PART V

Features in phonological development

Features in child phonology

Inherent, emergent, or artefacts of analysis?*

Lise Menn & Marilyn Vihman
University of Colorado/University of York

The emergence of features plays a key role in any theory of phonological development that does not assume that they are innately available before the onset of speech. After reviewing Jakobson's claims for universal orders of emergence, we consider possible criteria for feature or segment acquisition, and then discuss data from nearly 50 children (10 languages including Estonian, Finnish, Japanese, and Welsh as well as several Germanic and Romance languages). Small early vocabularies and phonetic variability make minimal pairs rare in most children. While a few children show clear evidence of utilizing features or segments, others show none, and most fall between these extremes. Gradually increasing evidence of segmental structure and systematicity reflects the slow transition to a more orderly phonology. These data support an emergentist model of feature acquisition that has many possible routes to (re-)creating phonological organization within the individual child's mind.

1. Introduction

The more one scrutinizes it, the more complex the notion of acquiring a phonological feature becomes, yet the role of features is a basic question for any theory of phonological development. Jakobson's magisterial account of the emergence of features is so far removed from what children actually do that it is essentially useless as a serious basis for the study of acquisition. The nature of features remains controversial; however, it is clear that they serve a useful role in the

* We thank Amy Bidgood, Philippa Claxton, Michelle McGillion and Helena Sears, who collected and transcribed the UK data presented in Appendix II, and we are also grateful to the National Science Foundation (United States), which supported the collection of Vihman's American English, French, Japanese and Swedish data, and to the British Economic and Social Research Council (United Kingdom), which supported collection of both the British English and the Welsh data. We also thank colleagues who have discussed these problems with us and who have given us editorial feedback.

description of adult languages. It is puzzling, then, to find that, for most children, they are of little help in characterizing the very first words identified. On the other hand, they can be helpful at a later stage for describing both the child's output and the model-to-output relationships. This change over time is one reason to claim that phonology emerges.

The primary reason for the inefficacy of features as descriptors at the early stages of phonological development is that the first word forms tend to be the least systematic. Furthermore, generalization across children is difficult for the early word forms but somewhat easier later on: Different children, when studied in detail, approach the problem of acquiring phonology in rather different ways, only slowly converging as they move towards the structure of the ambient language. This evidence against early phonological systematicity, both within and across children, is a serious problem for claims that a well-defined phonological system unfolds or is reorganized within each child from the start. Instead, the fact that, with time, children gradually become more systematic and more similar to other children acquiring the same language supports the claim that phonology emerges on the basis of experience with language production and self-perception.

We present substantial amounts of data to back up these statements: Appendix I shows variability across many children acquiring different languages, and Appendix II focuses on variation within each of a small number of children acquiring British English. Dismissing the less-systematic early phonology as 'pre-linguistic' – as Jakobsonians have sometimes done – precludes the possibility of establishing that the system develops gradually from a precursor state. This gives spurious support for the claim that the phonological system appears fully formed, needing only to be unfolded or rearranged, as well as for the argument that it must therefore be innate.

2. Some history

Child phonology is still struggling with the challenging legacy of Roman Jakobson, whose writing on child phonology, most famously *Kindersprache, Aphasie, und allgemeine Lautgesetze* 1941 (*Child Language, Aphasia, and Phonological Universals*, 1968), included many resounding claims that have proven quite difficult to evaluate (Kiparsky & Menn 1977). Jakobson is selective in his presentation of data (which at an early developmental point are sparse to begin with), and his categorical statements overwhelm the hedges by which he occasionally acknowledges this. For example, consider this 1949 statement from *The sound laws of child language*, reprinted in Waugh & Monville-Burston (1990) [emphasis as in the original]:

> By eliminating these particular facts [about apparently onomatopoeic utterances recorded by various linguist parents] and following step-by-step the formation of the child's phonemic system, we discern a rigid *regularity* in the succession of his acquisitions, which constitute for the most part a strict and invariable temporal sequence…. (p. 297)

The categorical statements themselves pose a problem of interpretation: how can something 'for the most part' show a 'rigid *regularity*' or a 'strict and invariable temporal sequence'?

Similarly, consider Jakobson's well-known statements about phonemic oppositions:

> Ordinarily the vowel system originates in a low vowel and the consonantal system simultaneously in a stop with an occlusion at the front of the mouth… The first opposition within the consonantal system is between nasal and oral, and the second between labials and dentals…Following the appearance of these two consonantal oppositions a high vowel arises in opposition to the wide vowel in the child's speech. (p. 297)

Again, we have a problem of interpretation: what does it mean to say that something happens 'ordinarily', followed by an un-hedged statements of what is claimed to happen next? It also turns out that these claims about oppositions are hard to falsify, even in their strong form, because early vocabulary, by definition, contains only a few words per child: When there are only five or even ten words, the minimal pairs that would be needed to test them are often lacking. Nevertheless, there is indeed counterevidence, as we shall see.

Jakobson's famous progression of phonemic oppositions remains popular partly because it is fairly close to the general progression of acquisition of phoneme classes as phonetic targets: English stops in general do precede fricatives and affricates; glides precede liquids.[1] And velars are indeed usually later than labials and dentals, in English and several other languages (Finnish, with common use of early words with /k/, is one exception: cf. Kunnari 2000, Savinainen-Makkonen 2001, 2007; Japanese is another: Beckman, Yoneyama & Edwards 2003). In languages employing complex articulations, simply articulated consonants such as [k] are reported to be mastered before more marked phones such as glottalized [k'] (Pye, Ingram & List 1987). The influence of phonetic frequency (and probably functional load) in the ambient language is evident, however: a phone is generally mastered

1. Interestingly, although Jakobson famously dismissed babbling as irrelevant to phonology – and of course no really accurate description of babbling could have been made before the development of the tape recorder in the 1950's – the progression of oppositions also looks closer to what happens in babble than to what happens in speech (see Vihman 1996, Ch.2).

earlier in a language in which it is more frequent (*ibid*.; Ingram 1992). But as is well known from, for example, the very late mastery of /ð/ in English, frequency cannot be the sole determining factor.

The crucial theoretical point obscured by the rough correspondence between Jakobson's prediction and the reported group data is that *general statements cannot speak directly to the question of development in individuals*. 'Laws of grammar' – in the linguistic sense of 'grammar' – operate within the individual mind. The implicational hierarchy is about individuals, not about groups, and it is explicitly about contrasts, not phonetic targets – yet target acquisition orders are the only information we can extract from group pronunciation accuracy data. And individuals, as is now well known, occasionally violate Jakobson's proposed 'laws' and, more frequently, violate the inferences that are typically drawn from them (see, for example, several children's use of velars before alveolars in our Appendix I and, in Group 4, Virve's stop-fricative contrast preceding any place contrast in stops).

With regard to order of acquisition of vowels, there is a specific point where Jakobson's typologically limited child language evidence may have misled him. Note, in Appendix I, the several instances of a child's first vowel contrasts arising between front and back low vowels (e.g. Nicola, Group 3b; Jonathan, Group 3c). Jakobson's adult-language distributional data led him to claim a primacy for the /i – u – a/ vowel triangle, and therefore to predict, as in the paragraph we quoted above, that children would start differentiating their vowels by first developing either a height contrast or a front-back split between high vowels. He claimed that child language data bore out this prediction, but the data available to him in writing *Kindersprache* were derived from a very narrow range of adult languages. Of the languages he cited, only English (Wellman et al. 1931), Danish (Jespersen 1916, 1925; Rasmussen 1913) and Norwegian (Kock 1901) have [æ] as a phonetic target (in the Scandinavian languages, as a possible realization of one of several front mid-vowel phonemes), and Jakobson cites none of these sources in the context of vowel development. Wellman et al. indicate relatively late [æ], which might have contributed to Jakobson's impressions, but they have only group phonetic target information and, as we have said, neither group production data nor phonetic target data can sustain claims about individual paths of phonological development, given the extent of individual differences and the absence of information about contrasts.

For the children in Appendix I who are acquiring English and expanding their vowel systems very early, like Sean, T., Will, and Jonathan, a contrastive or near-contrastive early a/æ split actually seems to be reasonably common. Appendix I also shows that low front [æ] is quite common as a target for additional English-learning children who lack the minimal pairs needed to render the phonetic [a]: [æ] distinction contrastive. So (although children *do* sometimes create contrasts absent from the ambient language to preserve a distinction that they cannot make

in the way that the ambient language does: Cf. [i]: [y]: Timmy, acquiring American English, in Vihman, Velleman & McCune 1994), it is likely that Jakobson inadvertently biased his acquisition data against finding early low-vowel front-back contrasts because he depended so heavily on reports of French, German, and Russian child phonology. This, coupled with the English data showing several cases of early low-vowel front-back splits, negates his claims about the generality of the implicational splits in the vowel triangle during development, whatever may be their status in the world's languages.

Before dealing with the theory of child phonology, we need to alert readers to a methodological problem, which we will return to in more detail in Section 6: One reason that it is difficult to establish contrast in young children is the variability of their renditions of a given adult target. The variability in child production (and also the relative unreliability of transcription of child forms, which typically gives no better than about 80% agreement on segmental identity across transcribers) makes it difficult to judge whether, for example, Alice's use of [pɛ-] for *baby,* [dæ] for *daddy,* and [mɑ -] for *mommy* (in Appendix I, Group 1) should be taken to reflect an intention to match the target and maintain the adult contrasts.

3. Theoretical perspectives

Our field has learned a lot in the 60 years since *Kindersprache* was published, but positions on the origins of systematicity still seem to be very much dependent on the investigator's theoretical preferences. The basic difficulty may lie in the fact that gradual change is inherently difficult to describe. Also, individual differences among children as well as across languages defeat all but the loosest general statements, unless they are formulated in probabilistic terms. Furthermore, the question that we address here, namely, how best to determine whether features actually play a role in early phonology, has rarely been raised explicitly.

What would it mean for features to be 'artefacts of analysis'? One should not use such a dismissive term lightly. Certainly, features are not always artefactual. If a feature functions in ways generally taken to reflect psychological reality – for example, if it spreads, plays a role in generalizations, or divides sounds into classes that are treated consistently within class but differently across classes – then it may serve as a valid descriptor in the child's phonological system. So what can be said now about the development of features?

3.1 The search for criteria: How may one define 'having a feature'?

As frequent transcribers know, resorting to descriptive features is very useful when a child's articulatory control seems poor and transcription of segments is

dicey, but where one may be able to say that a particular sound is, say, an anterior stop or a low vowel. Description/transcription in terms of segments in cases like this is over-interpretation, and necessarily introduces guessed-at segments that can lead to inappropriate analyses; description in features is appropriately conservative. The usefulness of features in descriptions like these is neither a problem nor a theoretical issue. The theoretical problem is whether features are *units* that are learned (or inherited, developed or acquired) as such. More concretely: Does it makes more sense to speak of a child as (a) acquiring features (or demoting constraints against those features), (b) acquiring segments that happen to have particular features, or (c) acquiring words that have those features?

Cristià and colleagues (Cristià & Seidl 2008; Cristià, Seidl & Francis this volume) make an important distinction between features as *distinguishers* (describing what is different between two speech sounds) and features as *classifiers* (describing what two or more speech sounds have in common). In our usual thinking about adult phonology, both of these aspects of the feature are part of its definition, but there is no *a priori* reason to assume that if a child can make a distinction between two sounds that we represent as being [+Feature] and [–Feature], she can classify all the sounds belonging to the set which we would call [+Feature] as being 'the same' in opposition to all those which we would call [–Feature]. As Cristiá and her colleagues point out, for some features, classificatory ability has been demonstrated for pre-speech children, but for features with little direct correspondence with acoustic properties, such demonstrations are largely lacking.

