriminate this non-native contrast (Werker and Tees 1984, Werker and Logan 1985). Other studies have shown that adults can be trained to discriminate nearly any non-native contrast (Tees and Werker 1984, Pisoni et al. 1982, Morosan and Jamieson 1989), so it is clear that the decrease in perceptual ability is not a matter of decreasing auditory sensitivity,
but of a change in processing strategy – what Werker and her colleagues call a ‘developmental reorganization’ of perception (Werker and Pegg 1992).

This reorganization of perception is not particularly surprising. It would be inefficient for listeners to continue to attend to differences which are either predictable or irrelevant in speech. What is remarkable about this change from a universal phonetic sensitivity to a native-language phonemic sensitivity is the evidence of recent years about when children begin to perceive in terms of native-language phonemes. Werker and Tees (1983) found that children aged 12, 8, and 4 years old performed perception tasks like adults: that is, these English-speaking children performed as poorly on the Hindi (non-English) contrasts as English-speaking adults did. They then tested English-learning infants 6-8, 8-10, and 10-12 months of age for their ability to discriminate two contrasts: Hindi retroflexes vs. dentals, and Nhlakapmx (Thompson) glottalized velars versus glottalized uvulars. The infants were exposed to speech stimuli from a single phonetic category, and were conditioned to respond with a head turn to the presentation of a sound from a contrasting category. The absence of a head-turn thus represented, in effect, a judgement of ‘same’; its occurrence, ‘different’. Almost all of the 6-8 month old infants could discriminate the Hindi and Thompson contrasts, but among the infants 10-12 months of age, only two of ten could distinguish retroflex from dental, and only one could distinguish velars from uvulars (Werker & Tees 1984, Werker & Lalonde 1988). A different procedure used by other researchers (Best & McRoberts 1989) reportedly produced the same finding for the velar-uvular contrast. There is thus important evidence that a ‘developmental reorganization’ takes place within the child’s first year and results in something much like phonemic perception.

This evidence for early development of phonemic perception seems to confirm the conclusions of many observers of individual children – that the child’s perceptions, even at the onset of speech, resemble adult phonemic representations (e.g. Stampe 1969, 1972; Smith 1973, Pupier 1977). There are several bases for claims for the phonological accuracy of the child’s perception and representations:

1) Perceptual confusions or systematic misperceptions on the child’s part are rare (and transient, when they do occur). E.g., a child who pronounces baby, bird, and bee as [bi] might be expected to mis-hear one as the other, but this never seems to occur systematically.

2) Striking regularities in the child productions are often explainable only with the assumption of accurate representation, e.g. A child pronounced adult /æ/ as [æ] before adult /t/, but as [ɛ] elsewhere, even though he deleted the /t/ (a word like serviette was [sæːjɛt]) (Pupier 1977). Further, the child seems to go beyond the phonetic form to perceive in terms of adult intentions, as when adult [baʔn] or [baʔtn] is pronounced by the child as [baðon], although the adults around him never used a released [t] or [d], or a second vowel.

3) Mastery of a new articulation affects known words, without the child’s having to re-hear them. There are exceptions, and some words may lag behind, but such exceptions may be attributed to the influence of the child’s own pronunciations on his underlying forms (cf. Macken 1980). (The absence of some information from the child’s phonological forms would not substantially affect this basic claim.)


It is important to consider here what is meant by ‘phonemic perception’, ‘phoneme’, etc. Werker and Pegg (1992) seem to hedge on Werker’s earlier claim that the reorganized, language-
specific perceptions of children at the end of their first year are actually 'phonemic perception'. Instead, they refer to this as 'language-specific phonetic perception', because, they say, they have no evidence that the child actually distinguishes lexical minimal pairs on the basis of these sound differences.

But this assumes that the phoneme is primarily a unit of lexical contrast, rather than a unit of perception, memory, and intention – a 'possible sound' in a given language. It was the structuralists, in an effort to make phonological analysis 'empiricist', who took the criteria of contrast and complementary distribution of phonemic analysis and changed their status from heuristics to definitions. But as Chomsky (1964) and Stampe (1987) have argued, when the distributional theory of the phoneme and the perceptual theory disagree, the distributional theory always turns out to be wrong. This seems a small loss, since distributional analysis seems to be an impossible model for acquisition (Donegan 1985).

A key to better understanding of the relationship between phonemes and their allophones and thus, the relationship between different phonemes, was offered by Bazell (1954). Bazell contrasted the phonemicists hesitation to group initial [h] and non-initial [n] in English as allophones of a single phoneme with the general willingness to group Japanese [ɸ] (which occurs before [u]) and [h] (which occurs elsewhere) as a single phoneme, /h/. He pointed out that this was simply because there was no phonetic motivation for an alternation of /h/ with final [n], while the appearance of [ɸ] before [u] is motivated.