In addition, as regards phonological development over the first year of life, we must take into account the developmental shifts identified by Werker and her colleagues: Early speech-sound discrimination does not guarantee later word-based discrimination. This is true not only for the well-known case of sharply decreased discrimination of most of the contrasts not found in the ambient language(s), such as English-learners' loss of the ability to distinguish voiced from murmured stops, or velar from uvular ejectives (Werker & Tees 1984), but also for the difference between being able to discriminate such minimal pairs perceptually and the more stringent requirement of distinguishing them in the context of use as labels for different referents (Stager & Werker 1997; Werker & Stager 2000). This is a task difference that is all too familiar to the adult second-language learner, for whom the need to not only distinguish a contrast but retain it in long-term memory may prove an insuperable barrier to native-like production and perception.

Rather than promulgate a particular position as to when it would be reasonable to speak of 'having' features, we consider various patterns of early output data that one might possibly take as criterial. We hope that this strategy will enable our colleagues to apply their preferred criteria while making those criteria explicit, to facilitate genuine discussion. But given the following spectrum of conceivable

positions, we think that the data that we discuss under Section 4, below, and present in full in Appendix I are clear enough to rule out both the first (which is too weak) and the last (which is too strong, since it is not true of adults either). For convenience, we use expressions like 's/he has a feature', but – since grammar is inside a speaker's head and we have no direct access to it – what we really mean is that 's/he functions as if s/he had a feature'.

1. If the child produces a word-like utterance, s/he has all the features needed to describe the segments that an adult can hear in that utterance.
2. If the child produces a segment, s/he has all of the features involved in defining that segment
3. If the child produces two segments that share one or more features, s/he has the features that they share.
4. If the child produces a minimal pair of word-like utterances, s/he necessarily has the feature that separates them.[2]
5. The child has a feature only if it plays a contrastive role in minimal pairs involving two different pairs of segments contrasting in that feature (e.g. [continuant] functions as a feature if both /p/: /f/ and /t/: /s/ contrasts are present).
6. The child has features only if those features are freely recombinable.

3.2 Problems with the idea of the feature as 'unit'

One line of argument about whether a feature is a unit goes like this: For an adult, a phone is a very tight feature bundle, a gestalt not easily decomposed. We know this because even when the features involved are distinctive in the first language, second language learners cannot freely re-combine them: English speakers must work to learn to make velar fricatives, and Arabic speakers often have persistent difficulties in differentiating English /p/ and /b/, in spite of having the voicing contrast between other pairs of stops.

However, this observation in itself does not preclude the possibility that features are units, because they could start out being relatively independent (freely combinable); the tight bundling into segments that is found in adults might be something that develops over time. But does it? Or is tight bundling present from the outset? Or might that vary across children?

Before we look at the available evidence, let us consider the logic a bit more. In other sciences, units exist without being freely combinable: atoms are fussy about which others they make molecules with. Units don't have to be strictly

2. With respect to [±voice], this is precisely the point at which Jacob had [b], [k], and the alveolar /t/: /d/ contrast but not [p] and [g].

hierarchical: humans belong to multiple overlapping social groups. They don't have to be discrete: micro-level ecozones overlap and interpenetrate in a landscape. So if arguments about 'being a unit' are not to descend to the level of *Is not! Is too!*, proponents of either position have to make it clear what they intend their claim to imply. We will stipulate, then, that we are willing to call a feature a unit for a given child if and when that child's patterns of contrast *or* her patterns of altering adult words are more effectively described in terms of features than in terms of segments or longer units (although we recognize that there is an element of subjectivity in judging the 'effective'-ness of a description, so there will still be room for debate).

3.3 Criteria based on class omission or feature spreading

How could one support the claim that a particular child is acquiring features as opposed to segments, words or templates? We often have evidence that a group of sounds sharing a feature are all being treated in the same way, which, as we said above, would seem good grounds for considering that feature as a unit. Take a specific example: after his two early words beginning with /h/ (which, incidentally, were maintained intact throughout his development), Menn's Daniel (Danny) omitted all initial continuant consonants, from fricatives through glides. Because Danny did produce most of these sounds (except for the voiced fricatives and /h/) in word-final position, let us take it as a working assumption that he at least perceived and therefore differentiated them as consonants-that-were-not-stops in word-initial position. But how did he perceive them: as segments? as parts of words? or as a natural class? It is not clear that this question can be answered, and yet only the last choice would be equivalent to having acquired the feature [continuant]. (Even if we were certain that he had a [continuant consonant] natural class, its precise description would be subject to the usual debates about the particular feature(s) that should be used to characterize it.)

Perhaps omission of all the members of a set is too weak a phenomenon to argue from. Let us consider a stronger one: the widespread harmony constraint against two positions of articulation in the same word $C_xV_xC_y(V)$ (which under some interpretations of vocalic features – e.g. Levelt [1995] – may include the articulation of the intervening vowel). This constraint is one of the strongest arguments for an autosegmental approach to child phonology (Menn 1978, 1983; Macken 1995), and it has many implementations within and across children, such as selection of only words with one place of articulation in the adult form, omission of one of the disharmonic segments (often the earliest clear manifestation of the constraint), replacement of one of them by /h/ or glottal stop, and most dramatically although usually later, place assimilation across the intervening vowel.

Assimilation is a form of feature spreading and is probably the best evidence one can have for the reality of the feature that has spread. If systematic, the other manifestations of the harmony constraint – avoidance of disharmonic forms like *cup*, omission of one of the disharmonic segments, and replacement of a disharmonic supraglottal segment by a glottalic one – are of course nicely handled formally by a harmony (or a single-place-feature-per-word) constraint. But if one takes a skeptical stance, the only thing we are sure that the child knows about disharmonic words is that she can't say them, or that she can't say them unless she gets rid of one of the supraglottal consonants. If she is systematic about *which* of the supraglottal consonants she gets rid of, and the system is based on the place of articulation (e.g. always deleting a labial if a velar is present in the same word) rather than the position of the consonant in the word (e.g. always deleting the final consonant in a disharmonic word), *and* the system applies to two or more consonants that share a place of articulation (*pick, fig, pig, big,…*), then that is good evidence for the validity of a place-featural description of the consonant that is retained – but not for the one that is deleted. Why so? Because all that the child may know about the deleted segment is that it does not have the same point of articulation as the one that she retains.

Some caveats are warranted even in the case of assimilation, however. First, the main reason that a feature-identity constraint seems preferable to a whole-segment-identity constraint (C_1VC_1[V]) is the fact that the two C's may differ in voicing. However, with the English prevoiced-to-short-lag VOT for 'voiced' stops and long-lag VOT for stressed-syllable-initial voiceless ones, it is common for all of an English-learning child's early stops to be transcribed as voiced in word-initial position but voiceless in word-final position (see Macken 1980; Macken & Barton 1980 on the relatively late acquisition of voicing contrast). Children who are young enough to have an active (as opposed to residual) consonant assimilation pattern rarely have voicing contrast in both word-initial and word-final position. Some cases of apparent same-*place* constraints for stops might be produced by a same-*segment* constraint, coupled with onset voicing and coda devoicing, as in Hans (Lindner)'s [gak] for *gasse* 'street' [gasə] or Menn's Danny's [gak] for *sock*. Another difficulty for a feature-based place-assimilation rule is that for some consonant-assimilating children, like Danny (although unlike Hans), final /s/ escapes assimilation almost as soon as it is attested: Danny's *mouse* and *glasses* were [mæʊs] and [gæs].

We note finally that breaking out of the constraint typically occurs one segment pattern at a time; labial-V-alveolar (Macken 1979) and velar-V-coronal (Menn 1971; Berman 1977) have both been reported as the first disharmonic sequence. This makes it hard to support any more abstract or acoustic type of feature such as grave or acute that would group places of articulation together.

3.4 The minimal pair criterion

Some children, like Stoel-Gammon's Daniel (Stoel-Gammon & Dunn 1985), provide little or no evidence that adult-like features or phonemes play a role in their early words, because the contrasts among the first few words are so far from minimal. Daniel's earliest words are wildly orthogonal to one another; at 12 months, he had just these four:

(1) *banana* [nænæ]

(2) *light* [(d)ai]

(3) *uh-oh* [ʔʌʔo]

(4) *what's that* [wəsæ]

Such cases are quite common in lists of children's first few words; see Appendix I, Group 1, where 20 children learning one of eight languages – nearly half of our sample – provide first-word data that resist any straightforward analysis in terms of phonological contrast. From these examples, we derive two important implications: First, the data afford no pre-theoretical basis for determining that the locus of lexical contrast is at any sublexical level, whether consonant features, vowel features, or some aspect of the prosodic structure, although if one of these dimensions is taken to have priority over the others, it may be possible to describe the words as contrasting in that dimension. Second, the child who lacks minimal or near-minimal pairs cannot be shown to be using sublexical dimensions to facilitate the construction of new words; nothing in her behavior indicates that she has yet succeeded in generalizing any aspect of word production from one word to the next (beyond the bare ability to carry out phonation and articulation simultaneously).

If we were to write rules to derive these forms from adult-like targets, or constraints to permit their emergence, we would have to use most of the features required by a full description of English without the slightest evidence that the child can even re-combine the segments to make other words, let alone being able to recombine any of the features to make other phones/phonemes. In short, features postulated as a basis for such rules or constraints are indeed artefacts of analysis, because there is nothing in such a child's production behavior to indicate that he has made use of any sub-word resemblances.

At the other extreme, some children – like Virve at 10 months (Vihman 1976), Leslie at 11 months (Ferguson, Peizer & Weeks 1973), Hans at 14 months (Lindner 1898, cited in Ferguson 1978) and Jonathan at 15 months (Braine 1974), all of whom show tidy phonemic contrasts – seem to have tuned into features from the start of speech. In Section 5.1 we explore further the gradient between these extreme types of early phonology, and in Sections 3 and 5 we consider the

theoretical problem of deciding when analysis in terms of features is justified for a given child.

3.5 Gestural vs. feature analysis

Another important way of looking at feature-like phenomena in the acquisition of phonological production capacity is to consider which of them might be explainable in terms of gestures and their coordination, as defined in gestural phonology (Browman & Goldstein 1989, 1991, 1992). Gestures are movements of articulators (including the vocal folds and the velum), so place of articulation, nasality and the vocal-fold-approximation aspect of voicing are gestures (or complexes of gestures) as well as being features (at least in some systems). But other descriptive features, such as [continuant] and [vocalic], are not also gestures. In Danny's rule deleting pre-nuclear liquids, glides, and fricatives, 'gesture' cannot be a candidate for the shared property, because these sounds have no one articulatory movement in common. So this rule has a featural or a perceptual basis but not a gestural one.

For early words, however, atomization into basic gestures is comparable to analysis into features. The children who have the wildly orthogonal patterns seem to have acquired a word-length complex of gestures as an unanalyzed whole; those who have the tidy systems under a feature account have necessarily learned to recombine or modulate their articulatory gestures. Most children follow a path that lies between these extremes; gestures may better characterize their preferences than features, but a discussion of this here would take us too far afield.

4. Order of emergence and the effect of the ambient language

The order of emergence (of features or of phonemes) is now known to be modified by patterns of the ambient language, so even aside from individual differences, the invariance idea cannot be sustained in a strong form, nor can such modifications of it as were proposed by Rice & Avery (1995; but see now Rice 2007).

The earliest ambient language effects on production have been revealed by acoustic analyses of the vowel space sampled in the babbling of infants exposed to British English, French, Arabic and Cantonese (at 10 months: Boysson-Bardies, Hallé, Sagart & Durand 1989). Similar production effects are reported for prosody (rising pitch is more common in the babbling of French than of American infants in the age range 6–12 months: Whalen, Levitt & Wang 1991) and for consonants (a larger proportion of labials are found in the vocalizations of 10-month-olds exposed to English and French than in those of infants exposed to Japanese or Swedish: Boysson-Bardies & Vihman 1991). In each case the differences reflect a

biasing of the child's output in the direction of the ambient language while production nevertheless remains within the expected range for infants.