According to Bazell, the distribution itself meant little without reference to the intrinsic character of the segments, and the aim of phonemic analysis 'is to reach a system whereby intrinsic features and distribution are mutually explanatory ... The phonemes are the arbitrary residue left after the deduction ("discounting") of whatever is to be regarded as motivated' (1954, 134).

Stampe (1987) has claimed that this principle of phonemic analysis applies to hearers as well as to phonemicists. He observes that allophones are in a relationship of substitution with the basic phoneme, and the substitutions are phonetically motivated. If there is no phonetic motivation for the substitution, there is no motivation for a phonemic unity. Stampe also pointed out that there are phonetic motivations for the basic phonemes, as well as for allophones. For example, the nasality of vowels before nasals is motivated or optimal, and therefore discounted, but the non-nasality of vowels is motivated, too. Likewise, the appearance of fricatives for stops in weak or medial positions is motivated, but the basic stops are motivated, too. The motivations toward optimal segments, first described with Jakobson's implicational laws (Jakobson 1968), interact with the motivations for allophonic substitutions and create a set of possible intentions, which underlie the phonetic realizations.

A phoneme is thus a sound that can be perceived and produced as itself (not as a variant of some other sound). Thus a phoneme is what the hearer perceives as the sound the speaker intended to say; the hearer arrives at the speaker's intention by identifying the speaker's limitations as a speaker with his own limitations and by attributing to the speaker the same kinds of substitutions he himself would make, if he were speaking. It is important to realize that phonemic representation varies with the actual pronunciation, or utterance. Variant pronunciations of a word like sixths, like [siks0s], [siks], [siksts], may have different phonemic representations. Phonemic representation refers to the utterance; it is not essentially lexical.
But let us return to the original question of the child’s perceptions and how they become phonemic. Obviously, phonemic perception is not ‘innate’ in the sense that each child is born with this kind of language-specific perception. So how does a child determine which characteristics of adult speech must be attended to and remembered, and which others are predictable from the phonetic limitations of the adult speakers? How does the English-learner discover that vowel nasalization and stop aspiration may be ignored, while the Hindi-learner discovers that these differences must be noticed and remembered? What happens to the child’s perceptual abilities within the first 10 months of life?

2. The Discovery of Phonetic Features and Processes

An important part of the answer to this question seems to be that the child learns about the abilities and limitations of her own vocal apparatus, through vocalization and babbling. Although the infant vocal tract is straighter and shorter than the adult’s (Bosma 1975, Lieberman et al. 1972), it begins to assume a more adult-like form by about 3-4 mos (Sasaki et al. 1977). The infant’s earliest sounds are mostly phonatory, or vowel-like, with only occasional closing gestures, but by about 4-6 months, they begin to include closing, consonantal gestures with varying places of constriction. Then these come to alternate with open, fully resonant vowels, eventually becoming more like the sequences of CV syllables that have come to be called ‘canonical babbling’. In all of this prelinguistic vocalization, babbling, and early imitation, the child acquires three important kinds of knowledge.

2.1. Features

First, through auditory (and proprioceptive) feedback, the child begins to establish motor-auditory-kinesthetic connections, connections between articulatory gestures or positions and their acoustic or auditory effects (Fry 1966, 188-190; Locke & Pearson 1992, 115; Menn 1992). All of the infant’s own vocalizations can contribute to the establishment of these acoustic-articulatory connections as she creates a mapping of gestures or articulatory positions to sensory outputs, a sort of ‘phonetic guidance system’. Other factors that may influence the development of the system include observation of speech activity in others and active articulatory practice with auditory self-monitoring (Locke and Pearson 1992). These associations of particular gestures or kinds of gestures with particular acoustic effects can be identified with what phonologists call ‘phonetic features’. For example, the articulatory action of complete oral closure (and release) results in an interval of silence or of low amplitude, followed by an abrupt onset of energy over a range of frequencies; an incomplete oral closure (and release) results in more sound during the constriction and a less abrupt increase in amplitude at its release. When the child recognizes this association, she has discovered the feature [continuant]. So-called ‘feature analysis’ is thus, in a sense, ‘feature synthesis’: it is the establishment of a connection between articulation and effect. The connection of auditory characteristics to articulations is, of course, essential for imitation (including self-imitation).2

---

1 The infant vocal tract has a broader oral cavity, a shorter pharynx, a gradually sloping oropharyngeal channel, a relatively anterior tongue mass, a closely approximating velum and epiglottis, and a relatively high larynx (Kent 1992, 69).

2 This view of a ‘feature’ as an aspect of speech that is independently controllable and has a detectable (often quantal) result allows that, in babbling, children learn about more features than
2.2. Fortitions

Second, the child begins to experience the articulatory and perceptual optimality of certain combinations of phonetic features, and to learn about the difficulties the less-optimal combinations represent. Such problems include the inhibition of voicing by the close constrictions obstruents require, the superior audibility of varying vowel articulations when the velic port is closed, the audibility of nasality in voiced sounds and its relative inaudibility in voiceless ones, or the finer adjustments required for fricatives as opposed to stops or for front or rounded vowels that are mid as opposed to high. These may represent phonetic constraints on the child’s productions: obstruents become voiceless, vowels become non-nasal, nasals become voiced, fricatives become stops, front or rounded vowels become high.

2.3. Lenitions

Third, the child experiences co-articulation and the context-dependent variation in the sounds that she intends to produce. An articulation may have variant forms that occur in different articulatory contexts, and such combinatory variation has varying acoustic effects; these variations must be integrated into the child’s phonetic awareness. For example, the child finds that an optimal dorsal consonant, the velar [k], may be fronted to [c] before or after front vowels, or backed to [q] before or after low back vowels. Or an optimal stop, the voiceless [p], may be voiced between voiced segments. Or an optimal vowel, which is non-nasal, may be nasalized adjacent to a nasal consonant.3

The child’s vocal explorations thus result in an implicit body of knowledge about the constraints that limit her ability to produce particular individual sounds (simultaneous feature combinations) or sequences of sounds, and about the alterations or substitutions that result when she submits to these constraints. The perceptual aspect of this body of knowledge includes both the recognition of optimal segments, and the realization that deviations from these optima occur in certain circumstances so as to optimize sequences. The features, the constraints the child discovers, and the adjustments or substitutions that respond to these constraints form the basis of phonology. These constraints are universal. Likewise, the interactions of these substitutions are

---

3 Some co-articulation may be universal and inevitable, the result of mechanical properties of the vocal tract. There are other aspects of articulation, however, which represent articulatory optima, but which a given language may require its speakers to learn to avoid. For example, although velar stops appear to be the optimal dorsal stops, some languages distinguish velar stops from palatals or uvulars (regardless of the following vowel). Similarly, although continuous voicing represents an articulatory optimum, some languages distinguish voiced and voiceless consonants intervocally.
universal: the substitutions that optimize phonetic properties of individual segments apply before those that optimize sequences (Donegan and Stampe 1979).

3. Phonological Implications of the Phonetic System

But, of course, phonologies of different languages differ in what they require their speakers to learn. One can overcome a constraint by learning to pronounce the difficult configuration. Doing so requires some effort and attention, so the speaker or learner prefers, for perceptual and/or articulatory reasons, that constraints apply. In learning to pronounce a language, the speaker learns to overcome only those constraints his language requires him to. Speakers of English may allow the constraint, *vowels are non-nasal*, to apply; speakers of Hindi or French may not. Speakers of Hawaiian may allow the constraint, *obstruents are voiceless*, to apply; speakers of English may not.

The context-free, segment-optimizing, fortitive constraints (like ‘Vowels are non-nasal’, or ‘Obstruents are voiceless’) limit the inventory of ‘possible’ sounds, in the sense of sounds the child can actually produce; they may also limit the inventory of sounds the child will perceive as ‘possible’ – that is, intended or memorable or significant. If a phoneme is ‘the mental image of a sound’ (Baudouin 1895), these constraints limit the child’s phoneme inventory. But by considering other phonetic constraints which optimize sequences of sounds, the child may discount certain phonetic features (like vowel nasalization before nasals, or voicing in obstruents between vowels) as the inevitable results of context-sensitive, sequence-optimizing, lenitive constraints that apply in her own productions and, presumably, in those of adult speakers around her. Thus, sounds that are ruled out by the segment-optimizing constraints – but which in fact occur – may be perceived as variants, or allophones, of sounds that are allowed.