Consider some cross-linguistic data relevant to an emergence story for a prosodic feature, consonant length. The differential production of geminates or long consonants by children exposed to Finnish, with its categorical long: short contrast, can be compared with that of infants exposed to Welsh, in which consonants are automatically (i.e. non-contrastively) lengthened medially after a stressed syllable falling under nuclear accent. Children exposed to Welsh actually hear more long consonants in input speech, since lengthening is associated with the prosody of running speech in Welsh, whereas in Finnish long consonants occur in only about a third of the content words mothers use with their children (Vihman & Kunnari 2006). As would be predicted on the basis of frequency of exposure, then, Welsh infants produce longer consonants on average than do Finnish infants in their babble and first words. But by the time that the children reach the end of the single word period the situation has reversed, with the mean length of medial consonants virtually doubling for the Finnish children, while the Welsh children show little change from the earlier stage (Vihman 2009; Vihman & Kunnari 2006).

This can be ascribed to two factors, neither of them supported by a markedness account (which would presumably predict that geminates are learned relatively late). First, long consonants come naturally to children, whose articulation is sluggish; second, the geminates in adult speech, when lexically licensed, are highly salient (Vihman & Croft 2007). Once one or two of the common words with geminates (such as Finnish *anna* 'give', *kukka* 'flower', *loppu* 'finished, all done', *pallo* 'ball', *tyttö* 'girl') have been produced, production as well as perception experience will lend further salience to words with long medial consonants in Finnish despite their relatively low overall frequency of occurrence in the input speech stream. That salience can be considered to be due to 'top-down' (cognitive, lexical) processing in contrast to the 'bottom-up' or purely 'signal-based' salience of long medial consonants in a language like Welsh, where they occur more frequently due to their function as phonetic markers of accent. Thus we find that with lexical advance the Finnish children begin to show the bipolar pattern of the adult language, but with overproduction of geminates in comparison with the adult model.

Ambient language frequency effects are also documented for fricatives and liquids: They both tend to occur in early words in French, which has 21% liquids in running speech and also 21% prevocalic fricatives in content words, but not in English, which has 16% incidence of each (Vihman & Boysson-Bardies 1994); see also Pye et al. (1987) for liquids in Quiché (K'iché Mayan).

An ambient-language phenomenon that is even more troublesome for some acquisition theories is the loss, within the single word period, of an early sound for which the ambient language gives little or no support. Consider [h], a phone

whose featural description itself has often been a matter of contention. Although this sound occurs very marginally in adult French, it was produced to as great an extent by prelinguistic French babies as by the English, Japanese and Swedish babies participating in one study (with five infants in each group). However, only the English and Japanese babies, more than 10% of whose word targets included onset /h/, continued to include [h] in their word forms by the end of the single word period (Vihman 1992).

5. Our proposal: Features as emergents from first words

5.1 Micro-level analysis of features in a first typological gradient: Continuum in evidence for feature use in first words

We have suggested that the simplest argument against features as primitives is the large proportion of children whose early words differ from one another by multiple contrasts, sometimes by all of the factors of consonant position, consonant manner, vowel choice, and prosodic structure. As we noted for Stoel-Gammon & Dunn (1985)'s Daniel, in examples (1) to (4), above, each word is a thing unto itself, with little or nothing – neither features nor segments – carrying over from one word to the next. We have also noted that there are children like Ferguson et al.'s Leslie at the opposite extreme: each word in their first six is a member of a minimal pair. Overall, when we look over a reasonable-sized set of child data, we find a complete gradient of early phonologies from the 'each word unique' extreme to its opposite – a gradient from 'maximal dispersion' children like Stoel-Gammon's Daniel to 'maximal use of distinctive features' children like Ferguson et al.'s Leslie.

Appendix I provides an informal typology of this gradient of the first four to six word attempts for children speaking Dutch (N = 2), Estonian (4), Finnish (3), French (4), German (3), Italian (4), Japanese (5), Swedish (3) and Welsh (2) in addition to English (1 UK, 16 US); this includes the 27 children in Vihman 1996, Appendix B, plus any other children whose data were readily available to us. Appendix I divides the children into four groups, based on the orderliness or evidence for feature use provided by the child's first word forms. Children are listed by age at first words within each group, while words are listed alphabetically; finer temporal ordering information is available in some but not all cases.

Group 1 includes only children who lack minimal pairs entirely and whose words seem to contrast along multiple dimensions, like Stoel-Gammon's Daniel, who is by no means the most extreme. In each case the child seems to have learned each word as an unrelated item, with little or no reliance on similarity of onsets or vocalic nuclei. In Group 1a (N = 4) the children show little if any reuse of

segments and have no near-minimal pairs. Groups 1b and 1c show progressively more repeated use of segments and near-minimal pairs. Note that, if we give equal weight to a second (stressed) syllable, Stoel-Gammon's Daniel now looks relatively systematic, compared to the other 19 children of Group 1.

Group 2 data (N = 8 children) suggest contrast along two or three dimensions, based on repeated use of one or two segments in the same phonetic context but with no minimal pairs or too much variability to establish contrast rigorously.

Group 3 is the first in our typology to show minimal pairs based on rigorous criteria. These children look as if they were working with a recognizable phonology from early on. Group 3a children (N = 8) elaborate single vowel contrasts first. (Note, among these, the number who do *not* start with either a high – low or a high back – high front split, the vowel contrasts that Jakobson predicted would be first.) Group 3b children (N = 8) elaborate consonant contrasts first or alongside vowel contrasts. Group 3c includes only one child, Braine's Jonathan, who takes one consonant as an anchor and builds his vocabulary by elaborating his vowel system quite systematically.

Group 4 includes only three children; these are the truly impressive systematizers, who provide the best evidence we can find of clear featural organization, based on the child's use of two or more minimal pairs. This provides a valuable benchmark for assessing feature use, but although the three children were learning three different languages and began their word learning at three different points in time (with Virve and Leslie getting a start within the first year while Hans began at 14 months), the one thing shared by these data – and also by those of the only child in Group 3c – should give us pause: All of them derive from diary studies, unsupported by either recordings or acoustic analysis. The variability we see in the first word forms of children whose data were audio-recorded (see Section 6), suggests that the Group 4 word lists may reflect the (unintentional) tidying up of data by linguist parents or, in the case of Leslie Weeks, a linguist grandparent. This does not invalidate them completely, but it must cast some doubt on their generalizability.

An instructive contrast is provided by the case of another diary account, conducted this time by not one but two academic parents – namely, the remarkable study by Labov and Labov (1978), who exhaustively transcribed the speech of their fifth child, Jessie, for her first six months of identifiable word use. Here, the contrast between the first two long-lasting words was initially multidimensional: the words differ in consonantal nasality, glottality, and voicing, and over time, drew *further apart* in vowels (a low front vs. low central opposition, very unlike the pattern of the world's languages) and in number of syllables.

We have not, of course, exhausted the possible ways of supporting feature-based analysis of early child words; there are no doubt many viable arguments, and as we will see in Section 5.2, there will be correspondingly more of them for

children past the initial word-learning period, as more systematic phonology begins to emerge from the earliest whole-word production (based on lexical or item-learning). Data-rich studies since the early 1970s, however, support only some kind of gesture and some kind of long unit (whole-word or motor-sequence) as a viable early descriptor for most children's first word forms; we will not repeat those arguments here, but refer readers to Menn 1983; Waterson 1978; Macken 1995; Vihman 1996, Vihman & Croft 2007, Vihman and Vihman, in press, and the data in Appendix I. This point has now been recognized by some OT acquisitionists (e.g. Fikkert & Levelt 2006, who also note the relative accuracy of place of articulation of early words and describe the emergence of constraints from early lexicons in their data set).

5.2 Transition to a more orderly state: The emergence of phonological structure

A general observation is that at some point between the use of 30 and 70 different adult-based words, based on a diary record, or after the use of 25 or more different spontaneous words in a recorded half-hour free play session in the home (the '25 word point' in longitudinal studies: Vihman & Miller 1988), the mappings from adult to child forms may become more regular, the child's own forms may become more similar, and the rate of production of new words is likely to increase. Vihman and Velleman characterize some aspects of the evidence for emergent systematicity in this period: "When…the child begins to exhibit consistent patterning in the production of different adult words, including the distortion of some words to fit them into the child's individual production template, we identify this as the first evidence of phonological organization" (2000a, p. 312). That is, it now becomes critical to consider both the child's match to target (or lack of it) and the regularity or apparent systematicity of the word forms she produces. Vihman and Velleman (2000b) use the terms 'selected' (for words that show a good match to target and thus seem to be 'selected' for their pronounceability within the limits of the child's vocal resources) and 'adapted' (for words that are formally similar to the 'selected' words but that show striking deviations from the adult target; see also Vihman & Croft 2007).

Glossing over some details and exceptions, this trend is clear enough to make the basic question urgent: How do the children whose first five to ten words seem to go off in all directions become more systematic? And how does this increasing systematicity relate to the emergence of features?

There are several aspects of phonological systematicity to consider, not all of which need develop simultaneously. Here we identify three: emergent evidence for analysis into segments, systematicity in the mapping, and structural systematicity.

We consider the relationship of each of these to the emergence of features and the ways in which the different manifestations of systematicity are themselves related. Fortunately, ample published materials and discussions have long been available to support this comparison (e.g. Waterson 1971, 1978; Menn 1971, 1976/78; Priestly 1977; Macken 1978, 1979). For the purposes of illustration we will here focus mostly on data from Danny (Menn 1971) and Jacob (Menn 1976/78).

The three aspects of systematicity that we consider are interrelated, but only in indirect ways. Their histories vary, depending on the extent to which the individual child is a 'selecter', like Menn's Jacob – that is, a child who generally maintains the early preference for saying words that s/he can closely approximate, or an 'adapter', like Danny, whose heavy use of consonant harmony led to a sharp spurt in expressive vocabulary.

A. Analysis into segments: Minimal pairs are increasingly in evidence as the child comes to combine consonants and vowels more freely and generally increases the number of output word types. This advance supports analysis into segments and also analysis into features (subject to the questions about possible 'long units' that also apply to adult language). In an autosegmental framework, Macken (1992) considers much of this to reflect the development of planar segregation.

Data: Jacob shows two probable minimal pairs by the point at which he has produced 8 different words, in spite of token variation. The pairs contrast [±nasal] in 'no' vs. 'toast' [nʌː], [ŋɛʌ]/[dœʌ] and contrast [±round] in 'there' vs. 'toast' [dɑ], [dʌm], [dʌh], [dɛ], [dæ]/ [dœʌ]; unfortunately, there is only one tape-recorded token of 'toast' this early. Danny at 10 words contrasts [±nasal] in 'banana' [næ] vs. 'Daddy' [dæ] and at 19 words shows [±continuant] when his new word 'giraffe' [wæf] contrasts with 'pacifier' [bæf].

But analysis into features and analysis into segments may develop separately, as we have been arguing. We will return to this topic below, in our concluding section, and we give just one more illustration here: Jacob at almost 70 words of active vocabulary had fairly free combination of C and V in CV syllables, and a syllable-initial d/t voicing contrast for alveolar place, but at labial place he had only [b], and at velar, only [k]. So a segmental analysis is well supported but a featural analysis is not. Optimality Theory can describe a configuration like this with the right ordering of markedness constraints against segments and features, but it would not predict that segments would become well-defined before features do.

B. Systematicity in the mapping from targets: The mappings between a child's word and its adult target become less idiosyncratic and more amenable to description

by rules or constraints. For Danny, an 'adapter' with a strong and rule-governed pattern (except when his two place assimilation rules conflict), mapping systematicity seems to increase as structural systematicity increases. But for Jacob, an almost pure 'selecter' until quite late in his vocabulary growth (and never a strong adapter), mapping was loosely systematic from the beginning, basically being a matter of approximating the adult target. And for Waterson's son, an 'adapter' but a template-matcher rather than a rule-user, systematic mapping must have been impossible until segmental organization began to supersede his idiosyncratic templates (cf. also Si, Macken 1979).

Features seem to be needed for describing this emergent systematicity in mapping from the adult form to the child's form, but as always, the features we use to write rules or constraints to describe the mapping may not be those used by the child. For example, when Danny – beginning with *light* at 10 words – starts to systematically omit all onset glides, liquids, and fricatives (except his fossilized earliest [h]'s in *hi* and *hello*) but not nasal or oral stops, there is a choice of features to use to describe this rule/markedness constraint elevation (which was probably his first move away from selection). Depending on your preferred feature system, glides + liquids + fricatives may or may not constitute a natural class. Danny appears to have unified them, so if we impute an underlying feature system to his observable behavior, we could set up a feature set in which these sounds group together, say, [–vocalic, +continuous]. (If one's theory dictates a universal feature set, that might limit the freedom to follow a child's behavior in this way, but if one takes an emergentist stance, there is no such limitation.)

C. Structural systematicity: We find clearer organization into templates (or sets of canonical forms), or a growing core of words that resemble one another, although 'marginal' forms generally remain (the child's grammar leaks).

Data: For Danny, this organizational advance seems tightly connected with the abrupt onset of consonant assimilation at 30 words; for Jacob, it is a very gradual evolution from the beginning of his speech, because his early output divides clearly into two classes: his two complex targets, *thank you* and *Jacob*, which were almost always produced as two syllables with extraordinary consonant and vowel variation, and his increasing set of simple targets, produced as CV monosyllables beginning with his second word, *toast*, and expanding slightly to CVn. For Waterson's son P, structural systematization into templates like ⟨ŋVŋV⟩ (used for *Randall, window, finger*…) follows a period of slow whole-word learning, and seems well in advance of analysis into segments.

The properties of words that affect which canonical forms they are assigned to may be prosodic, segmental, or – as in Jacob's case – both (cf. also the data in Priestly 1977, where the child's choice of target consonants to include in his fixed CVjVC template provides a parlor game for phonologists). Similar analytic challenges can be found in the 'long word' harmony patterns arrived at by Laura [Lleó 1990] and Virve [Vihman 1978]; both are summarized in Vihman (1996, ch. 9). To the extent that these properties can be described in terms of phonological features, they are evidence for those features as distinguishers, as in Cristiá et al. (2008).

Over time, as consonants with different places of articulation start to co-occur in the same word, organization into segments overtakes a child's templates, which weaken or fade to the point where they simply reflect the phonotactic patterns of the ambient language (Priestly 1977; Waterson 1978; Macken 1979; Oliveira 2008; Vihman & Vihman in press).

One additional aspect of systematicity, *language specificity*, does not seem to be linked to the three that we have focused on here, because the effects of the phonotactic patterns of different ambient languages become evident at different times – depending on the level of challenge posed by those patterns for infants and also on the notable individual differences across children (e.g. some children acquiring languages with closed syllables show them quite early, others quite late: Vihman & Boysson-Bardies 1994).

Data: Danny has several real closed syllables by 13–14 words ([æp] 'apple', [gæk] 'cracker'); Jacob has a fairly decent final [n] after a couple of weeks of saying *down* as his sixth word; but weekly observations of Timmy (Vihman 1996, Appendix C), also acquiring English, showed that he was barely beginning to produce a final nasal at the 25-word point, which, as noted above, corresponds to a cumulative lexicon of about 50–70 words.

6. Variability based on recordings

For most of the children for whom we have transcriptions based on intensive audio recordings (such as those in Appendix II), we see that phonetic systematicity emerges gradually (but not necessarily monotonically): The range of variation in the way an adult phonetic target is realized is gradually reduced. In some cases, it seems as though another type of systematicity also needs to emerge: the range of variability needs to become more standard from one word to the next, so that it becomes justifiable to speak of phones that are comparable across words

(Ferguson & Farwell 1975). Emergence of phones in this sense is still a problematic area to study, because it requires recording multiple tokens of a number of phonologically similar words. This is the only way that we can establish whether or not a speech sound is stable when it recurs in different lexical items. Two of the three densest available data sets – the 1977 reprint of Ferguson and Farwell (1975), which includes the appendix omitted from the published version in *Language*, and the appendix of Menn (1976/78) – probably cannot sustain the statistical analysis necessary for rigorous support of the claim that the variation in the phonetic forms of the early words is lexical rather than random. For example, the Jacob corpus for the period 12;8–14;21 has 18 tokens of *down* with a wide range of vowels and, in the same time frame, four tokens of *(a)round* with a much smaller range. Some kind of statistical test is needed to know whether the [æʊ] of *down* was intended as the same phone as the [æʊ] of *(a)round*, because the apparent difference in range could simply be due to our having a smaller sample for *(a)round.* The raw data for Labov and Labov (1978) are probably the only longitudinal data set that could support an analysis capable of distinguishing lexical from random variation.

However, let us consider variability as we typically find it. Appendix II presents the first recorded words of nine children acquiring UK English in York, with all of their variant forms. The words are drawn from the first session in which the children, who were recorded one to four times a month from 9 to 18 months, produced four or more identifiable words spontaneously (but no more than six words: for the purposes of comparison with the 'first words' listed in Appendix I, many of them from diary studies, we exclude here children whose first recorded session with as many as four identifiable words already had seven or more such words).[3] The extent of variability differs from one child to the next but is consistently higher than the variation in published diary studies (where a single form is often provided for each word) or even in published accounts of observational studies based on recording and transcription. Let us see what kind of evidence of feature use might be obtained from these children's data.

The youngest of the nine children to produce four adult-based words in a recorded session, Ella (11 months), uses no supraglottal consonants with any consistency. Her three proper names seem to be only insecurely distinguished from one another: The core of all three is a VjV sequence, but a vagrant [n]

3. We exclude onomatopoeia lacking a conventional and stable adult target form. Only words produced one or more times spontaneously are included, but imitated tokens are also shown for those words, to provide as complete a picture as possible of the extent of each child's variability.

makes its way into tokens of both *Eva* and *William,* though not of *Amelia* – which, on the other hand, has no disyllabic tokens, unlike the other two, suggesting a distinct prosodic representation at least. *Hat* and *hiya* make use of glottals only, while one token of *hello* as well as two tokens of *row row* include an approximant 'r'. At 13 months Flora, similarly, has two words based on glides or glottals, while her other words (actually homonyms used for four adult words, *oh dear, star, ta* 'thanks' and *there*) share an onset [t] but differ in prosodic shape. Leila, at 14 months, uses different supraglottal consonants in each of her words but has nothing approaching a minimal or near-minimal pair, since the prosodic shapes for each word are also distinct. Finally, Patrick, at 16 months, uses completely different sounds in each of the words he attempts ('maximal dispersion' again). None of these children show any sign of feature-based organization or knowledge.

Rachel, 14 months, targets and/or produces fricatives in three of her first four recorded words; there is no evidence that the word forms are interrelated in any systematic way. In contrast, Lewis, age 15 months, makes repeated use of velars in two of his words – both of them with velar targets – and [d] in the remaining three words. Here again there are no minimal pairs, although there is more than a suggestion of interword relationships or 'cross-talk' (Menn & Matthei 1992): The occurrence of an otiose velar coda in *all gone* suggests item-contamination or blending, since *clock* has the target velar at both onset and coda; similarly, the vacillation between coda [s] and [t] for *duck* suggests a possible influence of *this* and *that,* as the child explores ways to deal with the classic difficulty posed by differing stop articulations in a single word shape.

Rosie, 15 months, has a simple reduplicative pattern for all of the words she attempts with supraglottal consonants, *mummy, teddy* and *tractor;* there is no real contrast, however, and one might predict an imminent merger of the latter two words. The variability shown by Tania, 17 months, in her productions of *book* with [b], [v] and [w] while *byebye* varies only between voiced and voiceless tokens, is reminiscent of the data that prompted Ferguson and Farwell's 'phone tree' analysis, designed to show the independent variability of the same consonant in different lexical contexts. Finally, Tobias, also 17 months old at the '4-word-point', shows a wide range of variant tokens for *baby* and a near-minimal pair to some of them in his productions of *digger*; his remaining words show re-use of some vowels but little between-word similarity otherwise.

What is the source of this quite considerable within-word variability across different tokens – within a single 30-minute recording session? Although the first answer that may spring to mind is articulatory or motoric immaturity, this would be more plausible if we saw even a hint of an age effect in the eight-month age-range

represented by the nine children, but we do not. The sources of this variability remain open to speculation but probably include variability in the input, instability in the speech production process – in motor planning as well as in control – and other factors relating to infant perception, attention and memory. We conclude, with Thelen and Smith (1994), that variability in production is part and parcel of the developmental process, wherever one looks at individuals over time rather than at groups. In fact, rather than expecting to see a steady fall in variability and rise in accuracy over the single word period, the evidence suggests non-linear trends in both of these dimensions of production. Here again, we should not be surprised, since 'controlled variability stands as the source of new forms in both real [i.e. 'on-line'] and ontogenetic time' (Thelen & Smith 1994, p. 134). Thelen and Smith remark further that 'in self-organization, the system *selects* or is *attracted to* one preferred configuration out of many possible states, but behavioral variability is an essential precursor ("order out of chaos")' (p. 55f.; emphasis in original), and finally, that 'infants' individual actions in context are the very stuff of development…[they] are selected as categories of knowledge from exploration of the inherent variability and noise of a biological system'. (p. 247)

In fact, at least one longitudinal case study, supported by acoustic analysis, was able to demonstrate a relationship between the increase in variability in a child's word forms and the move to a new 'attractor state' – specifically, the emergence of a child-based production pattern ('canonical form' or 'template'), which then stabilized production and facilitated new lexical learning for some period of time (Vihman & Velleman 1989; cf. also Vihman et al. 1994).

7. Concluding reflections: The forest and the trees

In this paper we have attempted to look squarely at the complexity of both theory and data bearing on the question of the initial stages of the development of features in children's speech production. There was no way to do this without getting down to the details – but what general conclusions can we draw now? Let us summarize what we have (and have not) accomplished.

First, we have not tried to define what it means for a child to 'acquire a feature'; instead, we have offered a way to understand the complexity of the notion of feature acquisition.

Second, we have offered a sample of the evidence on which we base our empiricist/emergentist stance; useful contrasting positions can be found in, e.g. Fikkert (1994), Macken (1995), Lohuis-Weber and Zonneveld (1996), Demuth, Culbertson and Alter (2006). Unlike these colleagues, we posit no stages: Note

that the '4-word point' is merely a heuristic designed to identify the relatively secure start of word use.[4] For any given child, postulation of a 'stage' should imply simultaneous advances on several fronts, whether reflecting qualitative reorganization due to an underlying cognitive shift, Universal Grammar-based parameter-setting, or something else entirely, but 'the boundaries of *progressive stages* are equally blurred by seeming regressions in performance and losses of previously well-established behaviors' (Thelen & Smith 1994, p. xvii [emphasis in original], speaking not of child phonology but of general principles of development).

Third, then, we argue against the notion that there are *universal* stages that might apply to all children, or even just to those learning a common language; the data that we have presented here may be sufficient to indicate why. Instead, we see each child as gaining knowledge, first, from distributional learning of many kinds in the period before the production of the first (recognized) words, and then additionally, in the period that we have focused on here, from the words s/he is using, out of which both units and patterns are gradually induced. Although we have spoken of 'selecters' and 'adapters', a child may vary in the extent to which s/he uses each strategy over time, perhaps because both vocal and perceptual resources change, as well as the child's cumulative lexical and phonological knowledge; all of these necessarily change the problem space for the child.

An emergentist position such as ours does not entail reducing universal tendencies to perception and articulation alone; grammar is clearly a matter of mental organization and is ultimately determined by the way the brain handles large quantities of complex and noisy data. Reduction to peripheral processing is especially untenable in the light of increasing evidence for top-down interaction with processing even in systems classically considered to be the most hard-wired (Li, Piëch & Gilbert 2004).

Fourth, we have tried to get closer to a concrete understanding of what it means for a phonological system to emerge (though we do not claim to have arrived at this point yet) and to present coherent arguments as to why phonology should be considered to emerge from its precursors. In development, nothing comes from nothing. If a researcher wants to call a particular apparent level of phonological (dis)organization pre-phonological rather than, say, pre-systematic, that is probably harmless – just so long as that researcher then takes the trouble to define what would be required for a system to qualify as phonological. The important general

4. Here again Labov & Labov (1978) is instructive: Jessie produced only two words consistently over the six-month period on which they report, although several other words appeared very ephemerally; it is unclear whether a '4-word point' might have been identified based on a single half-hour recording made at any time during that period.

point about phonological emergence is this: Fine-grained data show that the discontinuities of system-making are little ones: local, not global (this is the essential burden of Thelen and Smith's developmental model). And they need line up neatly neither with one another nor with other cognitive or developmental milestones. We cannot rule out dramatic invisible underlying changes, of course, but we hope that our focus on what is observable makes it clear what kind of thing would have to be postulated as intervening between any such invisible level and what can be observed. What we do observe is that, in a given child, the evidence for features, segments, and phonological organization accumulates incrementally. Phonological distinctions will propagate through different children's vocabularies in different ways, depending, we suggest, on the relevant aspects of structure that the child has already induced at each point in time.

We have presented so much data about variation that a reader might conclude that child phonology is a chaotic jungle, incapable of penetration by phonological theory: within child from moment to moment (token variation), within child over longer time intervals (developmental variation), and across children (individual variation, both within and across languages). We admit that it appears inhospitable, but this chaotic system, like many, does settle, albeit locally, lumpily, and gradually, into relatively stable and comprehensible systems at all levels.

On the time scale of a session or a week, we suggest that variation from token to token, in line with ideas proposed by Thelen & Smith (1994) and Blevins (2004), is like genetic variation: it provides a substrate from which selection can choose the most successful variants. Success is not a simple metric, however; tension among its several dimensions – including at least communicative effectiveness, perceived match to remembered adult target, and articulatory reproducibility – probably has much to do with the messiness of the token-level data that keeps turning up whenever we have multiple audio-recorded tokens of a given word.

On the developmental time scale, extending well beyond the initial period of word production that we have detailed in this paper, we see the emergence of phonology: the gradual systematization of various-sized portions of the lexicon, as described in the bulk of the child phonology literature since 1970. For 90% of the children in our sample, phonological features gradually emerge after the initial words are produced, under the four reasonable versions of the six criteria that we set out in Section 3.1. If features are used to describe the sets of words produced by the children in our first few groups, where there are no similarities from one word to the next, they are indeed artefacts of description, but features used to describe our most systematic children seem to be as well justified as they are for any adult corpus (with the caveat, again, that we have no recordings against which to verify those well-behaved transcriptions).

Are features inherent? (We are deliberately avoiding the word innate.) We have argued that they are the result of an intricate interplay between the auditory-acoustic input signal (with its variations across speakers and noise conditions as well as across languages and cultures), the child's developing cognitive capacity (which induces system, yet also retains cumulative auditory statistical information about input tokens – not raw, but as perceived – plus multisensory statistical information about output tokens), and articulatory capacities, which also develop over time. Yes, of course they are inherent (biologically grounded, as opposed to arbitrary) – but as emergents of this complex chaotic system, not as pre-experiential cognitive givens. We claim that they become part of a mental grammar as they are discovered by the speaker, becoming more and more fully realized as they come to be more stably represented in production. There may be theoretical elegance in holding otherwise, but there is (at least so far) no good evidence for it.

Appendix I. First words

The convention often used for English – use of voiced stop symbols to represent unaspirated voiceless stops – is not followed here. Voiceless stop symbols are meant to reflect 0 VOT, more or less, with use of voiced stops for pre-voicing and a raised [ʰ] for perceptible voicing lag. However, for data taken directly from sources that do not mention aspiration in languages that have allophonic aspiration of voiceless stops, the VOT value of initial stops transcribed as voiceless is unknown.

Group 1. No minimal pairs

1a. Maximal dispersion: no minimal or near-minimal pairs; no clear reuse of syllables or segments across items.

Alice (Vihman, Velleman & McCune 1994): **American English, 9–10 months**

[beɪbi]	*baby*	[pɛpɛː], [tɛɪtiː]
[dædi]	*daddy*	[dæ]
[hɑɪ]	*hi*	[haːiː], [ʔɑːjɛ], [hɑɪje] [haɪjʌ]…
[mɑmi]	*mommy*	[m̩ːɑnːə]
[noʊ]	*no*	[njæ̃]

Joel (Savinainen-Makkonen 2007): **Finnish, 13–15 months**

heihei [heihei]	*bye-bye*	[eiei]
anna [ɑnːɑ]	*give*	[ɑnːɑ]
äiti [æiti]	*mother*	[æti]
loppu [lopːu]	*allgone*	[əpːu]
vettä [ʋetːæ]	*water (partitive singular)*	[etːæ]

Taro (Vihman unpublished data): **Japanese, 14–15 months**

[bɯːbɯː]	*car (nursery word)*	[βeːβɪː], [əβeːβɪː]
[haitːa]	*we got it, it's gone in*	[tʌːtəʔ], [dʌːtã]
[ija]	*no*	[ɪjæʔ], [ijɯː], [ĩːjə̃]
[joiʃːo]	*oof!*	[ʃɻ], [ʃɻʔʃ ɻʃ̬ʔ], [ʃɻʃ̬]
[tsɯita]	*(switched) on*	[cidːæʔ]

Haruo (Vihman unpublished data): **Japanese, 15 months**

[haitːa]	*we got it, it's gone in*	[hʌtːʰaː], [atːa], [hatʰa]
[goɾogoɾo]	*sound of rolling object*	[ɣʰɔɻɣɔː], [ɣəɣə], [ɣəɻəɻəːː]
[kɯkːɯ]	*shoe(s)*	[kəkːə]
[mimːi]	*milk (request drink)*	[mɛɪβɪːʔ], [mɛːβɪʔ], [memɛʔ], [wɛːwɛʔ], [meːmeʔ], [mimiʔ]
[naɪnaɪ]	*all gone, no-no*	[nəʔ], [nəʔnaje], [njæʔjæᵒ]

1b. Moderate differentiation: Some reuse of segments across items, or near-minimal pairs suggesting possible contrast.

Molly (Vihman & Velleman 1989): **American English, 10–11 months**

Reuse of initial labials.

[beɪbi]	*baby*	[bæpæ]
[kˌɹækəɹ]	*cracker*	[pɑkæ], [kwɑ], [wæʰk], [pækwɑ], [kʌk]
[muːː]	*moo*	[meʔje]
[nɑɪʔnɑɪt]	*night-night*	[hʌnːʌ], [noʊnæ]

Anna (Keren-Portnoy et al. 2009): **Italian, 10–13 months**

Near-minimal pairs contrasting labial/alveolar place ([be], [den]), front/back mid vowels ([be], [bom]), and oral/nasal stops ([bom], [mam]).

bambola	*doll*	[bombœ]
bebè	*baby*	[bebɛ]
caffè	*coffee*	[kakɛ]
cavallo	*horse*	[kakːo]
dindon	*ding-dong (bells)*	[dende]
mamma	*mama*	[mamːa]

Joan (Velten 1943): **American English, 11–12 months**

Repeated use of [ba] and [za]; no near-minimal pairs.

[bæŋ], [bɑɾəl]	*bang, bottle*	[ba]
[bʌs], [bɑks]	*bus, box*	[bas]
[pʊɾɔn]	*put on*	[baza], [baːza]
[ðæt]	*that*	[za]
[ʌp]	*up*	[ap]

Sean (Vihman & Kunnari 2006): **American English, 12 months**

Near-minimal pairs contrasting front and central low vowels ([dæ:], [tak]) and mid -low front vowels ([tak], [tɛ^{h}]).

[ɑlgɔn]	*allgone*	[ʔɔdæ:]
[bu:]	*boo*	[pʊ]
[dɑg]	*dog*	[tak]
[tɪk]	*tick*	[tɛ^{h}], [tɪʔ], [tɪ̰], [tʊ̰t]
[wʊf]	*woof*	[wʊ], [ʔʊʔ], [ʔoʊ]

Fflur (Vihman & Kunnari 2006): **Welsh, 13 months**

Near-minimal pairs contrasting three places in stops, stop-nasal at alveolar place and stop-glide at labial place ([ba], [da], [ga]; [da], [na]; [ba], [wa]) – but with too much within-word variability to consider any contrast as properly established.

agor [agɑr]	*open*	[gəga], [gag:a]
blodyn [blɒdɪn]	*flower*	[ʔɛbə], [bəba], [bəʊwa]
ceg [kɛg]	*mouth*	[gag], [gɛ:g]
golau [gɒla]	*lights*	[baʊwa], [ʔʌwa]
na [na]	*no*	[na:], [ɲa]
sannau [sana]	*socks*	[dənə]
sgidie [sgidʒɛ], [sgidijɛ]	*shoes*	[gəga], [dadæ], [dada]

Maarja (Vihman & Vihman in press): **Estonian-English bilingual, 12–14 months**

Near-minimal pair contrasting voicing ([tæ^{h}], [dæ:]), based on diary transcription; reuse of diphthong [ai].

aitäh [aɪˈtæh]	*thanks*	[aiˈtæ^{h}], [tæ^{h}]
auh-auh [auhˈauh], aua [ˈaua]	*bowwow; doggy (nursery word)*	[wɑwɑwɑ], [wʊwʊwʊ] (growly)
[ˈdædi]	*daddy (English)*	[dæ:], [dæ | i:]
kuku [kuku]/uh-oh [ʌʔo]	*peek-a-boo, uh-oh (English)*	[uʔu:], [ʌʔʌ:], [ʔo], even [m̩ʔm̩:]
mõmm-mõmm [mʏm:mʏm:], mõmmi [mʏm:i]	*teddy (nursery word)*	[mʏm:mʏm:], [mʏm:i]
pai [paɪ]	*nice (petting and patting word)*	[ʔaɪ], [daɪ]

Raivo (Vihman & Gathercole, in prep.): **Estonian-English bilingual, 13–14 months**

Apparent near-minimal pair, but with mismatch to vowels in target [ta] for /-tæh/ vs. [pæ], [bæ] for /paʎ/); reuse of [b] and syllabic fricatives, based on diary transcription.

aitäh [aɪˈtæh]	*thanks*	[ta], [taʔ]
ei [eɪ]	*no*	[ei]
pall [paʎ]	*ball*	[bæ], [pæ]

pömm [pʏm]	*boom*	[bm̩]
shoe [ʃu]	*(English)*	[ʃ], [ç]
viska [viska] (+ vesi [vɛsi])	*throw*	[is], [iɬ], [ɬ̪], [ʂ]

Kazuko (Vihman unpublished data): **Japanese, 13 months**

Near-minimal pair contrasting labial stop and glide ([pa], [wa]), with a (syllable-level) voicing contrast as well.

[baŋ], [bã]	*bang*	[p^hą̊], [b^uæ̃], [p^hã]
[bɯː]	*car sound*	[βɯɯː:]
[nen:e]	*sleep (nursery word)*	[næn:ɛ], [næ:nɛ::]
[wãwã]	*doggie*	[wɑ:wɑ], [uwɑ:wɑ], [wɑ:w:ɑ]

Nina (Keren-Portnoy et al. 2009): **Italian, 13–18 months**

Reuse of labials.

baubau	*bow-wow*	[ba:ˈba:]
mio	*mine*	[mio]
mamma	*mama*	[mem]
zia	*aunt*	[ia]
caffè	*coffee*	[aɛ]

Daniel (Danny) (Menn 1971): **American English, 20 months**

The phonetic variability of forms for *byebye* means that contrast between those disyllabic forms and the monosyllable form for *squirrel* is best viewed as prosodic, not segmental. Diary study, no tape recording.

[gʊbɑɪ], [bɑɪbɑɪ]	*goodbye, byebye*	[bæbæ, baba, gægæ]
[hɛloʊ]	*hello*	[hwoʊ]
[hɑɪ]	*hi*	[hæ, haɪ]
[noʊ]	*no*	[o^{n}o, no, nu]
[noʊz]	*nose*	[o]
[skwəɻl]	*squirrel*	[gæ, goʊ]

1c. Greater differentiation: More use of repeated segments or syllables, or contrasting syllables – but still no minimal pairs

Laurent (Vihman 1993): **French, 10 months**

Reuse of the syllable [ljo]; near-minimal contrast at alveolar place.

allo [alo]	*hello*	[hailo], [ailo], [haljo], [aljo], [alo]
donne (le) [dʌnlø]	*give (it)*	[dlə], [də], [ldɛ], [heldɔ]
l'eau-l'eau [lolo]	*bottle (nursery word)*	[ljoljo]
non [nɔ̃]	*no*	[ne]
tiens [tjɛ̃]	*here, take it*	[ta]

Daniel (Stoel-Gammon & Dunn 1985): **American English, 12 months**

Nasal/fricative contrast in second syllable (stressed in target).

[bənænə]	*banana*	*[nænæ]
[lɑɪt]	*light*	[ai], [dai]
[ʔʌʔoʊ]	*uh-oh*	[ʔʌʔo]
[wəsðæt]	*what's that*	*[wəsæ]

Noël (Vihman & Gathercole, in prep.): **French, 13 months**

Reuse of onset [p], coda [m].

coucou [kuku]	*peek-a-boo*	[tətə]
miam [mjɑm]	*yum*	[ʔam], [ʔʌm]
papa [papa]	*papa*	[pæpæ]
poum [pum]	*boom*	[pɔ̃m]

Elen (Vihman & Gathercole, in prep.): **Welsh, 13–14 months**

Reuse of [d] and [g].

[tʃutʃu]	*choochoo*	[tʃʊtʃʊə], [$t^hʌt^hʌ$]
gogalw [gogalu]	*bird (family word)*	[gak:u]
[kwækkwæk]	*quack-quack*	[gəgɒg]
[sɪt]	*sit (said to dog)*	[$s^hɪt$], [s:$ɪ^h$], [sɪ:t]
tân [t^ha:n]	*fire*	[d^ha:]
tata [tata]	*bye-bye*	[dɪdə]

Eelis (Vihman & Gathercole, in prep.): **Finnish, 18–19 months**

Some reuse of both vowels and consonants.

äiti [æiti]	*mother*	[æiti:]
anna [an:a]	*give*	[an:a], [æn:æ]
ei [ei]	*no*	[ei]
heppa [hɛp:a]	*horsie*	[bap:a]
kiikkuu [ki:k:u:]	*swinging*	[kik:u], [ka:k:u]
kukka [kuk:a]	*flower*	[ka], [kak:i]

Joost (Lohuis-Weber & Zonneveld 1996): **Dutch, 20–21 months**

Reuse of [k]; contrast of coda [p], [k]; near-minimal pair [pa]/[mɔ].

kikker	*frog*	[kɪ]
klok	*clock*	[kɔk]
mond	*mouth*	[mɔ]
paard	*horse*	[pa]
stop	*stop*	[tɔp]

Group 2. Systematic building from two places of articulation; no true minimal pairs

Hildegard (Leopold 1939): **German-English bilingual, 10–12 months**

Reuse of labials and alveolars; near-minimal pair contrasting high/mid vowels at alveolar place.

[bɔl], Balle [balə]; [opa]	*ball* (Eng., German), *grandfather*	[pn̥]
[pɑpɑ]	*papa*	[pn̥pn̥]
[pɻ̊ɪti]	*pretty*	[tpr̥tɻ̊]
[ðɛəɹ], da [da]	*there (Eng., German)*	[dɛː]
Ticktack [tɪktɑk]	*ticktock*	[tɪta]

Luca (Keren-Portnoy et al. 2009): **Italian, 10–13 months**

Reuse of labials, oral and nasal, and of velars; near-minimal pair contrasting mid/high front vowels at labial place ([be], [bim]).

acqua	*water*	[akwa]
bebè	*baby*	[beˈbɛ]
bella	*beautiful, nice*	[beja]
bimba	*child*	[bimba]
mamma	*mama*	[momːœ]
occhio	*eye*	[aˈgo]

Timmy (Vihman, Velleman & McCune 1994): **American English, 11 months**

High variability precludes any categorical statement about contrast between *ball* and *block* or *car, kitty*, and *quack-quack*; nevertheless, Timmy seems to be building systematically around the labial and velar places of articulation, making it reasonable to claim that he has that place contrast.

[bɔl]	*ball*	[pæ], [bæ], [ʔəpæ], [abːa]
[blɑk]	*block*	[əpʰə], [ʔʌβæ], [pæ]
[kʰɑɹ]	*car*	[kɑə], [ɑkːɑʰ]
[kʰɪɾi]	*kitty*	[khn̥], [kʰn̥], [kɑkɑ], [ʔukɑ]…
[kʰwæʔkʰwæk]	*quack-quack*	[kʰn̥], [kʰɑ], [kʰɑʰkɑʰ], [gaga]…

Sarah (Stoel-Gammon & Dunn 1985): **American English; 11 months**

Systematic building from labial and alveolar places and also oral/nasal stops; reasonable claim of place and manner contrasts but no minimal pairs.

[beɪbi]	*baby*	[bebi]
[bɑɪbɑi]	*byebye*	[baɪbaɪ]
[dɑgi]	*doggie*	[dɔgi]

[dʒu:s]	*juice*	[dus]
[mɑmɑ]	*mama*	[mama]

Emily (Vihman, unpublished data): **American English, 13 months**

Near-minimal labial/alveolar contrast; reuse of labial.

[bæ:bæ:], [baʊwɔʊ]	*baa, bow-wow*	[pæpæ], [bæbæ], [ʔapɪæ], [pæ:]
[bi:dz]	*beads*	[bi], [p^{h}i]
[dædi]	*daddy*	[tæ], [hadatɛ]
[ʌp]	*up*	[ʌp], [ʌpə], [ʌpije], [æb]

Jacob (Menn 1976): **American English, 13 months**

Building from alveolar place, incipient alveolar/velar word-medial place contrast, reasonable claim of nasal/oral contrast but variability makes minimal pairs hard to define.

[noʊ]	*no*	[nʌ:::], [ŋɛʌ]
[dʒeɪkəb]	*Jacob*	[dikʌ], [dɛikʌ], [gɛikʌ], [æku], [dɛikʌ], [æku]
[θæŋkju:]	*thankyou*	[didʌ], [dɪdɛjdi], [ɬɛjʌ], [daˈza], [di], [daˈdʌ], [bɛ], [dɜt],[gɑdu],etc.; varied stress & pitch patterns; [dVdV] forms dominant but velars recur.
[ðɛɹ]	*there*	[dɑ], [dʌm], [dʌh], [dɛ], [dæ]
[t^{h}oʊst]	*toast*	[dœʌ]

Atte (Vihman & Kunnari 2006): **Finnish, 17 months**

Two-or three-feature-based contrasting pairs (*starred) for manner + voice (stop/fricative) and place (labial/coronal with stop/nasal).

anna [an:a]	*give*	*[na]
haua [hauva]	*doggy*	*[ʋa], [ha:ʋa]
heppa [hep:a]	*horse*	*[pa], [ap:a]
mummo [mum:o]	*grandmom*	[mo], [am:o]
pappa [pap:a]	*grandpa*	[pa], [pap:a]

Tomos (Vihman & Gathercole, in prep.): **UK English, 17 months**

Non-minimal contrast, labial stop/alveolar nasal.

[bædʒə]	*Badger*	[babm̩:], [bʌbm̩]
[bæŋ]	*bang*	*[ba], [bæ], [baʊ], [da]
[haɪja]	*hiya*	[jaja], [dajæ:]
[nəʊ]	*no*	*[na], [næ], [nə]
[ta]	*ta* 'thank you'	*[ba], [pa], [ba:], [ɹa:]

Group 3. Minimal pairs; contrast in vowels or in consonants

3a. One clear minimal pair (*): Vowel contrasts only.

Deborah (Vihman & Gathercole, in prep.): **American English, 10 months**

Re-use of labial place; one minimal pair contrasting mid/low front vowel (and with match to target contrast).

[bæ:]	*baa*	*[bæ:]
[beɪbi]	*baby*	*[be], [pipe], [bebe]
[hɑɪ], [hɑɪjʌ]	*hi, hiya*	[hai], [ai], [haie], [aie], [e:], [a:]
[mʌŋki]	*monkey*	[mam:ɛ]
[ʌʔ:oʊ]	*uh-oh*	[ʌʔ:ɛ]

Lina (Vihman 1996, App. B): **Swedish, 10 months**

Building system from labial; mid/low vowel contrast.

blomma	*flower*	[bɔmbɔ̃]
boll	*ball*	*[ba], [bʊ^{h}]
brum	*vroom*	*[bɛ]
bulle	*bun*	[bɯl]
oj	*oh!*	[ʔɔɪ]
titta	*look*	[tit:a]

T. (Ferguson & Farwell 1975): **American English, 11 months**

Building from alveolar and glottal articulations: High/low vowel contrast.

[dædi]	*daddy*	[dæji, dæi]
[dɑg]	*dog*	[dɔ]
[hɑɪ]	*hi*	*[ai], [hai]
[si:]	*see*	*[hi]

Charles (Vihman & Kunnari 2006): **French, 11 months**

Mid/low vowel contrast – but disputable because of variation in *boom* as well as lack of match to target.

au-revoir [ɔʀvwaʀ], [ɔvwaʀ]	*byebye*	[awa, haɥa]
boum [bum]	*boom*	*[ba, bœm]
beau [bo]	*beautiful, good*	*[bo]
donne, tiens [dʌn, tjɛ̃]	*give/here*	[dæ]
mama [mamã]	*mama*	[mama]
non [nɔ̃]	*no*	[nɛ]

Hanna (Vihman 1996, App. B): **Swedish, 11 months**

Building from alveolar place with contrast between front/central low vowels.

bok	*book*	[bβuʔ]
dar	*there*	[dæʔ]
tacktack	*thanks*	*[dad:a]
titta	*look*	[dit:a]
tittut	*peekaboo*	*[dæd:a]

Will (Stoel-Gammon & Dunn 1985): **American English, 12 months**

Systematic building from alveolar place; low/high front vowel contrast.

[ɑldʌn]	*all-done*	[dada], [ada]
[dɑʊn]	*down*	*[dæ], [dʌ], [dau]
[lɑɪt]	*light*	*[di]
[ʃu:z]	*shoes*	[tsis, θiz]
[ʌʔ:oʊ]	*uh-oh*	[ʔʌʔo], [hʌho]

Kenji (Vihman 1996, App. B): **Japanese, 12 months**

Building from alveolar nasal: apparent low/mid central vowel contrast – but variable production makes contrast uncertain; near-minimal alveolar/velar contrast.

[aɽigato:]	*thank you*	[ɑoi::], [ɑɪ::ja], [ʔɑɪə::], [ɪ:jɑʰ], [ɑwɪ:wɑ:], [əɪ::jə]
[bɯ:bɯ:]	*car (nursery word)*	[βɯ::βɯ::]
[do:zo]	*here you are, if you please*	[ədo::]
[koɽe]	*this*	[kɔ:jeʔ], [kolɪ], [koɛʔ], [kore], [gɔɾi::], [gəʔ],[korɪ], kɔɾʲɻ:]…
[nainai]	*all gone, put away*	*[nɛ:nɛ], [nɛ̃:nɛʰ], [nə: | ɲæ:], [nɑ:ɪnɑʰ], [nɛ:ne]
[njãnjã]	*kitty, meow*	*[næʔnʌʰ]

Eeriku (Salo 1993): **Estonian, 17 months**

Low vowel contrast, with match to target contrast.

tita [titɑ]	*little girl*	[tit]
onu [onu]	*uncle, man*	[en:]
paber [pɑbeʀ]	*paper*	*[paba]
päkapikk [pækɑpik:]	*elf*	*[pæpa]
suur [su:r]	*big*	[u:]
tiss [tis:]	*nursing (nursery word)*	[ts:]

3b. One or more minimal pairs, including consonant contrast.

Annalena (Elsen 1996): **German, 8–9 months**

Minimal pair involving oral/nasal contrast; near-minimal pair contrasting labial/alveolar place, including consonant contrast.

da	*there*	[da]
ei	*egg*	[aɪ]
Mama	*mama*	*[mama]
nein	*no*	[naɪn]
Papa	*papa*	*[baba]

Siri (Savinainen-Makkonen 2001): **Finnish, 10–13 months**

Labial/velar contrast as well as high/low vowel.

kukka [kuk:a]	*flower*	[ʔka], [aˈka]
kiikkaa [ki:k:a:]	*swinging*	[ki:k:a:]
vettä [vet:æ]	*water*	[et:æ]
nappi [nap:i]	*button*	*[pi], [əp:i]
kissa [kis:a]	*cat*	*[ki]

Nicola (Keren-Portnoy et al. 2009): **Italian, 10–16 months**

Oral/nasal consonant and low/mid vowel contrasts.

mamma	*mama*	*[mama]
bimba	*child*	[bœb:ɛ]
nonna	*granny*	*[nen:a]
nanna	*to sleep (nursery word)*	*[nan:a]
Vale	*(name)*	[ae]
papà	*papa*	*[pap:a]

Carole (Vihman & Kunnari 2006): **French, 11 months**

One pair contrasting voicing, another overlapping pair contrasting low/mid vowel, and a two-feature contrast of alveolar nasal/labial oral stop.

balle [bal]	*ball*	*[ba], [baba]
bébé [bebe]	*baby*	*[bebe]
nounours [nunuʀs]	*teddy bear*	*[ne], [nene]
Mickey [mike]	*Mickey Mouse*	[kə]
papa [papa]	*papa*	*[papa]

Jonah (Vihman 1996, App. B): **American English, 13 months**

Labial/alveolar contrast.

[bɑɾəl]	*bottle*	[bwɪdʊ]
[bɑʊwɑʊ]	*bow(wow)*	[bɑʔ], [bᵛɑ], [bæ]
[ɛdgəɹ]	*Edgar (dog's name)*	*[dɑdɑ]
[noʊ]	*no*	[ənæ::]
[ɹɑkɹɑk]	*rock-rock*	*[bɑʰbɑʰ], [bɑːbɑʔ]

Emi (Vihman unpublished data): **Japanese, 14 months**

Manner contrast: stop /nasal /glide at labial place; near-minimal labial/alveolar stop contrast.

[ɑtːa], [hɑitːa]	*we got it, it's gone in*	[a̰tʰa̰], [haʔta], [tæʰ], [ɑ], [ta̰tʰa̰ʔ]
[bɑ]	*peekaboo*	*[pa̰], [wɑʔ]
[mɑmɑ]	*mama*	*[mɑmːɑː], [mæmæ], [mæmːɑː]
[nenːe]	*sleep*	[nenːeʔ], [nænɑʔ], [ɛnæ], [nænɑ]
[wɑ̃wɑ̃]	*doggie*	*[wɑwɑːː], [wɑwɑʔ], [wʌwɑʔ], [wɔwɑʔ]…

Thomas (Elbers & Ton 1985): **Dutch, 15 months**

Minimal pair in coda position.

[autoː], [oːtoː]	*car*	*[at], [atə], [aut], [autoː], [oːt], [oːtə], [oːtoː]
[hap], [hapjə], [hapi]	*a (little) bite*	*[ap], [apə], [hap], [hapə], [hab], [habə]
[paːrt], [paːrtjə]	*hors(i)e*	[paːt], [paːtə], [baːt], [baːtə]
[pus], [pusjə]	*cat, kitty*	[pəx], [bəx], [pux], [bux]

Stig (Vihman 1996, App. B): **Swedish, 16 months**

Labial/alveolar contrast as well as near-minimal contrast of low/mid vowel at alveolar place.

blomma	*flower*	*[bæ], [bʰɔːbɔː]
(skol)buss	*(school)bus*	[kɯlbɯ], [gɛbɯkəç], [klmbæ]
dar	*there*	*[dæ]
klocka	*clock, watch*	[gɔːka]
titta	*look*	[tɪta], [tɛtə]

3c. Large minimal set: Vowel contrast.

Jonathan (Braine 1974): **American English, 15 months**

[hɑɪ]	*hi*	[ʔai]
[dʒus]	*juice*	*[du]
[noʊ]	*no*	*[do]
[siː]	*see*	*[di]
[ðæt], [ðɛɹ]	*that, there*	*[dæ, dʌ, da, dɛ]

Group 4. Two or more minimal pairs; clear featural organization

Virve (Vihman 1976; Vihman & Croft 2007): **Estonian, 10 months**; some exposure to English
Vowel contrast, mid/low; also stop/fricative contrast at alveolar place.

aitäh [aɪˈtæh]	*thanks*	*[ta]
[haɪ]	*hi (English)*	[aɪ]
isa [isa]	*daddy*	*[saʔ]
see [se::]	*this*	*[seʔ]
tere [tere]	*hello*	[tete]

Leslie (Ferguson, Peizer & Weeks 1973): **American English, 11 mos**

Every word a member of a minimal or near-minimal pair with respect to the features labial /alveolar and oral/nasal.

[dædi]	*daddy*	*[dædæ]
[dɔgi]	*doggie*	*[gaga]
[mami]	*mommy*	*[mama]
[pæti], [pædi]	*patty(-cake)*	*[bæbæ]

Hans (Lindner 1898; cited in Ferguson 1978): **German, 14 months**

Two minimal triples showing consonantal contrasts: monosyllables showing labial/alveolar/velar, and disyllables showing labial /alveolar and oral /nasal.

Birne [birnə]	*pear*	*[bap]
das [das]	*that*	*[da], [dada], [dat]
Gasse [gasə]	*street*	*[gak]
Mama [mama]	*mama*	*[mama]
Papa [papa]	*papa*	*[papa]
Wewe [veve]	*weewee*	[we:we:]

Appendix II. Variability in first word forms: UK English

(i) indicates imitated form; (x2) etc. indicates number of tokens of form.

Ella: 11 months

Amelia	ɪdʒi:ijaʰ	zi:ʔijaʰ	i:ʔɪjaʰ	i:ijaʰ	i:::hiæʰ (i)
Eva /i:vʌ/	i::jaʔ	i::jaʔ	i::jaʰ	ni:jæʰ	i::jæʰ (x3)
	aɪ:ija (x2)(i)	nijɪʔija (i)			
hat	aʔ (x2)				
hello	əɹaʊ	eʔaʊ:	eʔɬə	ɪhəʊ	
hiya	aɪjɛ	aʔaɪ:			

row row	ɪvəʉwə (i)	ɪbvɐuβæ (i)		m:ɹəʉ:u:wa	ɪbuɹau:wa:
	əwau:evəva	mβauʔvəu			
William	i::jɛ	ihija^{n}(i)	ungijæ (i)	i:jæ^{h} (i)	

Flora: 13 months

hello	ɛjəu	aɻjə (i)	ælə (i)	
oh dear, star, ta, there	ata	ætæ (x3)	æt^{h}a:	ətɪ:
	tæ:h	həʔtæ		
uh-oh	ʔaʔəu	ʔæʔæ:ə	ʔæʔa (i)	

Leila: 14 months

bump	bou	bə: (i)	bɪ (i)	p^{h}ɪ (i)	ba: (i)	bijab (i)
no	i:ɜ:nau	nəu	anɑ:	ŋou	ŋounə:u (i)	
		na:ou (i)	ənɑ: (i)			
that	daʔt^{h}					
tickle	ŋtəgɑwə (i)	kɪtʉɪkxkxdj	glʌlʌbli:glɪ	kɔtlʌð (i)		
		gʌ:ɯgʊldə (i)				

Rachel: 14 months

hiya	ʔəʔa:ɪja	heʒə:a^{h}	ɑ:jə^{h}	heɪdɜ: (i)	he:ɪjə^{h}	
juice	dæ	tsɜ:dɜ::	dɪdʊs:ʊ	t^{h}əs:	dʊs	təðʊs
	gʊs:	dəʔɜ:ʃ				
sheep	ta (x2)	dɛ:ɪ	tθa:	t^{h}ə̰	t^{h}ahə (i)	
this	dɪəʃ	dʊɪθ	i:hɪθ			

Lewis: 15 months

all gone	ɔ:kʲə:	ɔ::gʊ	ɔ:gə	ɔgɒk^{h} (x2)		ɔgə	əg:ɒk^{h} (i)
	ɔˈɣɒə	eʊk^{h}ɒk^{h} (i)	ɔ:gʊk^{h}	əgɒ::	ŋ gɒk^{h} (i)		əʊgɒ::
	ɔgɒ:^{h} (i)	ɔ:ʔgə	ɔ:k^{h}ək^{h} (i)	ə::gɒk^{h} (i)	ʔɔ:gɑk^{h}		ɔgɒk^{h} (i)
clock	k^{h}ɒ:k^{h}	k^{h}ɒ:ʔk^{h}	k^{h}ɔ:k^{h} (x2, i)				
duck(s)	dəs	dʒɑʔt^{h} (i)					
that	dɑ:t^{h}	daʔət^{h}	ndæ:t^{h} (i)				
this	ndə:ʃ	dɪ:ʊʃ:	t^{h}ɪs (i)				

Rosie: 15 months

mummy	məmp^{h}æ	mama	mam		
teddy	tɛtɪ	t^{h}ət^{h}ɪ	t^{h}ɛt^{h} (i)	dət^{h}ə (i)	
tractor	t^{h}ət^{h}ɛ:	tɪdə (i)			
yeah	ɛ	ja (x2)	jæ	jɛ (x8)	əjɪ (i)

Patrick: 16 months

baa	bɑ	baʔ (i)	
hello	ɛlɔːʰ	əɹə	
hiya	haɪ ji	aɪ jəʔjɛ	heɪʔjɛʰ
two, three, four	paː dʒeɪ fʷɑ	paː dʷeɪ dæ (i)	

Tania: 17 months

book	vəʊ	bʌː	wəʊ	bəʊː (i)	
bye(bye)	bajiʲjɪ (i)	baɪː	paiː (x2)	pæiː (x2)	
		pajiç (i)		bɛjʊɸɛɪ (i)	
daddy	tɛdɛ				
fish	tʉʃs				
mummy	mʔəməmæh				

Tobias: 17 months

(a) baby	bɛːbɪʰ	mbiːbɪç	bɪːbɪç	əbiːbɪʰ	
	beɪbiːʰ	nːːbiːbiç	biː	biːbiːθ (i)	biːbiːʃ
	biːːbɪʰ	biːbiːʰ	biːbɪʰ (x2, i)	biːbiʰ (x3)	bibiːːç
	bibiːç (i)	biːbiʃ	bɹiːbɪʰ	bɪbɪːʰ (i)	bɪːːbɛʰ
bear	nabɛː	beɪ bijɪː	bɛːʰ bɪʰ	pɪʔə	
digger	gigɪjə		di tʰɪʔɪ		
hiya	ɑɪjɪ	ɑɪjɪː	əɪjɛːʰ		
there/that	də	dəʔɛ			
uh-oh	ɛʔəː				

References

Beckman Mary E., Yoneyama Kiyoko & Edwards Jan. (2003). Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. *Journal of the Phonetic Society of Japan* 7. 18–28.

Berman, Ruth A. 1977. Natural phonological processes at the one-word stage. *Lingua* 43. 1–21.

Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns.* Cambridge: Cambridge University Press.

Boysson-Bardies, Bénédicte de, Pierre Hallé, Laurent Sagart & Catherine Durand. 1989. A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language* 16. 1–17.

Boysson-Bardies, Bénédicte de & Marilyn M. Vihman. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language* 67. 297–319.

Braine, Martin D. S. 1974. On what might constitute learnable phonology. *Language* 50. 270–299.

Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6. 201–251.

Browman, Catherine P. & Louis Goldstein. 1991. Gestural structures: Distinctiveness, phonological processes, and historical change. In Ignatius G. Mattingly & Michael Studdert-Kennedy (eds), *Modularity and the Motor Theory of Speech Perception,* 313–338. Hillsdale, NJ: Lawrence Erlbaum.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49. 155–180.

Cristià, Alejandring & Amanda Seidl. 2008. Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development* 4. 203–227.

Cristià, Alejandrina, Amanda Seidl & Alexander L. Francis. Phonological features in infancy. this vol.

Demuth, Katherine, Jennifer Culbertson & Jennifer Alter. 2006. Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech* 49. 137–174.

Elbers, Loekie & Jos Ton. 1985. Play pen monologues: The interplay of words and babble in the first words period. *Journal of Child Language* 12. 551–565.

Elsen, Helke. 1996. Two routes to language. *First Language* 16. 141–158.

Ferguson, Charles A. 1978. Learning to pronounce: The earliest stages of phonological development in the child. In Fred D. Minifie & Lyle L. Lloyd (eds), *Communicative and Cognitive Abilities – Early Behavioral Assessment,* 273–297. Baltimore: University Park Press.

Ferguson, Charles A. & Carol B. Farwell. 1975. Words and sounds in early language acquisition. *Language* 51. 419–439. Reprinted in William S–Y. Wang (ed), 1977*, The Lexicon in Phonological Change.* The Hague: Mouton.

Ferguson, Charles A., David B. Peizer & Thelma A. Weeks. 1973. Model-and-replica phonological grammar of a child's first words. *Lingua* 31. 35–65.

Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. Amsterdam: Holland Institute of Generative Linguistics.

Fikkert, Paula & Clara Levelt. 2008. How does Place fall into place? The lexicon and emergent constraints in children's developing grammars. In Peter Avery, Elan Dresher & Keren Rice (eds), *Contrast in Phonology: Theory, Perception, Acquisition*, 231–270. Berlin: Mouton.

Ingram, David. 1992. Early phonological acquisition: A cross-linguistic perspective. In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds), *Phonological Development: Models, Research, Implications,* 423–435. Timonium, MD: York Press.

Jakobson, Roman. 1941/68. *Child Language, Aphasia, and Phonological Universals*. The Hague: Mouton. Eng. tr. of *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala, 1941.

Jespersen, Otto. 1916. *Nutidssprog hos born og voxne.* Copenhagen: Gyldendal Nordisk forlag.

Jespersen, Otto. 1925. *Die Sprache, ihre Natur, Entwicklung, und Entstehung.* Heidelberg: C. Winter.

Keren-Portnoy, Tamar, Marinella Majorano & Marilyn M. Vihman. 2009. From phonetics to phonology: The emergence of first words in Italian. *Journal of Child Language* 36. 235–267.

Kiparsky, Paul & Lise Menn. 1977. On the acquisition of phonology. In John Macnamara (ed), *Language Learning and Thought,* 47–78. New York: Academic Press.

Kock, A. 1901. Om Barnspråk. *Nordisk Tidskrift*. (Cited in Jakobson, 1941/68; no publisher given.).

Kunnari, Sari. 2000. *Characteristics of Early Lexical and Phonological Development in Children Acquiring Finnish.* Acta Universitatis Ouluensis, Humaniora, B34. Oulu: Oulu University Press.

Labov, William & Theresa Labov. 1978. The phonetics of cat and mama. *Language* 54. 816–852.

Leopold, Werner F. 1939. *Speech Development of a Bilingual Child,* 1: *Vocabulary Growth in the First Two Years*. Evanston: Northwestern University Press.

Levelt, Clara C. 1995. The segmental structure of early words: Articulatory frames or phonological constraints. In Eve Clark (ed), *Proceedings of the* 27th *Child Language Research Forum,* 19–27. Stanford, CA: CSLI.

Li, Wu, Valentin Piëch & Charles D. Gilbert. 2004. Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience* 7. 651–657.

Lindner, Gustave. 1898. *Aus dem Naturgarten der Kindersprache*. Leipzig: Th. Grieben's Verlag.

Lleó, Conxita. 1990. Homonymy and reduplication: On the extended availability of two strategies in phonological acquisition. *Journal of Child Language* 17. 267–78.

Lohuis-Weber, Heleen & Wim Zonneveld. 1996. Phonological acquisition and Dutch word prosody. *Language Acquisition* 5. 245–283.

Macken, Marlys A. 1978. Permitted complexity in phonological development: One child's acquisition of Spanish consonants. *Lingua* 44. 219–253.

Macken, Marlys A. 1979. Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua* 49. 11–49.

Macken, Marlys A. 1980. The child's lexical representation: The 'puzzle-puddle-pickle' evidence. *Journal of Linguistics* 16. 1–17.

Macken, Marlys A. 1992. Where's phonology? In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds), *Phonological Development: Models, Research, Implications,* 249–269. Timonium, MD: York Press.

Macken, Marlys A. 1995. Phonological acquisition. In John A. Goldsmith (ed), *The Handbook of Phonological Theory*, 671–698. Oxford: Blackwell.

Macken, Marlys A. & David Barton. 1980. The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language* 7. 41–74.

Menn, Lise. 1971. Phonotactic rules in beginning speech: A study in the development of English discourse. *Lingua 26.* 225–251.

Menn, Lise. 1976. Pattern, control and contrast in beginning speech: A case study in the development of word form and word function. Unpublished Ph.D. thesis, University of Illinois at Urbana-Champaign. Reprinted by the Indiana University Linguistics Club (Bloomington, IN, 1978).

Menn, Lise. 1978. Phonological units in beginning speech. In A. Bell & J. B. Hooper (eds.), *Syllables and Segments,* 157–171. Amsterdam: North-Holland.

Menn, Lise. 1983. Development of articulatory, phonetic, and phonological capabilities. In Brian Butterworth (ed), *Language Production*, vol. 2, 3–50. London: Academic Press.

Menn, Lise & Edward Matthei. 1992. The 'two-lexicon' account of child phonology: Looking back, looking ahead. In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds), *Phonological Development: Models, Research, Implications,* 211–247. Timonium, MD: York Press.

Oliveira, Daniela. 2008. Percurso de construção da fonologia pela criança: uma abordagem dinâmica. (The course of phonological construction in children: A dynamic approach.) Unpublished Ph.D. thesis, UFMG, Belo Horizonte, Brazil.

Priestly, Thomas M. S. 1977. One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language* 4. 45–66.

Pye, Clifton, David Ingram & Helen List. 1987. A comparison of initial consonant acquisition in English and Quiché. In Keith Nelson & Anne van Kleeck (eds), *Children's Language* 6, 175–190. Hillsdale, NJ: Erlbaum.

Rasmussen, W. 1913. *Barnets sjælelige udvikling i de første fire aar.* Copenhagen. (Cited in Jakobson, 1941/68; no publisher given).

Rice, Keren & Peter Avery. 1995. Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In John Archibald (ed), *Phonological Acquisition and Phonological Theory*, 23–42. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rice, Keren. 2007. Markedness in phonology. In Paul de Lacy (ed), *The Cambridge Handbook of Phonology*, 79–98. Cambridge: CUP.

Salo, Ave. 1993. Muutelõppude ilmumine ühe eesti lapse keelde vanuses 1;5–2;5. [The emergence of inflectional endings in one Estonian child's speech, age 1;5–2;5.] Undergraduate thesis, Tartu University.

Savinainen-Makkonen, Tuula. 2001. *Suomalainen lapsi fonologiaa omaksumassa.* [The acquisition of phonology in Finnish children.] Publications of the Department of Phonetics, University of Helsinki, 42.

Savinainen-Makkonen, Tuula. 2007. Geminate template: A model for first Finnish words. *First Language* 27. 347–349.

Stager, Christine. L. & Janet F. Werker. 1997. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388. 381–382.

Stoel-Gammon, Carol & Judith Dunn. 1985. *Normal and Disordered Phonology in Children.* Baltimore: University Park Press.

Thelen, Esther & Linda B. Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action.* Cambridge, MA: MIT Press.

Velten, H.V. 1943. The growth of phonemic and lexical patterns in infant language. *Language* 19. 231–292.

Vihman, Marilyn M. 1976. From prespeech to speech: On early phonology. *Stanford Papers and Reports on Child Language Development* 12. 230–244.

Vihman, Marilyn M. 1992. Early syllables and the construction of phonology. In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds), *Phonological Development: Models, Research, Implications,* 393–422. Parkton, MD: York Press.

Vihman, Marilyn M. 1993. Variable paths to early word production. *Journal of Phonetics* 21. 61–82.

Vihman, Marilyn M. 1996. *Phonological Development: The Origins of Language in the Child.* Oxford: Basil Blackwell.

Vihman, Marilyn M. 1978. Consonant harmony: Its scope and function in child language. In Joseph H. Greenberg (ed), *Universals of Human Language,* 281–334. Stanford, CA: Stanford University Press.

Vihman, Marilyn M. 2009. Word learning and the origins of phonological system. In S. Foster-Cohen (ed), *Advances in Language Acquisition,* 15–39. Luton: Macmillan.

Vihman, Marilyn M. & Bénédicte de Boysson-Bardies. 1994. The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica* 51. 159–169.

Vihman, Marilyn M. & William Croft. 2007. Phonological development: Toward a 'radical' templatic phonology. *Linguistics* 45. 683–725.

Vihman, Marilyn M. & Virginia Mueller Gathercole. In prep. *Language Development: Issues, Models, Empirical Evidence.* Oxford: Basil Blackwell.

Vihman, Marilyn M. & Sari Kunnari. 2006. The sources of phonological knowledge: A cross-linguistic perspective. *Recherches Linguistiques de Vincennes* 35. 133–164.

Vihman, Marilyn M. & Ruth Miller. 1988. Words and babble at the threshold of language acquisition. In Michael D. Smith & John L. Locke (eds), *The Emergent Lexicon*, 151–183. N.Y.: Academic Press.

Vihman, Marilyn M. & Shelley L. Velleman. 1989. Phonological reorganization: a case study. *Language and Speech* 32. 149–170.

Vihman, Marilyn M. & Shelley L. Velleman. 2000a. Phonetics and the origins of phonology. In Noel Burton-Roberts, Philip Carr & Gerry Docherty (eds), *Phonological knowledge: Its nature and status*, 305–339. Oxford: Oxford University Press.

Vihman, Marilyn M. & Shelley L. Velleman. 2000b. The construction of a first phonology. *Phonetica* 57. 255–266.

Vihman, Marilyn M., Shelley L. Velleman & Lorraine McCune. 1994. How abstract is child phonology? Towards an integration of linguistic and psychological approaches. In Mehmet Yavas (ed), *First and Second Language Phonology*, 9–44. San Diego: Singular Publishing.

Vihman, Marilyn M. & Virve-Anneli Vihman. In press. From first words to segments: A case study in phonological development. In Eve V. Clark & Inbal Arnon (eds), *How children make linguistic generalizations: Experience and variation in learning a first language. Trends in Language Acquisition Research.* Amsterdam: John Benjamins.

Waterson, Natalie. 1971. Child phonology: A prosodic view. *Journal of Linguistics* 7. 179–211.

Waterson, Natalie. 1978. Growth of complexity in phonological development. In N. Waterson & C. E. Snow (eds), *The Development of Communication*, 415–442. New York: John Wiley.

Waugh, Linda & Monique Monville-Burston (eds). 1990. *On Language/Roman Jakobson.* Cambridge, MA: Harvard University Press.

Wellman, Beth L., Ida Mae Case, Ida G. Mengert & Dorothy E. Bradbury. 1931. *Speech Sounds of Young Children.* University of Iowa Studies in Child Welfare, 5.

Werker, Janet F. & Christine L. Stager. 2000. Developmental changes in infant speech perception and early word learning. In Michael B. Broe & Janet B. Pierrehumbert (eds), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*,181–193. Cambridge: Cambridge University Press.

Werker, Janet F. & Robert C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7. 49–63.

Whalen, Douglas H., Andrea G. Levitt & Qi Z. 1991. Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language* 18. 501–516.