It will help to look at a couple of very simple examples of this interaction in some adult languages. Fortitive and lenitive constraints differ in phonetic motivation, and may consequently have opposite effects. For example, the constraint \(-\text{son} \rightarrow -\text{voi}\) (Obstruent Devoicing, or OD) creates articulatorily and perceptually optimal obstruents by substituting voiceless for voiced obstruents. The constraint \(-\text{son} \rightarrow +\text{voi} /+\text{voi} +\text{voi}\) (Intervocalic Voicing, or IV) creates articulatorily optimal sequences, making voicing continuous, by substituting voiced for voiceless obstruents (especially in syllable-final or unstressed positions). Devoicing applies in Southern Chinook, Hawaiian, Tamil, Yidîn, etc., but not in English, Sanskrit, Danish, French, etc. Intervocalic Voicing applies in Yidîn, S. Chinook, Sanskrit, and Danish (with some qualifications, in most of these languages), but not in English French or (usually) in Hawaiian.

Since each process is phonetically motivated, the preferred state for each is application (+), but speakers of any language must learn to master some difficulties, thus acquiring the skills that allow them to limit or ‘turn off’ some processes (−). Other processes continue to apply. The interaction of fortitions and lenitions that apply for adults creates a language-specific phoneme inventory and a pattern of automatic alternations.

We can see this in the schematic example in the table below. A ‘+’ indicates that the process applies; a ‘−’ indicates that process does not apply: the speaker must learn to product the more-difficult configuration that the process avoids.
In each case, the phoneme inventory is defined by the interaction of the universal constraints: the fortitions limit the set of intendable sounds to a set of relatively optimal segments, and the lenitions determine their pronunciation in context (sometimes creating ‘impossible’ sounds like [b] in S. Chinook) as modifications of the ‘possible’ sounds.

The learner, then, can arrive at the phoneme inventory of his language and the relations of phonetic to phonemic forms – not by analyzing distributions of sounds, but by discovering which phonetic processes fail to apply in his language. In effect, he learns the phonemic inventory of the language by discovering which of his own phonetic limitations he is going to have to learn to overcome. The child does not have to perform a distributional analysis in order to discover this; decisions can be made on the basis of a single form.

The S. Chinook learner, for example, hearing forms like [paba], can assume that OD applies, ruling out */b/ as a possible phoneme. She can attribute the [b]’s she hears to the application of IV. The Hawaiian learner, hearing only forms like [papa], can also assume that OD applies, and also rules out */b/, but because she hears intervocalic [p], she must note that IV does not apply. (Her initial pronunciations may undergo IV, until she acquires the articulatory control to overcome it.) The English learner, hearing [p]-initial and [b]-initial forms like [pa] and [bi], must recognize that OD does not apply. He need or remember or compare minimal pairs; it is the absence of a phonetic motivation for the voicing difference that tells him that both /p/ and /b/ are intendable sounds. He also hears voiceless intervocalic stops, and must recognize that IV does not apply, either.

Like the English learner, the Sanskrit learner hears both [p]-initial and [b]-initial forms like [pa] and [bi], and must recognize that OD does not apply. The intervocalic stops he hears, however, are always voiced (at least in certain prosodic environments), so he may, unlike the English learner, allow IV to apply. This means that some intervocalic [b]’s might arise from /p/’s. But since /b/ is an intendable sound, a phoneme, these [b]’s are at first perceived as /b/’s. Only later, when the child begins to identify some morphological variants, does he begin to identify these phonetic [b]’s with /p/’s. But even this does not require distributional analysis. His phonemic representations of morphemes can be revised to morphophonemic representations one at a time, as he recognizes the morphological unity of the varying forms, because the substitution that accounts for the alternations is there all the time, available to him, part of his knowledge of articulatory optima and the substitutions speakers with vocal tracts like his may use to achieve them.

Note that the application of both automatic processes means that the S. Chinook learner has the least to learn with respect to the articulatory skill it takes to control voicing. The English
learner, who must overcome both the OD and the IV processes by learning to produce voicing distinctions in all environments, has the most to learn in this respect.

Compare the application and non-application of Vowel Denasalization (VD), a fortition, and Vowel Nasalization, a lenition, in English, Hawaiian, Hindi, and French:

<table>
<thead>
<tr>
<th>Process</th>
<th>English</th>
<th>Hawaiian</th>
<th>Hindi</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>VD: V -&gt; -nas</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VN: V -&gt; +nas / _+nas</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

The inventories that result are:

<table>
<thead>
<tr>
<th></th>
<th>/a/</th>
<th>/a/</th>
<th>/a/</th>
<th>/a/</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a~ a]</td>
<td>[a]</td>
</tr>
<tr>
<td>Hawaiian</td>
<td>[a]</td>
<td>[a]</td>
<td>[a~ a]</td>
<td>[a]</td>
</tr>
<tr>
<td>Hindi</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
</tr>
<tr>
<td>French</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
</tr>
</tbody>
</table>

and the phonetic realizations are:

<table>
<thead>
<tr>
<th></th>
<th>[a~ a]</th>
<th>[a]</th>
<th>[a~ a]</th>
<th>[a]</th>
<th>[a]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
</tr>
<tr>
<td>Hawaiian</td>
<td>[a]</td>
<td>[a]</td>
<td>[a]</td>
<td>[ä]</td>
<td>[ä]</td>
<td>[a]</td>
</tr>
<tr>
<td>Hindi</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
</tr>
<tr>
<td>French</td>
<td>[a]</td>
<td>[ä]</td>
<td>[ä]</td>
<td>[a]</td>
<td>[ä]</td>
<td>[a]</td>
</tr>
</tbody>
</table>

In order to arrive at these inventories, children learning these languages must consider the phonetic realizations on the light of the phonetically motivated processes. The English learner may allow both VD and VN to apply, and can perceive nasalized vowels as variants of non-nasal ones, attributing their occurrence to VN. The Hawaiian learner may also allow VD to apply, since phonetic nasalized vowels are not ordinarily encountered. (Any nasalization that does occur can be attributed to VN, which may apply optionally.) The Hawaiian learner is thus in nearly the same position as the English learner, except that 1) he has fewer phonetic instances of vowel nasalization to account for and 2) he must learn to produce non-nasalized vowels before nasals, at least in some circumstances – VN is an option which may apply in faster or more careless speech, in unaccented syllables, etc., but it is not allowed to apply always, as it does in English.

The Hindi learner, hearing nasalized vowels where their occurrence cannot be attributed to the presence of a following nasal, must recognize vowel nasalization as the intention of the speaker. In doing so, he admits he must master the difficulty represented by VD, and learn not to denasalize vowels. He can, however, allow VN to apply, since he hears no non-nasalized vowels before nasals. This means, however, that a nasalized vowel before a nasal is perceived as a nasalized vowel (cf. Lahiri and Marslen-Wilson 1991). (A nasalized vowel may eventually be interpreted morphophonemically, as a non-nasalized vowel, when the learner recognizes a morphological identity of forms with nasalized and non-nasalized vowels and attributed the nasalized variant to the nasalization process.)

Like the Hindi learner, the French learner hears nasalized vowels in non-nasal environments, and must thus admit them to the inventory, overcoming VD. Unlike the Hindi learner, the French learner must also learn not to nasalize vowels before nasal consonants.

<table>
<thead>
<tr>
<th>Phonetic categorization</th>
<th>Process marking</th>
<th>Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>English learner hears [dän], [dat]</td>
<td>VD applies: /a/, */ä/</td>
<td>/a/, */ä/</td>
</tr>
<tr>
<td></td>
<td>VN applies:</td>
<td>[a] [ä]</td>
</tr>
<tr>
<td>Hawaiian learner hears [käpu], [kanë]</td>
<td>VD applies: /a/, */ä/</td>
<td>/a/, */ä/</td>
</tr>
<tr>
<td></td>
<td>VN does not apply:</td>
<td>[a]</td>
</tr>
</tbody>
</table>
Hindi learner hears forms like [ba], [hã]  
VD does not apply:  /a/, / ā/

Hindi learner hears forms like [bïn]  
VN applies:  l  l  
[a]  [ã]

French learner hears forms like [ta], [pi]  
VD does not apply:  /a/, / ā/

French learner hears forms like [bon]  
VN does not apply:  l  l  
[a]  [ã]

4. Summary

As children discover their phonetic abilities, in early vocalization, babbling, and imitation, they also discover the limitations on these abilities and the substitutions that optimize their productions and make them conform to these limitations. These limitations tend to keep a child’s own inventory of productions small, while the set of different segments the child can hear may be quite large. To learn to speak like adults, children have to overcome some of these limitations, learning how to pronounce some segments and sequences which are not optimal. As they begin to recognize adult productions, they learn which of these phonetic difficulties they must master – always hoping to have to master as few as possible – always hoping to allow the substitutions which optimize segments or sequences to apply. To admit a phoneme to the inventory is to recognize that one will have to learn to produce that sound as itself, and not as a variant of some other sound.

Phonemic perception is perception of sounds as intended or intendable. The ‘developmental reorganization’ that results in phonemic perception occurs when children begin to perceive in terms of what they might themselves pronounce. The limitations on our phonetic abilities cause us to reduce the variety of sounds we hear to a small, controllable, intendable inventory. So phonemic perception is innate because these phonetic abilities and limitations are innate. Phonemic perception is learned by learning which of these innate limitations we must overcome.

References


***

*Full reference data: