# The emergence of phonological system

- Whole word phonology
  - Some history
  - Babbling practice and the first words: Role of the ‘articulatory filter’
  - Word templates
  - Some questions about word templates
  - Extent of evidence
    - Word template analysis: A diary study
    - Word template analysis: Research studies
      1. English (UK): Tomos
      2. English (US): Deborah
      3. Finnish: Eelis
      4. French: Noël
      5. Italian: Marco
      6. Welsh: Elen
  - Discussion of the evidence for whole word phonology
    - Extent of evidence
    - Typology
    - Sources of patterning
    - Prosodic vs. segmental structure
    - Ambient language effects
    - Representation for perception vs. production
  - The emergence of segments
- Learning mechanisms
  - Distributional or statistical learning
  - Lexical or symbolic learning

## References
In Chapter 3 we reviewed a range of different approaches to early phonological development, distinguishing between formal and functional models. In this chapter we will focus on one functional approach which has developed over a period of over thirty years, which we will term ‘whole word phonology’. Development of this model has naturally been influenced by others, particularly those which share the view that phonology is emergent, not innately known. It differs from other functional accounts primarily in its emphasis on individual differences and on nonlinear advances. In addition, the emergence of a link between perception and production within the first 12-18 months receives particular attention here, as does, although to a lesser extent, variability within child forms at a single point in time. Both observational and experimental data are available to support the ideas presented here. However, formalization within the terms of any of the theoretical frameworks used to describe adult phonology has been attempted only to a minimal extent (for an early attempt at a nonlinear model, see Menn, 1978; for a recent description drawing on metrical phonology, see Wauquier-Gravelines, 2006; for an analysis in terms of OT, using proportions of word forms reflecting different rankings, see Velieaman & Vihman, 2003). On the other hand, an attempt to quantify the emergent use of templates – a way to approach differences between typically developing and late talking toddlers – is currently under investigation.¹

**Whole word phonology**

In the field of child phonology Charles Ferguson is best known for a single paper, Ferguson and Farwell (1975), one of the early products of the Stanford Child Phonology Project which Ferguson established in 1968 and which continued, with repeated funding from the National Science Foundation, until 1988.² The method of analysis used in the paper was the ‘phone tree’, which demonstrated, for three children beginning to learn English, that sounds that were accurately and stably produced in some early words might well be substituted, either consistently or inconsistently, by other sounds in other words, and furthermore that the longitudinal profile itself was subject to change, with periods of greater or lesser variability. Ironically, the method itself has had no real uptake, while the findings it yielded have been cited repeatedly over the decades, with a notable increase in their acceptance in recent years.

The early word form variability identified in this paper was only one of the ‘surprises’ noted by Ferguson and Farwell. Another was the relative accuracy of early words, which tend to resemble their targets in overall segmental content and sequencing, although there may be omissions; this may be followed by a period of word production that is less faithful to the target – in other words, by a ‘regression’ (Ferguson, 1978). The third surprise, a complement to the second, was the patterning apparent in the target forms on which the early words are

¹ Thanks are due to the participants in an informal workshop on this topic at the 2007 Biennial Meeting of the Society for Research in Child Development, in Boston: Tamar Keren-Portnoy, Lorraine McCune, Lise Menn and Shelley Velieaman, together with Marilyn Vihman.

² A conference held at Stanford in 1989, published as Ferguson et al., eds. (1992), brought a formal close to the project – ending a productive period of studies inspired and organized by Ferguson that was also marked earlier by the NIH conference published as Yeni-komshian, Kavanagh and Ferguson (1980).
modelled (see also Ferguson, Peizer & Weeks, 1973). All of these discoveries – the reward, in a sense, for Ferguson’s insistence on letting the data themselves guide the interpretation – have proven highly fruitful in the decades since the publication of that paper.

What is most memorable about Ferguson and Farwell (1975) is undoubtedly the insight that the authors provided regarding ‘whole word phonology’, namely, the idea that, even in adults, the lexical item has primacy, with ‘a complex array of phonological elements and relations’ (p. 437) deriving from it in development and being represented alongside it in the adult system (see Beckman & Edwards, 2000). In a now classic statement, Ferguson and Farwell made the assumption ‘that a phonemic core of remembered lexical items and articulations which produce them is the foundation of an individual’s phonology...’ (Ferguson & Farwell, 1975, p. 437). In this chapter we will (1) briefly trace the history of the idea of ‘whole word phonology’, (2) raise several questions about just what the idea might mean in relation to available data from the children and languages that have been studied so far, (3) provide data and analyses drawn from a number of different languages, based on diary studies as well as longitudinal and cross-sectional research, to support discussion, and finally (4) attempt to answer at least some of the questions raised.

Some history
The view that phonological development begins with ‘whole word representations’ began to gain support only very slowly when it was first put forward in the 1970’s (e.g., Menn, 1971, 1983; Waterson, 1971; Ferguson & Farwell, 1975; Macken, 1979; some notable expressions of the general idea that appeared in print even earlier include Ferguson, 1963; Fry, 1966; Francescato, 1968; Tervoort, 1969). The work of Waterson alone among these early promoters of the ‘whole word’ idea was rooted in a full linguistic theory, namely, that of J. R. Firth.

Anticipating some aspects of the current nonlinear models by several decades, Firth's prosodic analysis (1948) went beyond the segment-oriented analysis of structural linguistics to allow for the phonetic consequences of the continuous and partially overlapping flow of articulatory gestures which characterizes speech. Phoneme-like or 'phonematic' (C and V) units are supplemented in this model by prosodies, phonological structures of any length. Like nonlinear analysis within the generative framework, prosodies have been effectively used to deal with relations of length, stress, and tone, as well as with harmonic constraints and with ‘spreading’ phonetic features such as palatalization, retroflexion, nasalization, and glottalization. 3

3 For illustration and references to prosodic analyses of a range of different languages and phenomena, see Palmer (1970) and Robins (1989), in addition to the first chapters of Waterson (1987), a partially revised collection of earlier papers. Goldsmith (1992) provides an account of the relationship of Firthian Prosodic Analysis to Autosegmental Phonology, as he sees it; Ogden & Local (1994) provide a response from the point of view of phonologists continuing to work within the Firthian tradition.
Waterson attributes to Firth many of the elements of her own work which are most original and which have proven most influential. Above all, her focus on whole words rather than segments as the relevant unit for early phonologies derives directly from the theory:

It is the whole-unit or holistic approach of prosodic phonology which brought to light the patterned relationships between a child's hardly recognizable words and the adult models and thus made it possible to explain the somewhat bizarre forms of early words, as well as why sounds that children are capable of producing are not used in all the contexts in which they occur in adult forms. (1987, p.2)

Waterson (1971), the most widely cited of Waterson's studies of her son's phonological development, has three explicit goals: (1) to illustrate and defend the value of discovering the child's system on its own terms; (2) to demonstrate that the patterning of the child's forms is best understood to be holistic rather than based on segment-by-segment substitutions for adult sounds; (3) to support the contention that the child's perception is different from the adult's, and is at first schematic and incomplete. Each of these issues remains of central concern today.

Waterson's detailed and discerning analyses of her son's early word patterns provides a rich illustration of the notion of 'the child's own system'; these patterns are taken to reflect the child's perceptual filtering of the input. For each pattern, according to Waterson, the child forms reproduce only the most salient acoustic features shared by a set of adult words. Waterson insists on the differential perceptual salience of adult words as the basis for the child's choice of a production pattern; what is salient for the child is related both to what is already in his repertoire and to the 'most strongly articulated features' (1971, p. 41). Unfortunately, no independent evidence is provided for the characterization of certain features as 'most strongly articulated', so that the explanation has an inescapably circular ring.

While acknowledging the child's active role in constructing and systematizing a lexicon, Waterson wants to explain phonological development in terms of a gradual loosening of constraints on the complexity of internal lexical representations. Permitted complexity constraints are in turn assumed to reflect limitations on what the child is capable of perceiving linguistically, at any given time. (Queller, 1988, p. 465).

In the United States, where Firth's ideas had less influence than in Britain, the idea of a holistic start on phonological representation in the child grew largely out of close attention to child data (Menn, 1971; Ferguson et al., 1973 as well as Ferguson & Farwell, 1975). It emerged well ahead of its time in terms of mainstream American phonological theory. Chomsky and Halle's classic distinctive-feature based account of 'generative phonology' was published in 1968, while non-linear accounts of adult phonology began to appear only some years later (e.g., Goldsmith, 1976). The idea that holistic patterning might play a role in adult linguistic systems began to be considered in mainstream American linguistics only some 10 to 20 years after that, with the publication of various cognitive grammar approaches to syntax (e.g., Croft, 1995, 2001; Fillmore, 1988;
Langacker, 1987; Goldberg, 1995) and several variants of non-linear phonology, notably gestural phonology (e.g., Browman & Goldstein, 1990; cf. Studdert-Kennedy & Goodall, 1995).

Waterson (1971) was the first to provide strong evidence that only a whole word approach could identify the systematicity in some child forms. She articulates the position clearly:

Some of [Waterson’s son] P’s early forms seemed so different from the corresponding adult forms as to appear to have no relationship to them at all, but they are known to be the same by their function in context...Examined segmentally such child’s forms show very little congruence with the adult forms...’ (Waterson, 1971, p. 179).

To illustrate, she provides a set of four child forms that begin with the palatal nasal (which does not occur in adult English, except as the cluster [ŋ], as in *canyon*), as shown below.

Table 1. The nasal structure (Waterson’s ‘P’: 1971)

<table>
<thead>
<tr>
<th>adult forms</th>
<th>child forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>another</td>
<td>/ənʌðə/</td>
</tr>
<tr>
<td>finger</td>
<td>/fɪŋɡə/</td>
</tr>
<tr>
<td>Randall</td>
<td>/rændl/</td>
</tr>
<tr>
<td>window</td>
<td>/wɪndəʊ/</td>
</tr>
</tbody>
</table>

When compared segment by segment with the adult forms, the correspondences defeat any plausible analysis. Considered as phonetic patterns on their own, the forms clearly constitute a simple reduplicated production routine, the use of palatal [ŋ] possibly having its source in an articulation the child developed in babbling.

Although there is no close linear correspondence with the apparent adult targets, the relationship is not random but merely holistic: The child deploys the disyllabic ‘nasal structure’ (as Waterson terms it) in response to multisyllabic adult words that feature a nasal in the stressed syllable. A comparable pattern is reported in Vihman (1981): Her son Raivo, acquiring Estonian in the home, settled on the monosyllabic form [nən] to express the (phonologically spelled) Estonian adult forms *kinni* ‘closed’, *king* ‘shoe’, *lind* ‘bird’ and *rind* ‘breast (in nursing)*. In this case it is easy to see what it is about the auditory image left by the adult forms that suggests the child’s form: A medial [m] sequence unites all four adult target words, essentially constituting the rime across all four words (rime of first syllable of *kinni*, rime excluding final consonant of cluster in the remaining words). (We return to Raivo’s early word phonology below.)

The three primary arguments used to support the claim that the earliest phonological structure is whole-word based are summarised in Vihman & Croft (2007, pp. 689f.) as follows:
1. **Variability of segment production**: A child may produce the same sounds differently in different words, and some words may be more variable than others. This suggests that the child has knowledge of particular words but has not yet developed abstract categories of sounds for production (Ferguson & Farwell 1975).

2. **Relationship of child word to adult target**: The relation of early child words to their adult models is often found to be difficult to account for on a segment-by-segment basis. Instead, the child seems to be targeting a whole gestalt (Waterson 1971). The resulting patterns have been described as ‘whole word processes’, sometimes characterized as either HARMONY (assimilation of non-contiguous vowels or consonants) or MELODY (patterning in the sequencing of non-contiguous vowels or consonants) (Grunwell 1982; Macken 1992, 1995; Vihman 1996).

3. **Relationship between child words**: The interrelation between the child’s own words may be more evident than the relation to the adult models (Macken, 1979). This is due to the child’s eventual reliance on one or more word templates, specific phonological patterns which fit many of the words that the child attempts (these words may be said to be SELECTED) but which are also extended to words that are less close to the template (these words are then ADAPTED to fit the template: Vihman and Velleman, 2000; see also Jaeger, 1997).

Vihman and Croft add an additional argument, based on the fact that child patterns often suggest effects rooted in the entire adult word, such as inter-consonantal or inter-vowel effects across syllables or non-contiguous syllable positions:

4. **Source of child patterns.** The dominant child patterns of the early word production period are responses to challenges posed by the adult target words as a whole pattern, primarily, the challenge of producing distinct consonants or distinct vowels, or both, in different word positions or different syllables (typically, initial and final consonants in a monosyllable or successive onset consonants in a disyllable).

**Babbling practice and the first words: Role of the ‘articulatory filter’**

Just what a ‘whole word’ start on phonology should be taken to mean has been disputed, even by supporters of the idea (see the discussion in Studdert-Kennedy & Goodell, 1995). If it is meant to refer to a holistic representation for production, then how do we explain the child’s achievement of a word form recognizable to others, with all the detailed motoric instructions that must involve, given the ‘vagueness’ and/or underspecification of at least some segments or features that a holistic representation seems to imply (cf. Ingram, 1974)? This question, which goes to the heart of the issue, can briefly be addressed here (see also Vihman & Kunnari, 2006).

The developmental underpinnings of ‘whole word patterning’ in production must be traced back to babbling. Vihman (1993) argued that an ‘articulatory filter’ served to highlight word forms in the input speech stream that provided a relatively close match to the child’s own production patterns, as established
through babbling or vocalisation ‘practice’ (see also Locke, 1986). Given that speech to the child is generally highly repetitive or routinised, accompanying the caregiving situations that arise on a daily basis or even several times a day (feeding, changing diapers, putting to sleep, etc.), many words are heard and heard again with exceptionally high frequency over the early months of life. It is from among those words that the first to be produced will be ‘selected’ (or highlighted perceptually) by the ‘articulatory filter’, based on the child’s implicit experience of a rough match of the adult word form to an existing vocalisation pattern. By ‘implicit’ here we mean that no attention need be directed to the adult word as it is spoken for the child to experience the match. Production of ‘context-limited’ words may then follow in the appropriate situational context, arguably due to implicit priming by the situation, at least in the case of the very first word production. The ‘filter’ would serve only to boost the salience or memorability of certain word forms, based on their emergent familiarity to the child through experience with his or her own vocalisations (‘output as input’: Elbers, 1997, 2000).

The idea that such a filter, developed through production, could affect the way the child experiences speech has been confirmed experimentally, although only individual consonants have been tested, not whole one- or two-syllable patterns (Vihman & Nakai, 2003; DePaolis, 2006). The principle remains the same: The child is able to produce words without having a detailed segmental representation of those word forms because they are, initially, supported by an existing production routine. We will return to these issues in the final part of this chapter, LEARNING MECHANISMS, where we lay out our understanding of the origins of child word form representations in the implicit and explicit processing of speech.

Word templates
The term ‘word template’ has been used by a number of child phonologists (Menn, 1978? 1983?; Macken, 1995, 1996; Vihman & Velleman, 2000) to express the idea of a word production pattern or routine specific to a particular child. The production pattern, originating in babbling practice, would be further shaped by its use in matching adult word forms. It can be termed a ‘template’ only when it has begun to be extended to produce adult word forms that do not match the child’s existing output patterns. Here the child no longer draws on experience of a match but, rather, projects his own well-practiced output routine onto adult words that require a more or less radical ‘adaptation’, such as metathesis, if they are to be accommodated within the child’s system. We see the child drawing on an internal schema here, abstracted away from his experience in producing some or all of his early words. The template permits more word learning and use without going beyond the child’s existing phonetic or motoric resources. Furthermore, the experience of attempting a wider range of adult word forms can be expected to stimulate the development of new, more complex phonological patterns, even while the child’s existing patterns continue to constrain the output lexical forms. Finally, the resulting pattern can be considered phonological, given its abstract source, and the term ‘template’ accordingly marks the origins of phonological systematicity. The child’s establishment or ‘discovery’ of the power of the template – or the increased flexibility in word production afforded by the freedom to adapt adult word forms to existing production resources – typically results in a more or less rapid increase in lexical learning, suggesting that
phonological advance is indeed a critical ‘control parameter’ for many children as they enter into language (McCune, 1992, in press).

**Some questions about word templates**
1. The *extent of evidence*, or the generality question: To what extent can the notion of word template be considered to apply to all children? We will begin here by presenting longitudinal data, both diary- and observational research-study based, for six different languages, illustrated with one child per language. Those data can then support discussion of the remaining issues.
2. The *typology* question: Are the template patterns themselves wholly idiosyncratic, or do they recur in different children?
3. The *sources of patterning*: If there are recurrent patterns, how can we understand their phonetic sources?
   a. *Challenges and opportunities*: Do the patterns give us an idea of what is difficult for children to produce (or plan, access, or represent) and what is not?
   b. *Prosodic vs. segmental structure*: Do the patterns involve primarily only prosodic or only segmental structure? Or is there an interaction between the two?
   c. *Ambient language effects*: To what extent do we find effects of the ambient language on the particular patterns produced?
4. The *representation* issue: If representation for production is ‘holistic’, how does this relate to the findings of experimental studies that report ‘fine phonetic detail’ in early perceptual representations, at least for familiar words?
5. The *emergence of segments*: If word template use is a developmental phenomenon, it can be expected to fade as the child’s control over the segments of the ambient language increases. Can we identify a shift to segmental representation in the developmental data?

**Extent of evidence**

Data analyses available in print to illustrate the idea of an emergent word template draw on both diary data (Vihman & Croft, 2007: children acquiring English, Estonian, German and Hindi are represented, for a total of eight child illustrations) and research studies, some of them longitudinal (Vihman & Kunnari, 2006: English, Finnish, French and Welsh, for a total of 11 children), others cross-sectional, based on data from the end of the single word period (Vihman, 2007: American English, British English, Finnish, French, Italian and Welsh, for a total of 12 child illustrations).\(^4\) In order to address the question as to how general the phenomenon of word templates may be we will add to the existing data base by reporting new analyses only, from the same two types of data.

1. To illustrate diary study data we present a more complete analysis of Vihman’s son Raivo than has appeared elsewhere, tracing the phonological patterns evident in his first 50 words.

---
\(^4\) This last study reported the findings of an analysis of 33 children each acquiring one of the five languages, with two dialects of English represented; data from only a few of the children could be included in the paper.
To illustrate research study data we will draw on the same children sampled in Vihman (in press), taking our illustrations from children not previously described.

Word template analysis: A diary study

The context of language development for Raivo (R), Vihman’s second child, has been described in several papers, beginning with Vihman (1981), which is in large part a phonological study. Briefly, R was exposed primarily to Estonian in the home but to English in the community. For a half-day only from 6-14 months he attended a daycare home where only English was spoken and until age 4 years a Day Care Center; until about age 3 years Estonian was the stronger of his two languages. R’s vocabulary in the first year or two of language use was approximately 75% Estonian (see Vihman, 1985).

Within the period covered by R’s first 50 words (age 13-16 months), including all recorded word forms in either language, imitated as well as spontaneous, two basic word-shape types may be distinguished, closed monosyllables and open disyllables (see Appendix A), although a very few exceptional words were produced as closed disyllables or open monosyllables.

I. closed monosyllables
   • with fricatives (+ one ts word). At first R produces the fricative alone, as a syllable nucleus (shoe [ʃ], vesi [s] and klotsid [ts | ts]). Later, he begins to produce monosyllables with a high vowel or schwa and a final fricative. The word up (1;4) is his first and only fricative word produced with [a] throughout this period. Two of his first five words (shoe, viska) fit the fricative only pattern; they are seemingly ‘selected’ for the fricative. In this category lack of C1-C2 harmony is tolerated – even when place differs (musi, müts, both [mɔs]).
   • with nasals. Again, the first form has a syllabic consonant, [m], as nucleus. But R imitates kell as [dɛn] early on (1;2.15); later he omits either the initial or the final consonant, but assimilates other such words to nasal only.
   • with obstruents. There are two imitations here in Estonian (kott, kiik), both produced with harmonizing consonants, and English that. In addition, peek-a-boo is produced as CVC+CV, with labial harmony: [bəp+bo:].

II. open disyllables
   • with glottals or glides. The first instances are monosyllabic diphthong-only words (ai, ei). Then R adds disyllabic forms with glottals or glides before producing his first disyllables with a true (supraglottal) consonant.

---

5 R’s acquisition of noun morphology is reported in Vihman (1982), of a bilingual lexicon in Vihman (1985), of verb morphology in Vihman & Vija (2006), and of early word combinations in Deuchar & Vihman (2001, 2006); his codeswitching in conversation with his sister is reported in Vihman (1998); Vihman (2002) addresses the question of bilingual early word phonology, with Raivo providing some of the examples.
• with obstruents. At first R uses primarily [a] in these forms (except for imit. tere); he then adds see [te(:), de, se], rinda [i ñæ] (deriving the CV syllable here from segments occurring in two different syllables in the target word), and only later more [mɔ, mɔʔ], muna [muma, mumu], auto [teðw, toto], ball [bɔ] (with ɔ, u, o). The child incorporates a high vowel in only three disyllabic word forms (rinda, muna; imit. of poll as [pi:]}, in contrast with the fricative words.

R’s phonology resembles that described by Waterson (1971) for her son P in that he has ‘small groups’ of words (cf. R’s ‘nasal structure’, mentioned earlier) rather than a single over-arching pattern. His phonology also resembles that of Daniel Menn (Menn, 1971) in that systematicity emerges only after a period of practice (with omissions constituting the only changes to target in the earliest words, which thus qualify as relatively ‘accurate’); the later words, which fall into two distinct word templates (CVC with a narrow vowel opening and CVCC with a wide vowel opening), reflect something of the patterning of the first words produced (‘while phonotactic rules have not yet crystallized in stage 1, something vaguely systematic, from which the rules will develop, is at work’: Menn, 1971, pp. 231f.). These data thus illustrate the progression predicted by the developmental hypotheses outlined above:

1. Relative accuracy in the early words, based on production practice through babbling (though no records were kept of Raivo’s babbling) and ‘selection’ of similar adult words to attempt (the function of the articulatory filter);
2. establishment of one or more templates, which are then projected onto non-matching as well as matching adult word forms, with resultant ‘regression’ in overall accuracy.6

In R’s case the first ‘adapted’ words – words which reflect the influence of the template as well as the adult word model itself – are seen in several of his template subgroup types from 1;2.7 onwards, or about three weeks after first identifiable word production. His earliest examples include monosyllabic [küs] for kiipsis ‘cookie’, [ba] for banaan ‘banana’ and harmonised [tut] for kott ‘bag’. As is already evident from these first examples, the main sources of difficulty for R are words of more than one syllable and changes in place of articulation across consonants within a word. The former challenge leads to syllable omission while the latter is generally met by consonant harmony (muna /muna/ ‘egg’ [mum]). Less than half of the forms he produces are disyllables, though he targets roughly even numbers of one- and two-syllable adult word forms (see Table 2).

6 Since we find regression, or a U-shape developmental curve, to be the rule rather than the exception in phonological development, Ingram’s (2002) ‘phonological mean length of utterance’ (pmlu) cannot serve to index advance, like MLU in syntax, at least in the earliest stages: Advance in syntax (although not morphology) is typically linear, with a gradual increase in numbers of words combined, at least in the early stages. This is simply not the case in the early stages of phonological learning (Savinainen-Makkonen & X, 2005; Majorano, 2005).
Table 2. Raivo’s word shapes, first-50-word period

<table>
<thead>
<tr>
<th>Word shape types</th>
<th>Monosyllabic targets</th>
<th>Disyllabic targets</th>
<th>Disyllables retained as such</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC/fricative</td>
<td>8 (4 English)</td>
<td>6 (1 English)</td>
<td>3 (1 English)</td>
</tr>
<tr>
<td>CVC/nasal</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>CVC/stop</td>
<td>7 (1 English)</td>
<td>1 [3-syll] (English)</td>
<td>1 (English)</td>
</tr>
<tr>
<td>(C)VCV/glide or glottal</td>
<td>2</td>
<td>5 (4 English)</td>
<td>5 (4 English)</td>
</tr>
<tr>
<td>(C)V(CV)</td>
<td>9 (4 English)</td>
<td>19 (2 English)</td>
<td>14 (2 English)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>32 (9 English)</strong></td>
<td><strong>35 (8 English)</strong></td>
<td><strong>25 (7 English)</strong></td>
</tr>
</tbody>
</table>

Most of the disyllables Raivo produces in this period fall into Glottal/glide or ‘hiatus’ categories (see Table 3), where ‘hiatus’ refers to a brief break between the syllables. This suggests that production was planned as two separate syllable-units. Once R begins producing disyllables in the shape CVCV with a ‘true’ (supraglottal) consonant, he maintains the target word length except for words with /l/ (põlle, tuli, prillid), where monosyllabic production means avoidance of the liquid (see Appendix A, part V).

Table 3. Developmental sequence for disyllable production

<table>
<thead>
<tr>
<th>Target word</th>
<th>Gloss</th>
<th>Child age</th>
<th>Child form</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>tere (I)</td>
<td>hi/hello</td>
<td>1;1.17</td>
<td>[teʔe]</td>
<td>glottal-filled hiatus</td>
</tr>
<tr>
<td>klotsid (I)</td>
<td>blocks</td>
<td>1;2.2</td>
<td>[ts</td>
<td>ts]</td>
</tr>
<tr>
<td>hiya</td>
<td>1;2.7</td>
<td>[aja ]</td>
<td></td>
<td>glide</td>
</tr>
<tr>
<td>allo, hello (I)</td>
<td>hello</td>
<td>1;2.7</td>
<td>[a</td>
<td>a], [awa]</td>
</tr>
<tr>
<td>rinda</td>
<td>breast (obj.)</td>
<td>1;2.15</td>
<td>[i</td>
<td>næ]</td>
</tr>
<tr>
<td>uh-oh (I)</td>
<td>1;2.15</td>
<td>[aʔa]</td>
<td></td>
<td>glottal</td>
</tr>
<tr>
<td>peek-a-boo</td>
<td>1;3.0</td>
<td>[baʔ</td>
<td>boː̂]</td>
<td></td>
</tr>
<tr>
<td>anna</td>
<td>give</td>
<td>1;3.2</td>
<td>[ana ]</td>
<td>flap n [cf. Daniel Menn: ono]</td>
</tr>
<tr>
<td><strong>hello</strong></td>
<td>1;3.10</td>
<td>[ajo]</td>
<td></td>
<td>glide</td>
</tr>
<tr>
<td>head aega</td>
<td>goodbye</td>
<td>1;3.15</td>
<td>[dada, daga]</td>
<td>First true CVCV</td>
</tr>
<tr>
<td>tere</td>
<td>hello</td>
<td>1;3.23</td>
<td>[tede, teda ]</td>
<td></td>
</tr>
</tbody>
</table>

7 Note that he does not attempt /r/ words at all, with the exception of tere ‘hello’ and rindi(a) ‘breast’). The Estonian short /r/ is a tap, the long /r/ a trill which can be further lengthened, like virtually all Estonian segments, in certain morphological contexts. Raivo was almost 5 years old before he began producing a recognizable Estonian /r/ of any kind.
Referring back now to the questions raised regarding sources of patterning, we have mentioned the ‘challenges’ R encountered and his way of meeting them. In his case, ‘opportunities’ could be said to include monosyllabic targets (readily available in both the languages to which he was exposed: see Table 2) and final fricatives. Beyond that, initial obstruents and nasals dominate his production, as they do, cross-linguistically, in virtually all cases analysed (Vihman, in press).

Turning to the next question, regarding the specific sources of patterning, both prosodic and segmental structure play a role in this case. The primary division is into monosyllables vs. disyllables. Within that division, there is a prosodic constraint (largely though not always observed), such that monosyllables are closed while both syllables of longer structures are open. However, there is also a segmental bias, affecting primarily the vowels (closed or high vowels in the CVC structures, open or low vowels in CVCV), but also to some extent the consonants (final nasals and fricatives are favoured in CVC, only stops and nasals occur in CVCV). Furthermore, consonant harmony is particularly characteristic of structures that include stops. Finally, notice that within the ‘nasal structure’ the closed-monosyllable template and the open harmonised disyllabic template are in competition, resulting in a particularly high level of variability (see Menn, 1971; Menn & Matthei, 1992). This set of often used words soon settled into a single monosyllabic pattern, [nˈn], mentioned earlier; this was the first sign of what Vihman (1981) identified as a ‘homonym strategy’, in which one form was used for a number of different words that lent themselves to incorporation or adaptation into that ‘schema’ (in Waterson’s terms) or template. Another example from within the 50-word period shown in Appendix A is the set juice, juustu ‘cheese’, vorst ‘sausage’ and, later, horse (see also Vihman, 2002, where the point is made that, for bilingual children, such templatic patterns do not always respect language barriers, based on the data illustrated there, from Raivo, exposed to Estonian and English, Tom, French and English, and Shelli, Hebrew and English).

Word template analysis: Research studies
While a diary study can provide a good deal of information regarding a child’s early words, including variant shapes for the same word as it evolves over time, data collected in a uniform way from a number of children exposed to the same

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Prosodic Constraint</th>
<th>Vowel Pattern</th>
<th>Consonant Pattern</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>kinni</td>
<td>closed</td>
<td>[ninîn], [ninin], [nîn]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>muna</td>
<td>egg</td>
<td>[muma], [mumu]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>kana (I)</td>
<td>chicken</td>
<td>[kaka]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>palun (I)</td>
<td>please</td>
<td>[mamum]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>auto (I)</td>
<td>car</td>
<td>[tōtō]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>auto</td>
<td>car</td>
<td>[te⁵w], [toto]</td>
<td>any vehicle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>byebye</td>
<td></td>
<td>[baba ]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Prosodic Constraint</th>
<th>Vowel Pattern</th>
<th>Consonant Pattern</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>kinni</td>
<td>closed</td>
<td>[ninîn], [ninin], [nîn]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>muna</td>
<td>egg</td>
<td>[muma], [mumu]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>kana (I)</td>
<td>chicken</td>
<td>[kaka]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>palun (I)</td>
<td>please</td>
<td>[mamum]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>auto (I)</td>
<td>car</td>
<td>[tōtō]</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>auto</td>
<td>car</td>
<td>[te⁵w], [toto]</td>
<td>any vehicle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>byebye</td>
<td></td>
<td>[baba ]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
language has complementary advantages, providing a more readily generalisable idea of the ‘challenges and opportunities’ as they are experienced by a group of children rather than a single child. Furthermore, it is only by studying several children learning each of several languages that we can begin to establish a typology of early word patterns or word templates.

For children learning the same language and especially for children learning different languages we have found it important to use developmental level rather than age as the basis for comparison, since children differ so dramatically in this period in their rate of lexical learning. Comparing two 14-month-olds may mean comparing a child who has just produced his first word with another child whose recorded lexicon has reached well over 50 words. While this may be appropriate for some ends (e.g., in order to relate specifically linguistic advances to those occurring in other domains, such as broader cognitive or communicative growth), it defeats the purpose of identifying similarities and differences in phonological development itself, since phonological and lexical development are closely intertwined.

To demonstrate that the earliest period of word use differs in important ways from later periods, with relatively accurate word form production in the first words but less accurate word forms (alongside more flexible word use: See Ch. 2) occurring alongside relatively accurate ones some weeks or months later, we present here analyses of longitudinal data from children recorded on a regular basis – weekly, biweekly, or monthly, in most cases – and sampled at two data points: The ‘4-word point’ (‘early words’), when the child has produced at least four different identifiable adult-based words spontaneously in the course of a 30-minute free-play session with her mother, and the ‘25-word point’ (‘later words’), when 25 such words are produced. Some of the data, which are drawn from a larger data base derived from several sources, have been reported elsewhere (D’Odorico et al., 2001; Kunnari, 2000; Veneziano & Sinclair, 2000; Vihman & Kunnari, 2006; Vihman, in press).

Table 4 provides an overview of the data included in Vihman (in press). Language groups were included in that study only when data were available for at least four different children, with a minimum of 25 different word types per child, whether produced spontaneously or imitated but excluding onomatopoeia which lack a stable target adult form. Diary studies were not included, in order to maintain methodological consistency. All of the data included in that study were collected longitudinally, but access to only one word point was available for the Italian children (D’Odorico et al., 2001). For the purposes of discussion, data from one child per language will be considered here, at each of the two word points, except for Italian, for which only a later data point is available.

Table 4. Languages, children and sample sizes
The children are ordered within language groups by number of identifiable word shapes (or ‘phonetic forms’) at the later word point. Names in **bold face** identify the data samples drawn on for illustration in previous studies (see text); names in **bold face italics** identify data samples used for illustration in this chapter. The count of ‘word shapes’ is based on different word types but additionally includes variants of the same word with differing phonological shapes.
<table>
<thead>
<tr>
<th>language</th>
<th>child</th>
<th>child age (months)</th>
<th>total word shapes</th>
<th>source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>early word session</td>
<td>later word session</td>
<td></td>
</tr>
<tr>
<td>English: UK</td>
<td>Jennifer</td>
<td>15</td>
<td>21</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Rebecca</td>
<td>12</td>
<td>19</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Jude</td>
<td>13</td>
<td>15</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Tomos</td>
<td>17</td>
<td>24</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Ali</td>
<td>15</td>
<td>25</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Sylvia</td>
<td>15</td>
<td>22</td>
<td>29</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>15</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Timmy</td>
<td>12-13</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Deborah</td>
<td>11</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Emily</td>
<td>13</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Alice</td>
<td>9-10</td>
<td>16</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Molly</td>
<td>11</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>12</td>
<td>16(^a)</td>
<td>32</td>
</tr>
<tr>
<td>Finnish</td>
<td>Matti</td>
<td>15</td>
<td>18</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Atte</td>
<td>17</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Eelis</td>
<td>18-19</td>
<td>22</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Mira</td>
<td>14</td>
<td>15</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Eliisa</td>
<td>13</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Venla</td>
<td>11</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>15</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>French</td>
<td>Camille</td>
<td>18</td>
<td>45</td>
<td>Veneziano &amp; Sinclair, 2000</td>
</tr>
<tr>
<td></td>
<td>Gaël</td>
<td>22</td>
<td>45</td>
<td>Veneziano, unpub.</td>
</tr>
<tr>
<td></td>
<td>Laurent</td>
<td>10-11</td>
<td>17.5</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Charles</td>
<td>11-12</td>
<td>15.5</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Carole</td>
<td>11</td>
<td>14</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Noël</td>
<td>13</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>12</td>
<td>17.5</td>
<td>36</td>
</tr>
</tbody>
</table>

\(^a\) The US English children (with the exception of Alice, one of 10 children followed by Lorraine McCune at Rutgers University) were followed as part of a Stanford study designed to look at ‘Phonological precursors to early speech’, from 9 to 16 or 17 months. The Rutgers children’s data were transcribed phonetically at Stanford through their 16-month sessions for comparative purposes, in collaborative McCune and Vihman research. The six children listed here were the only ones to produce more than 25 words in a session within the timeframe of the collaborative study.
Italian

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Gender</th>
<th>Year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>18</td>
<td>F</td>
<td>53</td>
<td>D’Odorico et al., 2001 [all Italian data]</td>
</tr>
<tr>
<td>Francesca</td>
<td>18</td>
<td>F</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Marco</td>
<td>23</td>
<td>M</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Andrea</td>
<td>24</td>
<td>F</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Alessandra</td>
<td>19</td>
<td>F</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Luca</td>
<td>24</td>
<td>M</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Linda</td>
<td>15</td>
<td>F</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Federico</td>
<td>18</td>
<td>M</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

mean | 22  | 34 | Total N = 26

Welsh

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Gender</th>
<th>Year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gwyn</td>
<td>12</td>
<td>F</td>
<td>46</td>
<td>Vihman, 2000</td>
</tr>
<tr>
<td>Elen</td>
<td>13-14</td>
<td>F</td>
<td>35</td>
<td>[all Welsh data]</td>
</tr>
<tr>
<td>Catrin</td>
<td>12</td>
<td>F</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Fflur</td>
<td>13</td>
<td>F</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

mean | 13  | 17 | 36 | Total N = 5

overall mean | 13  | 18 | 36 | Total N = 80

1. English (UK): Tomos

a. Early words (17 months)

Two supraglottal consonants, [b] and [n], occur as match to target in Tomos’ early words (Table 5). In addition, [d] occurs as an unsystematic variant of different onsets while [p/b] occurs, surprisingly, as onset in the word ta. The only vowels are low and/or central. One diphthong and syllabic [m] also occur following the labial stop. Both monosyllables and disyllables occur, usually in accurate relation to the target syllable count; three of the four disyllabic word tokens are reduplicated. All of the forms are relatively close to the target forms with the exception of Badger, a catch-phrase the child’s grandfather has taught him.

Table 5. Tomos (English/UK child). Early words, 17 months.

<table>
<thead>
<tr>
<th>Target word</th>
<th>Target form</th>
<th>Child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badger</td>
<td>[bædʒə]</td>
<td>[bæmː], [bæm]</td>
</tr>
<tr>
<td>bang</td>
<td>[bæŋ]</td>
<td>[ba], [bæ], [baʊ], [da]</td>
</tr>
<tr>
<td>hiya</td>
<td>[haija]</td>
<td>[ja], [dæː]</td>
</tr>
<tr>
<td>no</td>
<td>[naʊ]</td>
<td>[na], [næ], [na]</td>
</tr>
<tr>
<td>ta ‘thank you’</td>
<td>[ta]</td>
<td>[ba], [pa], [baː], [tæː]</td>
</tr>
</tbody>
</table>

b. Later words (24 months; total words in session: 40)

i. Prosodic shape. Tomos’ word production is dominated by monosyllables (60%). Most of these (83%) are open, CV(V). Some are ‘selected’, with long vowels (car, me, more, star [taː], yeah)⁹ or diphthongs (cow [daʊ], no, as well as the [Vɪ] words included in Table 6); two – oh, shoe – are produced with a short vowel. Other monosyllables show omission of the coda: ball, book, bowl, juice,

---

⁹ Recall that this child is learning British English, with long vowels for intrasyllabic /Vr/ sequences.
Mott, peas. However, whereas only one coda had occurred in the previous recording session (Tomos’ first with over 25 words used spontaneously, two weeks earlier: \[ p^\text{b}uk^\text{h} \] for box), now four CVC monosyllables occur, all ‘selected’ (dog, sit, (tick)tock, yuck). As in the previous session a small number of words are ‘adapted’ to include the diphthong \(<V_1>\): bang, water (see Table 6), where in the previous session milk was produced as \[ h\text{m˘u}i \] and Postman Pat as \[ pu\text{t˘b}ai \].

Table 6. Tomos (English/UK child). Later words, age 24 months. N words in session = 40.

<table>
<thead>
<tr>
<th>(&lt;…)CV_1&gt;)</th>
<th>adapted</th>
<th>adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>selected</strong></td>
<td><strong>adapted</strong></td>
<td><strong>adaptation</strong></td>
</tr>
<tr>
<td>target word</td>
<td>child form</td>
<td>target word</td>
</tr>
<tr>
<td>bye</td>
<td>[bəi]</td>
<td>bang</td>
</tr>
<tr>
<td>nain (im.) Welsh ‘grandmother’ [nəi:n]</td>
<td>[næʔi]</td>
<td>water</td>
</tr>
<tr>
<td>train (im.)</td>
<td>[tɛt]</td>
<td></td>
</tr>
</tbody>
</table>

\(<C_1,VC,V(C)>\)

| beep beep | [bibiː:] | bucket | \[pʰdplʰeɪt\] | Harmonise \(C_2\) |
| bumble bee | [bububiː] | frog: boing-boing | [bəbəp] | Harmonise \(C_3\) |
| choo choo | [tutuː] | jigsaw (im.) | [tʃɪtʃɔː] | Harmonise \(C_2\) |
| daddy | [dadnː] | playdough | [dæɪdəʊʔ] | Harmonise \(C_1\) |
| mummy | [mʌmɪː] |

In disyllables (but not in closed monosyllables) Tomos’ dominant pattern is CH, as shown in Table 6. Just three disyllables occur with differing \(C_1\) and \(C_2\): butterfly \([palaːi]\) (with likely interaction with the Welsh word for ‘butterfly’, pilipala, produced in the previous session as \[p^hlibæ]\), puzzle \([pæʒu]\), and digger \([dɪɡɪja]\). Interestingly, codas now occur on disyllables as well (bucket, boing-boing, and one anomalous form, quack-quack \([baʔbək^h]\)).

**ii. Segmental inventory.** Tomos’ has a large number of match-to-target segments, especially vowels in stressed syllable (Table 7). Fricatives or affricates occur in all word positions, although infrequently and not always accurately; coda consonants are emergent, with the velar stop occurring as a match in more than one word only in this word position.

Table 7. Tomos: Segmental inventory in later words

<table>
<thead>
<tr>
<th>Consonant inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Word onset</td>
</tr>
</tbody>
</table>
2. Word medial

\[(p) \quad (t) \quad ([tS])\]

3. Word final

\[(p) \quad t \quad k \quad ([?])\]

Vowel inventory

1. Stressed syllable

\[i \quad u\]

\[i \quad u\]

\[(e) \quad ou, o\]

\[[e]\]

\[[?e]\]

\[a\]

2. Unstressed syllable

\[i \quad u\]

\[(e) \quad (o)\]

\[a\]

2. English (US): Deborah

a. Early words (10-11 months)

Deborah’s early words include only one supraglottal consonant, [b] (Table 8). A range of different vowels are represented and both monosyllables and disyllables. There is no one identifiable pattern, but the phonetic inventory is restricted. In addition, the child forms are noticeably similar to the target forms.

Table 8. Deborah (English/US child) Early words, 10-11 months.

<table>
<thead>
<tr>
<th>target word</th>
<th>target form</th>
<th>child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>baa</td>
<td>[ba:], [?ae:]</td>
<td>[?ae:]</td>
</tr>
<tr>
<td>baby</td>
<td>[beibi]</td>
<td>[bebi]</td>
</tr>
<tr>
<td>hi</td>
<td>[hai]</td>
<td>[?ai]</td>
</tr>
<tr>
<td>uh-oh</td>
<td>[?a?:ou]</td>
<td>[?a?:?e]</td>
</tr>
</tbody>
</table>

b. Later words (17 months; total words in session: 37)
i. Prosodic shape. Deborah’s word forms are, on the whole, quite ‘accurate’ – if we make allowance for the lack of codas, which are consistently omitted from the nine ‘selected’ words which have them – and for the sporadic omission of glottals and glides at onset (initial /h/ in hello [ΛwO] and /w/ in what’s that? [Λsæ]). Altogether, 27 (73%) of the different word types produced in the session can be considered to be ‘selected’. Eighteen of these are open monosyllables, eight of them of the shape (C)VV, the rest CV. The remaining ‘selected’ words are either fully or partially reduplicated (harmony forms) or VCV, namely, hello, what’s that, and uh-oh.

As was the case with Tomos, the front rising diphthong is an ‘attractor’, or a shape that attracts child production even for non-matching target words. Whereas in Tomos’ case it led sporadically to adaptation even in the case of longer words, Deborah adapts only monosyllables (and the imitated form, bunny) to the <CV₁> template, while longer targets (plus bus) are fitted into the harmony template. The vowels in these words are largely taken from the target. In addition, [w] serves as the basis for CH in cracker, presumably inspired by the /t/ of the onset cluster of the target form, while giraffe is the only form besides (cock-a-)doodle-do to be produced with /d/ - with harmony again spreading from the onset (where [d] for /dʒ/ is a common substitution) to the medial position. Finally, spaghetti, one of only three long forms attempted (with kitty-cat and around-and-around), shows what looks like metathesis, with the onset harmonizing to the medial velar stop but the stressed syllable taking its vowel from the omitted final syllable.

Table 9. Deborah (English/US child). Later words, age 17 months. N word shapes in session = 37.

<table>
<thead>
<tr>
<th>selected</th>
<th>adapted</th>
<th>adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>target word</td>
<td>child form</td>
<td>target word</td>
</tr>
<tr>
<td>A</td>
<td>[eI]</td>
<td>bunny (im.)</td>
</tr>
<tr>
<td>bye</td>
<td>[baI]</td>
<td>car</td>
</tr>
<tr>
<td>eye</td>
<td>[aI]</td>
<td>one (im.)</td>
</tr>
<tr>
<td>hi</td>
<td>[haI]</td>
<td></td>
</tr>
<tr>
<td>my, mine</td>
<td>[maI]</td>
<td></td>
</tr>
<tr>
<td>&lt;CV₁&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>around-&amp;-around</td>
<td>[waIwau]</td>
<td>bagel (im.)</td>
</tr>
<tr>
<td>baby</td>
<td>[bebi]</td>
<td>bus</td>
</tr>
<tr>
<td>(cock-a-)doodle-do</td>
<td>[dΛʔdu]</td>
<td>cracker (im.)</td>
</tr>
<tr>
<td>kittycat (im.)</td>
<td>[kikæ]</td>
<td>giraffe (im.)</td>
</tr>
<tr>
<td>mama</td>
<td>[mama]</td>
<td>pickle (im.)</td>
</tr>
</tbody>
</table>
Here, as in other cases, we disregard systematic omission of codas and of unstressed syllables (around-and-around, kittycat) in determining whether a word is ‘selected’ (‘roughly accurate’) or adapted. Thus, for Deborah, cheese is quite accurately reproduced as [tʃi] and we consider that word to be ‘selected’. Following the same logic, we disregard the substitution of [d] for /dʒ/, as in [dɪdæ] for giraffe, and of [w] for /kr/, as in [wæwɔ] for cracker (though for other reasons both of these happen to count as ‘adapted’ forms in Deborah’s case).

When, as in the case of both R and Tomos, detailed above, codas are regularly produced in some words but are omitted in others, we treat omission of codas as an ‘adaptation’. Where a child fails to produce any codas or words of more than two syllables, as here, then coda or unstressed syllable omission can be disregarded: Possibly, with no codas or three-syllable forms available in the child’s own production repertoire for ‘matching’, these may not at first be fully ‘registered’. That is, although under ideal experimental conditions the child would surely prove able to ‘perceive’ them, she would be less likely to accurately retain the full pattern in an everyday word learning situation.\(^\text{10}\)

On the other hand, the child’s production repertoire does not normally remain ‘stuck’ at a particular developmental point (although this may be one source of phonological disorder: Velleman & Vihman, 2002). Instead, we can hypothesize that it is the ‘adapted’ forms, with their tension between adult target – which may or may not be fully represented early on, at least as regards unprimed access for production – and the child’s own form, that lead the child on to more advanced phonetic attempts, as the discrepancy may focus child attention on what fails to serve as an effective match. Although no experimental evidence for this hypothesis is available, to our knowledge, it gains credence from the fact that, in the case of therapy for phonological disorder, it is often sufficient to bring the child’s attention to bear on her systematic mispronunciations for a change to be effected (Elbert, Dinnsen & Weismer, 1984).

**ii. Segmental inventory.** Deborah’s inventory of match-to-target sounds is surprisingly small (Table 10). There are only seven consonants produced in more than one word, none of them medial. This reflects the fact that although almost half of Deborah’s word shapes are disyllabic, most of these are ‘adapted’, with the ‘adaptation’ typically affecting the medial consonant. The stressed vowel inventory is about the same size as the consonant onset inventory; unstressed vowels, like medial consonants, tend to occur in only one form as a match.

\(^{10}\) A test of this hypothesis could be carried out using the Werker and Stager novel-word learning paradigm: We would predict that children’s fast-mapping reflects their own production repertoire, so that codas or multisyllabic forms are rapidly registered as such if and only if comparable structures are available in the child’s production repertoire.
Table 10. Deborah: Segmental inventory in later words

Key: [ ] = segments produced only as substitutions for adult segment, never as match-to-target; ( ) = segments produced in only one word

Consonant inventory

1. Word initial
   t (tʃ) k
   b d g
   s (h)
   m
   (w)
2. Word medial
   (k)
   (b)
   (m)
   [w]

Vowel inventory

1. Under stress
   i u
   (u)
   eɪ, e (ou, o)
   (ɛ) û ç
   æ a
   aɪ (au)
2. Unstressed
   (i) ([u])
   (ou)
   (ʌ)
   (a)

3. Finnish: Eelis

   a. Early words (18-19 months)

   Eelis, the oldest of the children examined here at the point when word use is first established (i.e., 4 or more words in a session), has a full set of voiceless stops, a full set of front vowels and also the high back vowel [u], and both one- and two-syllable words. His vowel-initial words (4 out of the 6) are all quite accurate. The remaining two forms show adaptation of the adult target: in heppa, harmony for both consonants (replacing the adult /h…p/ pattern) and vowels (replacing the higher vowel – lower vowel sequence). In kukka the vowels are adapted, by omission of the first syllable in one variant and by imposition of the <a…i> pattern seen in äiti in the other. Furthermore, in one variant even kiikkuu, with its two high vowels, is adjusted to a <low vowel – high vowel> pattern.
Table 11. Eelis (Finnish child) Early words, 18-19 months.

<table>
<thead>
<tr>
<th>target word</th>
<th>target form</th>
<th>child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>äiti ‘mother’</td>
<td>[äiti]</td>
<td>[äiti:]</td>
</tr>
<tr>
<td>anna ‘give’</td>
<td>[an:a]</td>
<td>[an:a], [äen:æ]</td>
</tr>
<tr>
<td>ei ‘no’</td>
<td>[ei]</td>
<td>[ei]</td>
</tr>
<tr>
<td>heppa ‘horsie’</td>
<td>[hep:a]</td>
<td>[bap:a]</td>
</tr>
<tr>
<td>kiikkua ‘(is) swinging’</td>
<td>[kik:u:]</td>
<td>[kik:u], [ka:k:u]</td>
</tr>
<tr>
<td>kukka ‘flower’</td>
<td>[kuk:a]</td>
<td>[ka], [kak:i]</td>
</tr>
</tbody>
</table>

b. Later words (22 months; total words in session: 32)

i. Prosodic shape

Like many Finnish children, Eelis ‘selects’ a good many CH forms (9 out of 32 words; cf. kiikkua ‘swing, s3’, mamma ‘mama’); this is in fact the shape of all of his CVCV words. However, he has no adapted CH words. Instead, he adapts non-harmonic target words by omitting either the final (unstressed) syllable after a long vowel, to fit the template <CVV>, or the onset consonant in words with medial geminates, to fit the template <VCCV> (see Table 12). Both of these patterns also occur in ‘selected’ words; note, however, that all clusters are reduced to a single stop, even among the ‘accurate’ selected words; this stop is then realised as a geminate.

Here, as elsewhere, we disregard omission of the coda in designating forms as ‘selected’, since the omission is consistent for Eelis (despite one fully accurate variant reported for noin). Note that codas are in any case rare in Finnish (only alveolars occur in that position); Eelis is the only Finnish child to attempt any words with codas.

Table 12. Eelis (Finnish child). Later words, age 22 months; N word shapes in session = 32.

Note: Finnish orthography is close to a phonemic transcription, with ä for the low front vowel [æ] and doubling of letters to represent length.

<table>
<thead>
<tr>
<th>selected</th>
<th>adapted</th>
<th>adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;CVV&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>target word</td>
<td>child form</td>
<td>target word</td>
</tr>
<tr>
<td>ei ‘no’</td>
<td>[ei]</td>
<td>kiinni ‘closed’</td>
</tr>
<tr>
<td>puu ‘tree’</td>
<td>[pu:]</td>
<td>kiiitos ‘thank you’</td>
</tr>
<tr>
<td>tuu ‘come’</td>
<td>[tu:]</td>
<td>suu ‘mouth’</td>
</tr>
<tr>
<td>noin ‘so, in this way’</td>
<td>[noi]</td>
<td>tuo ‘that’</td>
</tr>
<tr>
<td>pois ‘out’</td>
<td>[boi]</td>
<td></td>
</tr>
<tr>
<td><a href="">VC:V</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Eelis produces no C1…C2 forms and his consonant inventory is unusually small:
Only three voiceless stops and /m/ occur in more than one word either initially or
medially. Furthermore, although he produces geminate clusters ‘accurately’, with
a perceptually long consonant, he shows no evidence of contrastive consonantal
length, as he neither targets nor produces any short or singleton medial consonants
in disyllabic forms.

Two adaptations serve to achieve the <CVV> pattern (which accounts for all 10
of his monosyllabic forms): omission of the second syllable, apparently to avoid
the C1…C2 sequence, and simplification of the coda into a long vowel. Eelis, like
both of the English-learning children discussed above, has only front-rising
diphthongs in repertoire (although both /ie/ and /uo/ are common in Finnish).

Of the onset consonants Eelis omits, one is a fricative, three are liquids and two
are nasals, all consonant types which are absent from Eelis’ productive lexicon as
sampled here. The omission of the onset may thus be ascribed to phonetic
resource limitations, but in a larger context, Finnish children omit onset
consonants far more often than do children acquiring English (Savinainen-
Makkonen, 2000). This group difference has been ascribed to the salience of
medial geminates, which attract children’s attention away from word onset
(Vihman & Kunnari, 2006; Vihman & Croft, 2007).

ii. Segmental inventory
As noted above, Eelis has an unusually small consonant repertoire (Table 13).
However, he does have the basic five (peripheral) vowels, plus one use of [æ] in
stressed syllable. In unstressed syllable only /i, u, a/ occur in more than one word.
The diphthong /oi/ occurs in two words, /ei/ in one.

Table 13. Eelis: Segmental inventory in later words
Key: ( ) = phones produced in only one word

Consonant inventory
1. Word onset
   p t k
   m (n)
2. Word medial
   p t k

| ‘duck’       | anna ‘give’    | loppu       | omit final syllable and lengthen medial C
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘sit’</td>
<td>istuu ‘sits, is sitting’</td>
<td>makkara ‘sausage’ (im.)</td>
<td>[ak:a]</td>
</tr>
<tr>
<td>‘cry’</td>
<td>itkee ‘cries, is crying’</td>
<td>nappi ‘button’ (im.)</td>
<td>[ap:i]</td>
</tr>
<tr>
<td>‘apple’</td>
<td>omppu</td>
<td>räppää</td>
<td>omit onset C and final syllable</td>
</tr>
<tr>
<td></td>
<td>[op:u]</td>
<td>[ækːä]</td>
<td>omit onset C</td>
</tr>
</tbody>
</table>
m (n)
3. Word final
   (n)

Vowel inventory
1. Stressed syllable
   i u
e o
   (æ) a

2. Unstressed syllable
   i u
   (e) (ä) a

3. Diphthong
   (ei) oi

4. French: Noël
   a. Early words (13 months)
   Noël’s first recorded words show the usual characteristics of such forms (Table 14), with a restricted set of segments in target as well as child forms and relatively good ‘matches’ between child and adult forms. That is, although here adult /k/ is produced as [t], onset /m/ is omitted in onomatopoeic miam, and /u/ is produced as schwa in the reduplicated word coucou and as a nasalized low back vowel in poum, the relatively simple adult word shapes targetted are faithfully reproduced in the child forms.

   Table 14. Noël (French child) Early words, 13 months.

<table>
<thead>
<tr>
<th>target word</th>
<th>target form</th>
<th>child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>coucou ‘peek-a-boo’</td>
<td>[kuku]</td>
<td>[tata]</td>
</tr>
<tr>
<td>miam ‘yum’</td>
<td>[mjam]</td>
<td>[ʔam], [ʔʌm]</td>
</tr>
<tr>
<td>papa ‘papa’</td>
<td>[papa]</td>
<td>[pæpæ]</td>
</tr>
<tr>
<td>poum ‘boom’</td>
<td>[pum]</td>
<td>[pɔm]</td>
</tr>
</tbody>
</table>

b. Later words (17 months; total words in session: 31)
   i. Prosodic shape
   The dominant patterns for Noël’s ‘selected’ words are CV (an example of each word type is given in Table 15) and CH (maman, papa – both close to the models - and poupée ‘doll’ [pɔpe], dodo ‘sleep (BT)’ [tʃətʃə]). For adapted words CH is the strongest pattern, affecting six words (disyllabic canard ‘duck’ [tata], chapeau ‘hat’ [papa], lapin ‘rabbit’ [papɛ], peinture ‘painting’ [papa] and trisyllabic tatato for gâteau ‘cake’ and [papapɔ] for paletot ‘coat’).
However, Noël’s most interesting patterns are those with coda or medial nasal, also shown in Table 15. This is the strongest evidence for ‘whole word representation’ that we have seen in a research study child so far: The various processes listed in the rightmost column seem to mask the likely basis for the child producing, for example, [pøn] for *peinture* and [βan] or [banə] for *ballon*. Instead of a combination of disparate processes – ‘change nasalized vowel into vowel + n, omit/adapt second syllable…’, we hypothesize that the auditory image of the whole adult form serves as a model for the child’s production pattern or routine, which we term a template, so that the template – here, `<CVN>` – provides the schema or skeletal pattern for production, while the C, V and N slots are filled from segments or features perceptible in the adult target. The variant forms [papan] for *peinture* and [tE¯tE¯] for *tiens* draw on a competing template, `<CVNV>`, in which the reduplicated syllables draw their segmental content from the adult target (recall the trisyllabic harmony words for *gateau* and *palelot*, with reduplication of the initial syllable, as here). In the case of *tiens*, the child form has an added surprise in the palatalised final nasal, which must derive from the /tj/ onset cluster of the target form.11

Table 15. Noël (French child). Later words, age 17 months; N word shapes in session = 31.

<table>
<thead>
<tr>
<th>selected</th>
<th>adapted</th>
<th>adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>selected</td>
<td>adapted</td>
<td>adaptation</td>
</tr>
<tr>
<td>&lt;CV&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>çà ‘that’ [sa]</td>
<td>[ta]</td>
<td><em>couche</em> ‘lie down’ [kuʃ]</td>
</tr>
<tr>
<td>l’eau ‘water’ [lo]</td>
<td>[ljo]</td>
<td><em>poire</em> ‘pear’ [pwaʁ]</td>
</tr>
<tr>
<td>main ‘hand’ [mè]</td>
<td>[maE]</td>
<td><em>pomme</em> ‘apple’ [pɔm]</td>
</tr>
<tr>
<td>non ‘no’ [nɛ]</td>
<td>[nE]</td>
<td><em>poule</em></td>
</tr>
<tr>
<td>pain ‘bread’ [pE]</td>
<td>[pa]</td>
<td></td>
</tr>
<tr>
<td>&lt;CVN&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>donne</em></td>
<td>[dɔn]</td>
<td><em>ballon</em> ‘ball’ [balɔ]</td>
</tr>
<tr>
<td><em>pomme</em> ‘apple’ [pɔm]</td>
<td>[pam]</td>
<td><em>peinture</em> ‘painting’ [pɛtyʁ]</td>
</tr>
<tr>
<td><em>poum</em> ‘boom’ [pum]</td>
<td>[pam]</td>
<td><em>poire</em></td>
</tr>
<tr>
<td>&lt;C1VC1VN&gt;, &lt;CVNV&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ballon</em> ‘ball’</td>
<td>[balɔ]</td>
<td>[banə]</td>
</tr>
</tbody>
</table>

11 Note, however, that palatals are also inserted into the forms for *couche* [tjø], *l’eau* and *dodo* in the absence of any ‘hook’ or model in the adult forms.
Finally, it is worth noting the extent of Noël’s variability: Most of his forms fit neatly into one of three patterns, or broad templates – CV, CH and the nasal patterns – but the same words crop up in different patterns, illustrating the ‘cross-talk’ between emergent templates that Menn and Matthei (1992) emphasized. Furthermore, there are some problematic forms: banane is produced as [bada], possibly reflecting an early word that has not been ‘updated’ to the child’s current stock of possible production patterns, since the adult word form fits so closely into the nasal templates and thus the nasals could have been expected to be produced as such. Another puzzle is the form [w¨Z¨] for fraise, since no other comparable forms occur in the session. The consonant substitutions are not so surprising in themselves – [w] for the labial fricative /f/, [ʒ] for the voiced alveolar /z/. The adaptation from closed monosyllable to open disyllable is not seen elsewhere in the session, although [bada] banane and [banœ] ballon also reflect the C₁VC₂V pattern.

ii. Segmental inventory
Noël has a very small consonant inventory (Table 16). He makes far heavier use of /p/ than of any other consonant (13 different words, at syllable onset), and also four of his six CH words feature /p/. No other consonant occurs in more than three words as match to target; /t/ occurs in 6 words, but only tiens and gâteau are a match to target at syllable onset, one word-initially, the other medially. Only nasals occur word-finally.

Noël uses both oral and nasalized vowels, but /a/ occurs by far the most often (15 words as match to target); /o/ occurs in four words.

Table 16. Noël: Segmental inventory in later words.
Key: [ ] = segments produced only as substitutions for adult segment, never as match-to-target; ( ) = phones produced in only one word

<table>
<thead>
<tr>
<th>Consonant inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Word initial</td>
</tr>
<tr>
<td>p t</td>
</tr>
<tr>
<td>b (d)</td>
</tr>
<tr>
<td>m (n)</td>
</tr>
</tbody>
</table>
1. Word medial

2. Word medial

3. Word final

Vowel inventory

1. Accented syllable

2. Unaccented syllable

5. Italian: Marco

Later words (23 months; total words in session: 40)

Marco’s most common pattern for ‘selected’ words is CV (8 out of 11 forms), but these are the only monosyllabic forms; there are no adapted monosyllables. The dominant adapted form is CH (with only a single ‘selected’ CH form produced in this session: see Table 11), but VCV is also well represented (8 out of 32 disyllabic forms). Marco’s VCV template applies primarily to long words (e.g., biberon ‘bottle’ [ibe], cappello ‘hair’ [e:lo], kasetta ‘drawer (?)’ [eta]), but shorter words with onset consonants outside of Marco’s repertoire are also produced with no onset (cane ‘dog’ [en:e], mare ‘sea’ [ale]); the form nonna ‘grandmother’ [an:a] suggests competition between the <VCV> and harmony patterns.

In addition there is an <NVCV> pattern with onset /n/, present in the only ‘selected’ C1…C2 word forms, mano and naso. If the CH and VCV forms in /n/ are added to these, we have 14 out of 29 adapted forms based on the nasal template, including VnV (cane, gallina, nonna and scarpine), nVnVn (the sole coda, for camión), nVnV (Stefano, treno), nVC2V (see Table 17), and three long forms with nasals (in campagna, non c’è latte, unaltra). In addition, most of the n-template words show unexpected substitutions ([n] for /pl/, /bl/, /stl/, /kl/, and /l/), while the change of cucchiaio to [nad:o], which fits nicely into the template, retains of the target word form only the vowel pattern of the last two syllables and the length of the /k/ (orthographic –ech–). This is the kind of idiosyncratic patterning that is difficult to account for without positing an ‘attractor’ in the form of a word template.
Table 17. Marco (Italian child). Later words, age 23 months; N words in session = 40. Italian orthography is close to the phonemic form.

<table>
<thead>
<tr>
<th>selected</th>
<th>adapted</th>
<th>adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>target word</td>
<td>child form</td>
<td>target word</td>
</tr>
<tr>
<td>&lt;C1VC1V&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tata ‘child (BT)’</td>
<td>[tata]</td>
<td>acqua ‘water’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ecco ‘here it is’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>camión ‘truck’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in campagna ‘in the country’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>occhi ‘eyes’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>palla ‘ball’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pronto ‘hello (telephone)’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>scuola ‘school’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stefano</td>
</tr>
<tr>
<td></td>
<td></td>
<td>telefono ‘telephone’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>treno ‘train’</td>
</tr>
<tr>
<td>&lt;NVC2V&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mano ‘hand’</td>
<td>[mano]</td>
<td>bello ‘nice, beautiful’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>naso ‘nose’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non c’è latte ‘there’s no milk’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>peppe ‘shoes, BT’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>unaltra ‘another’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cucchiaio ‘spoon’</td>
</tr>
</tbody>
</table>

ii. Segmental inventory
Marco’s consonant inventory is heavily biased toward coronals, with labials used only rarely. There is a near complete set of vowels; only unstressed /u/ is absent.

Table 18. Marco: segmental inventory in later words
Key: [ ] = segments produced only as substitutions for adult segment, never as match-to-target; ( ) = phones produced in only one word

**Consonant inventory**
1. Word initial
   (p) t (tʃ)
   [d] [g]
   (s)
   (m) n [n]
   (l)
2. Word medial
   t
   (b) [d]
   n
   l
3. Word final
   (n)

**Vowel inventory**
1. Stressed syllable
   i u
   e o
   ü a
2. unstressed syllables
   (i)
   e o
   a

6. Welsh: Elen
   a. Early words (13-14 months)
   Elen’s early words show a larger consonant repertoire than is typical at this stage, with fricative [s] and affricate [tʃ] in addition to both alveolar and velar stops. The expected ‘accuracy’ is violated only in the case of *clap hands*, in which only the rhythm of the phrase appears to be reproduced. It is characteristic of the early word period that in this set of seven forms we fail to find any particular pattern, although both targets and child forms are in most cases either reduplicated disyllables or CVC monosyllables. The exceptions are both *clap hands* and the trisyllabic babtalk form *gogalw*, both of which are produced as disyllables with CH.

Table 19. Elen (Welsh child) Early words, 13-14 months.

<table>
<thead>
<tr>
<th>Target word</th>
<th>target form</th>
<th>child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>choochoo (x2)</td>
<td>[tʃutʃu]</td>
<td>[tʃutʃʊə], [t^bA_t^bA]</td>
</tr>
</tbody>
</table>
**b. Later words (17 months; total words in session: 35)**

i. Prosodic shapes

Elen produces a relatively large number of different segments in her later words as well and she also has a wider range of word shapes than do most children at this point. She is also the only one of the five Welsh children to produce more trisyllabic or longer forms than monosyllables (25 vs. 10). This phonetic versatility was prefigured in Elen’s tendency to use jargon, approximating the melody of adult speech without targeting specific sound sequences, which was apparent already at 13 months. Both Elen’s jargon and her production of words of more than two syllables may reflect her sensitivity to a variety of sound patterns, or good phonological memory, or both (Keren-Portnoy, Vihman & Armstrong, 2006).

It is difficult to discern a template in the case of Elen, who has greater resources for production than most children at her developmental level. CH is her one regular adaptation, but it affects only a few forms (see Table 20) and three targetted words with harmony in the adult form are produced with differing onset consonants (choochoo with deaffrication of the initial consonant, geegee with initial [g]. La-la with omission of the initial consonant).

Table 20. Elen (Welsh child). Later words, age 17 months; N word shapes in session = 35.

<table>
<thead>
<tr>
<th>target word</th>
<th>adult form</th>
<th>child form</th>
<th>target word</th>
<th>adult form</th>
<th>child form</th>
</tr>
</thead>
<tbody>
<tr>
<td>cloc ‘clock’</td>
<td>klɔkʰ</td>
<td>[gɒkʰ]</td>
<td>beic ‘bike’</td>
<td>bɑk</td>
<td>[gaɡʰ]</td>
</tr>
<tr>
<td>dad ‘dad’</td>
<td>[dad]</td>
<td>[dɑ]</td>
<td>llyga ‘eye’</td>
<td>lɔɡa</td>
<td>[ɡaɡa]</td>
</tr>
<tr>
<td>babi, baba</td>
<td>[baba]</td>
<td>[baba]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(BT) ‘baby’</td>
<td></td>
<td></td>
<td>geegee (said to horses)</td>
<td>[dʒidzi]</td>
<td>[tʰidi]</td>
</tr>
<tr>
<td>La-La</td>
<td>[lala]</td>
<td>[ Yapıwa]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tata ‘bye-bye’</td>
<td>[tʰaː]</td>
<td>[dɑːdaː]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 20. Elen (Welsh child). Later words, age 17 months; N word shapes in session = 35.
The words produced in this session show only two consistent patterns:

1. **Velar stops** (word initial and medial [g], coda [kʰ]) are strong for Elen; they are never substituted but tend to be the substituting consonant in the cases of consonant harmony. Thus, *beic* and *llyla* are produced with velar harmony, while *biscuit* is produced with retention of the final unstressed syllable only, as [gIt]. See also *geegee* (mentioned above) and *oh God!*, an expression learned from Elen’s older bilingual siblings, pronounced ‘accurately’ as [hÅ˘|ga˘t]: This is one of five (h)VCV(C) forms.

The only other change affecting a monosyllable is the shift from CVC to open disyllable (e.g., *trwyn* ‘nose’ /trwɪn/ [tʰuna]). The source of difficulty here may be the diphthong; the only diphthong Elen produces in the session is [əu].

2. **Long words or phrases** are generally reproduced with a full syllable count, but, again, harmony is resorted to for unstressed syllables, particularly at the expense of labials. In the case of *seesaw* Elen adds syllables, as she does in other variants of the counting routine presented in Table 20 as *dau, tri, pedwar* ‘two, three, four’.

### ii. Segmental inventory

Elen’s consonants are varied in all three positions (Table 21). She produces voiceless stops both initially and as codas and has a voiced/voiceless contrast in word initial position. She uses four consonants as codas; all occur in more than one word.

<table>
<thead>
<tr>
<th>Consonant inventory</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Word initial</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>tedi</th>
<th>[tʰedi]</th>
<th>[tʰedi]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>longer words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gwgalw</td>
<td>[gogalu]</td>
<td>[gogələu]</td>
<td><em>baba fyna</em> ‘baby here’</td>
</tr>
<tr>
<td>‘bird (family word)’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>llyw fana</td>
<td>[luivana]</td>
<td>[livena]</td>
<td><em>seesaw</em></td>
</tr>
<tr>
<td>‘spoon there’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pen fyna</td>
<td>[penvana]</td>
<td>[pena]</td>
<td><em>dau, tri, pedwar</em> ‘two, three, four’</td>
</tr>
<tr>
<td>‘on top’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pêl fyna</td>
<td>[p⁶elvana]</td>
<td>[balːla]</td>
<td>‘bal here’</td>
</tr>
</tbody>
</table>

---

### Table 21. Elen: Segmental inventory in later words.

Key: [ ] = segments produced only as substitutions for adult segment, never as match-to-target; ( ) = phones produced in only one word

**Consonant inventory**

1. Word initial
2. Word medial
\[ T \]
\[ b \quad d \quad (d) \quad g \]
\[ s \]
\[ (v) \]
\[ n \]
\[ l \]
3. Word final
\[ t \quad k \]
\[ n \]
\[ l \]

Vowel inventory
\[ i \quad u \]
\[ e \quad œu \]
\[ a \]

Discussion of the evidence for whole word phonology

Extent of evidence

We have now considered data from seven children, each learning one of six languages, including two dialects of English.\(^\text{12}\) The sample is by no means planned or typical. Instead, it was designed to consider data from children whose word production patterns were not so striking as to have been chosen for illustration of the concept of ‘word templates’ in earlier studies. Furthermore, for the research studies we have sampled only from individual data sets that included at least 25 non-onomatopoeic words in a single 30-minute recording session. This means that, for all but the US English and Welsh groups, the children tend to be a little older than the average age often given for the end of the single word period, ca. 18 months – although this is precisely the mean for the six groups overall, with a total of 36 children represented in Table 4 out of 80 followed longitudinally in the various studies on which we were able to draw. Unlike examples we and others have used in the past, there is no weighting toward vocally expressive or lexically precocious children here.

This sample is thus a reasonable foundation on which to base a response to the ‘extent of evidence’ question raised early in the chapter: Does the word template

\(^{12}\) Additionally, Raivo was exposed to English from 6 months on while the UK English child Tomos had a few Welsh words in his early lexicon and the Welsh child Elen had a few English words: This is the norm for children growing up in communities where more than one language is regularly spoken by many people.
concept apply to all children? Earlier treatments have tended to emphasise individual differences, as in this judgment from Macken (1996):

Some children are highly accurate, rapid learners, show few phonological rules, and very little evidence of prosodic positional constraints on features. Some children produce simplified forms that are consistent with the operation of strong distributional constraints but produce no error forms or generalization patterns that would indicate that their surface phonotactics are represented in their systems independently of the rules or constraints that are standardly taken to be producing the surface forms. Finally, some children produce highly regular constructions that not only exemplify coherent prosodic constraints but that also indicate that the relevant constraints are encoded as representational structures, a set of surface templates that, with the underlying representations, features and rules/constraints, are psychologically real components of their respective grammars. (p. 160)

Macken’s statement implies that only a subset of children will be found to have anything that could count as a template – namely, those for whom ‘the relevant constraints are encoded as representational structures…that…are psychologically real components of their respective grammars’. The question of the generality of templates as ‘representational structures’ is quite important, as the concept cannot be used to develop an operational definition of templatic patterns or a quantifiable approach to evaluating children’s phonological development as long as it appears to characterise the emergent systematicity of only a small proportion of children. If, instead, the establishment of one or more production patterns and subsequent dependence on those patterns for rapid lexical advance is found to be characteristic of most if not all children, then efforts to develop our understanding of the range and typology of templates becomes a promising undertaking with potential clinical implications (see Velleman & Vihman, 2002).

In order to evaluate the generality of the template concept on an empirical basis we begin by summarizing the findings here, indicating what templates we were able to identify for each child and also what the ‘weight’ of the template appears to be in organizing the child’s overall word production, based on the first fifty words in R’s case but on the single later session sampled for the other children.13

1. Raivo (Estonian and English): Closed monosyllables – $C_o VF + C_o VN$ (F for fricative, N for nasal) + $C_1 VC_1$ and open disyllables - $C_1 VC_1 V$, where C is a stop in both cases ($C_0$ signifies one or no consonants). Most of the words produced in the first-50-word period, spontaneous and imitated, fit one of these patterns, with

---

13 Macken (1996) notes that ‘children have to be studied individually and longitudinally…Arguments must be based on large, representative corpora from individual children, with many forms showing each pattern. Scattered examples result in heterogenous, underdetermined, nonexplanatory lists’ (pp. 169f.). We are entirely in agreement with this view, although we stop short of longitudinal analysis of the emergent templates in the present account. We would add that some quantitative indication of the role played by any putative pattern in the child’s lexical output is highly desirable.
the first templates emerging only after a period of practice which features independent item learning.

2. **Tomas (UK English and a little Welsh):** CV\textsubscript{1}; C\textsubscript{1}VC; VC\textsubscript{1}o. Here the two templates are of unequal value. The front-rising diphthong can be considered a template because it is applied to words which lack such a diphthong in the model (bang, water; previously, milk, Postman Pat) and because it cannot be taken to constitute a ‘simplification’ in any generally understood sense of the word. Instead, the application of this pattern as an ‘adaptation’ seems to reflect the role of production practice. There are, in the session sampled, seven words with [V\textsubscript{1}] as the stressed syllable, or 18% of all words produced. The harmony template is more prevalent, even though it applies only to disyllables; it accounts for 10 forms, or 25% of all words produced in the session, including the anomalous quack-quack. Together, then, the two templates, including words ‘selected’ for the pattern as well as words adapted to it, account for 43% of the words produced.

3. **Deborah (US English):** CV\textsubscript{1}; C\textsubscript{1}VC; V. Similarly to Tomos, Deborah’s front-rising diphthong template accounts for 22% of the words produced in the session while the harmony template accounts for 32% overall, even while being restricted to disyllabic forms, which Deborah produces slightly less frequently than monosyllables. It is perhaps not coincidental that words that occur only as imitations are disproportionately represented among the adapted forms: Such forms may be taken to represent the leading edge of the child’s phonological advance.

4. **Eelis (Finnish):** CVV, VC\textsubscript{1}C; VC\textsubscript{1}. Like the other children discussed so far, Eelis has two templates, one accounting for all monosyllables (31% of the words produced in the session), the other for disyllables (the remaining two-thirds of the words produced). His CVV forms include both long vowels and diphthongs, and the diphthongs are again exclusively front-rising (although, unlike Tomos and Deborah, there are no forms ‘adapted’ to include such diphthongs). Tomos lacks any ‘adapted’ harmony forms but instead has a ‘no onset’ template, applied to disyllabic words with different onset and medial consonants (and also to one monosyllable).

5. **Noël (French):** CV, CVN, C\textsubscript{1}VC, VN, CVNV. Noël’s word patterns are more complex than those we have encountered so far. Like Raivo and Eelis (but unlike the children acquiring English), Noël produces equal numbers of monosyllabic and disyllabic forms. However, he has two templates for monosyllables – simple CV and also CVN – and two for disyllables, both involving nasals. In addition, he produces both ‘selected’ harmony forms and reduplicated longer words, but there is no strong harmony template. Most of Noël’s words fit into one or more of these templates.

6. **Marco (Italian):** C\textsubscript{1}VC\textsubscript{1}V, VCV, NVC\textsubscript{2}V. Marco produces a few monosyllables, all based closely on their models in the adult target words, but adapts no words to this pattern, which thus cannot be considered a template. He has three patterns, all resulting mainly in disyllables (there are three trisyllabic nasal-template words, all three including harmonized syllables as well). Marco’s nasal pattern, like Noël’s, is not strictly linear: The nasal substitution occurs
differently in different words, suggesting that the ‘template’ inheres rather in a strong production bias toward nasals than in a set nasal-based frame; we find both ‘selection’ of words which include a nasal and imposition of an additional nasal.

7. Elen (Welsh): C1VC1(V). As discussed above, Elen shows little dependence on any template. Harmony appears to enter into her production only sporadically, but it does affect both one- and two-syllable words and also longer forms.

Of the seven children, then, there is evidence for reliance on one or more specific, individual templates – and in most cases at least two distinct templates – for all but one, Elen; two or three templates account for the majority of forms analysed for the remaining children. Given that analyses are already available for many of the other participants in the research studies named in Table 4, illustrating the children’s use of individual templates, we conclude that the concept is of broad value, although it may not apply to all children. Equally important is the observation that the templates emerge in the course of lexical development, so that sampling too early or on the basis of too small a corpus is likely to make it difficult to identify the child’s pattern. By the same token, as indicated by Macken (1996), longitudinal analysis is needed, both to facilitate pattern identification and to make it more reliable, as the templates can be seen to develop and stabilize over a period of time, as illustrated here with the analysis of Raivo’s early words (see also Priestly (1977 [Christopher]); Macken, 1979 [Si]; Vihman & Velleman, 1989 [Molly]; Lleó, 1990 [Laura]; Vihman et al., 1994 [Alice, Timmy]; Jaeger, 1996 [Alice]; Vihman, 1978 and Brulard & Carr, 2003 [Virve and Tom, resp.], both diary accounts that encompass a more advanced point in lexical and phonological development than was analysed here).14

Typology

The summaries we have provided, although based on close analysis of only seven children, suggest an answer to the question of the degree of idiosyncracy of the templates. Several patterns recur across these unacquainted children exposed to different languages. Harmony is the single most common pattern, accounting for at least one of the templates of all but one; this comes as no surprise, as the phenomenon of CH, only very rarely observed in adult phonology, has long been reported and even proposed as a ‘universal’ (Smith, 1973) in the literature on child phonology (for analyses of several children exhibiting harmony, with differing theoretical interpretations, see Vihman, 1978; Stoel-Gammon and Stemberger, 1994; Levelt, 1994, and for additional analyses, discussion and references, Vihman, 1996, ch. 9).

Less predictable is the occurrence of two patterns which were seen here for at least two children each: front-rising diphthongs and the nasal templates. The diphthong <Vi> has previously been reported for Alice, another child learning English (Vihman et al., 1994),15 as has the disyllabic <V…i> pattern (see also

---

14 The harmony template as applied by both Virve Vihman and Laura Lleó is summarised in Vihman (1996), ch. 9.

15 One child acquiring British English, Rebecca, was reported to have a <Vu> template (Vihman, in press: cf. bye [bau], two [tou]), as was one child learning
Davis & MacNeilage, 1990), while the more general pattern <low V – high V> has been reported for children learning Estonian, French and German (Macken, 1996; Vihman & Croft, 2007). In contrast, although three children out of the seven – Raivo, Noël and Marco – exhibit a nasal pattern of some kind, nasal templates have more rarely been noted previously (Vihman & Velleman, 1989, reported a final nasal + schwa template, used to produce final nasals and also nasals occurring in other target word positions).

Finally, some templates are used by only a single child out of the seven described here. Only Raivo has CVF as a template, although Waterson identified the same template for her son (cf. also Velten, 1943). Furthermore, Brulard and Carr (2003) find such a template in their bilingual son’s English words, Macken (1996) analyses the use of such a template by Hildegard Leopold (both in English and in German), and Vihman and Croft (2007) report the pattern for a monolingually exposed child acquiring Estonian as well. These findings suggest that a template of this shape may readily be arrived at by a child acquiring a language in which monosyllabic CVC words are common, as is the case for English, Estonian and German. The affinity of fricatives for word-final position has long been noted; cf. Ferguson, 1975/78: ‘Production of fricatives is easiest to acquire in post-vocalic, final position or intervocalically, and may precede the acquisition of stops in these positions’, p. 661).

The ‘no onset’ template, similarly, is used by only two children here – Eelis and Marco, both learning languages with medial geminate consonants - but such a template has been reported for other children exposed to a number of other languages, including Estonian, French, Welsh (Vihman & Croft, 2007) and Hindi (Bhaya-Nair, 1991), as well as for other children acquiring Finnish (Vihman & Velleman, 2000; Vihman & Kunnari, 2006).

Finally, Noël’s monomoraic template is not unique: Vihman et al. (1994) provide an analysis of the longitudinal emergence of the pattern for Timmy, with a very gradual increase in the choice of onset consonants and of vowels; Vihman (in press) reports such a pattern for another French child, Camille (drawn from Veneziano & Sinclair’s 2001 study) and for both a British (Jude) an American child (Emily) not described elsewhere.

Sources of patterning
The vowels and consonants that are rare in babbling – fricatives and liquids, front rounded or back unrounded vowels, for example – tend to be avoided in the earliest words (i.e., neither produced nor attempted), although, as we have seen, some children master at least one fricative and develop a template accordingly and some develop templates based on a well-practiced liquid (e.g., Laurent, as detailed in Vihman, 1993). Similarly, both consonantal and vocalic sequences (clusters, diphthongs) also present a difficulty for many children. In all of these cases the difficulty may be purely articulatory, as both the segments not commonly found in babbling and bi-segmental sequences require precise motoric control. According to some models, articulatory difficulty is the primary basis –

Welsh, Gwyn (Vihman & Kunnari, 2006: cf. dau ‘two’ /daʊ/, produced as [bau], cacen ‘cake’ /kakɛn/, produced as [gogɛn]).
the ‘control parameter’, in dynamic systems terms - for the slow pace of early phonological development (e.g., Labov & Labov, 1978; Studdert-Kennedy & Goodell, 1995), and it is the only reason generally given by non-phonologists to account for the possible relevance of phonological to lexical development (e.g., Bloom, 2000).

In Vihman (in press) the various challenges, or points of potential difficulty for a child, and ‘opportunities’, or aspects of the language that lend themselves to early learning, are characterised for each of the languages to which the participating children were exposed. In general, the single most common difficulty is a change in place or manner of articulation in consonants across different positions in a word, which is – at least in the earliest periods of lexical growth - most likely due to the planning difficulty of preparing first one, then another consonantal gesture within a single breath group or prosodic unit, whether monosyllabic CVC or disyllabic CVCV(C) (cf. Chiat, 1989); this is the difficulty most typically met with harmony, although a melodic template – permitting a pre-planned sequence to serve for production of a range of different lexical items – or omission of a consonant can serve as well. Typically only one of these production responses is adopted by a given child (Macken, 1995), although, as we have seen, there are exceptions.

Words of more than two syllables – common among early learned words in Italian and Spanish, for example – tend to be difficult for most children, leading to various truncation patterns; for the occasional child even disyllables present a problem (cf. the French child Camille, who consistently truncates longer words to achieve her CV template: Veneziano & Sinclair, 2001; Vihman, in press). This difficulty may also be rooted in the planning aspect of speech production, but it is at least as likely that it derives from the difficulty of remembering or representing large numbers of new lexical items, arbitrarily linking a phonological pattern with a meaning of interest to the child (Vihman, 1978, 1996). Memory is as well served by a well-established routine response, whether based on harmony or melody, as is planning for production, and there is now good experimental evidence from older children to show that familiar sequences facilitate the formation of novel sound-meaning links (Storkel, 2001).

**Prosodic vs. segmental structure**

Many of the patterns we have identified apply to monosyllables only or to disyllables only, rarely to both. In some cases we have found interaction between the prosodic and segmental structure, as when Raivo’s monosyllables but not his disyllables close with a fricative or nasal. In contrast, Noël uses a nasal as a default to close words of either one or two syllables. In Vihman (in press) an Italian child was reported to use <VtV> as a default sequence in words of more than two syllables, drawing the pattern from the accented foot of her target words. In earlier studies we have noted the use of [l] as a medial onset in one French child (Vihman, 1993) and palatalized consonants as medial onsets for an

---

16 Studies of bilingual children have reported a ‘specialisation’ in words of one length per language: Bhaya-Nair’s title, ‘Monosyllabic English or disyllabic Hindi?’ (1991) expresses the effect nicely. Ingram (1981/82) reported a similar distinction for a child learning English and Italian.
American child (Vihman et al., 1994). Similarly, Priestly (1977) provided a detailed analysis of the way that his son Christopher organized his phonology around the medial onset yod for a period of 13 weeks, from 22 to 26 months (e.g., farmer \([fajam]\), tiger \([tajak]\)). Even more intriguing – since the voiceless uvular fricative is not a sound commonly associated with babbling or early words (but see its use as a prelinguistic ‘protoword’ by the son of Brulard and Carr, 2003, p. 188, n. 3) – is the \(<V\chi V>\) pattern developed by one Swiss French child, Gaël, at a slightly more advanced stage in lexical development than we have sampled here (age 21 months; 45 words in the recording session [Veneziano, unpublished data]): cf. the ‘adapted’ forms roue ‘wheel’ \([i\chi u]\), grelot ‘rattle, bell’\([e\chi o]\); the ‘selected’ forms all are transcribed with the cluster \([\chi e]\) and reflect use of the cluster to substitute for different /C\(\chi/\) clusters in target words: accroché ‘attached’ \([\chi sof\text{e}]\), écrit ‘written’ \([e\chi bi]\), ouvrir ‘to open’ \([o\chi bi]\), trax ‘steamshovel’ \([\chi\text{ras}]\).

More generally, a variety of ‘melodies’ or sequential patterns have been reported, beginning with Ingram (1974), who argued for a universal fronting pattern (see the reply by Menn, 1975). Macken (1979) provided a meticulous longitudinal account of the emergence and later fading of what could be called the ‘coronal second’ pattern of one child acquiring Spanish in California (the sequences labial-coronal and dorsal-coronal both came to be included in the template). For vowels, we have identified only one pattern in the data presented here, the front-rising diphthong, although elsewhere melodic templates restricting sequences to ‘low-high’ have been reported, as noted above. No examples of the reverse sequence have been reported as yet, to our knowledge.

While harmony requires identity of consonants or, less commonly in children (for whom vowels appear to present less of a problem), of vowels across word positions, melodies may either broadly specify place of articulation – or vowel height or backness – or more precisely require a given segment in a given word position. Table 22 summarizes studies that have identified melodic templates, covering a range of languages.

Table 22. Melodies (cases of ‘adapt’ only)
[References in brackets report data from older child with phonological disorder; cf. Edwards, 1996]

<table>
<thead>
<tr>
<th>I. consonants (C1 – C2 onsets)</th>
<th>II. vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>+/- accords/disaccords with ‘fronting’: Ingram 1974; (\frac{1}{2}) accords with ‘start easy, end hard’: MacNeilage &amp; Davis, 2000, Davis et al., 2002</td>
<td>(\frac{1}{2}) a...u (\frac{1}{2})ë...ë, o...u/ü</td>
</tr>
<tr>
<td>+, (\frac{1}{2}) [LAB – COR]</td>
<td>(\frac{1}{2}) a...e, a...o</td>
</tr>
<tr>
<td>+ [LAB – COR, VEL]</td>
<td>Salo, 1993: EST</td>
</tr>
<tr>
<td>- [COR – LAB]</td>
<td>--</td>
</tr>
<tr>
<td>+ [COR – VEL]</td>
<td>Ingram, 1974: ENG</td>
</tr>
<tr>
<td>- [COR, VEL – LAB]</td>
<td>Vihman et al., 1994: ENG</td>
</tr>
<tr>
<td>medial [w]</td>
<td>Vihman, 1996: ENG</td>
</tr>
<tr>
<td>medial [l]</td>
<td>Vihman, 1993: FRE</td>
</tr>
<tr>
<td>medial [χ]</td>
<td>Veneziano, unpublished data: FRE</td>
</tr>
<tr>
<td>final [t]</td>
<td>Brulard &amp; Carr, 2003: ENG (and FRE)</td>
</tr>
<tr>
<td>final [l]</td>
<td>Elsen, 1996: GRM</td>
</tr>
</tbody>
</table>

**Key**

COR coronal
LAB labial
VEL velar
ENG English
It is striking to note on Table 22 that, although there are positional affinities for both place (velar-final) and manner features (fricative-final), it is only place features that appear to be subject to sequential constraints. This has been discussed in the literature, from both formal (Levelt, 1994) and functional perspectives (Davis et al., 2002). To our knowledge no theoretical account has been provided for the fact that, cross-linguistically, the only widely reported constraint on vowel sequencing is <low…high>, not the reverse.

Ambient language effects

In discussing the templatic patterns of various children we have not, so far, attempted to relate them to their ‘early word’ patterns, which occur, on average, five months before the ‘later words’, chosen for illustration on the basis of a sufficient sample within a single half-hour recording. We have already noted the main characteristics of the early words: They are relatively ‘accurate’, although segments and syllables may be omitted; they reflect a highly restricted phonetic inventory in most cases (one or two stops and one nasal and/or glide; one to three or four front or central vowels, typically); and there is little patterning or system evident among the various words produced. As this brief summary suggests, there is little hope of finding ambient language effects in the early words, given the robust rootedness of all of these characteristics in the prelinguistic babbling patterns generally reported. In fact, a finer acoustic or statistical analysis is needed to identify language-particular characteristics in this period (Boysson-Bardies et al., 1989; Boysson-Bardies & Vihman, 1991; Levitt, 1993; Whalen et al., 1991).

Once the child has begun to acquire a lexicon and thus ‘move into’ or acquire active productive control of some aspects of the ambient language, however, such effects are more easily identified. Segment types and word shapes that are not produced by the majority of children learning a given language may nevertheless be singled out for default use by one or two children, reflecting the joint effect of individual child phonetic capacities and frequent occurrence in the input. For example, <CVC> templates occur only in the case of children learning languages with a high frequency of such words, such as English, Estonian, German and also Dutch (see, for example, Elbers & Ton, 1985). As a natural complement to this, children are only likely to begin producing words of three or more syllables if they are learning languages that frequently present them with such models, as is the case of Italian, Spanish and also Japanese (Vihman, 1991), for example.

Similarly, with regards to segment types, of the 33 children whose data were presented in Vihman (in press) one French child (Laurent: Vihman, 1993) avoided the use of fricatives while another (Camille: Vihman, 1996) produced them as a matter of course and a third developed a template around the uvular fricative [χ].
In the same cross-linguistic study two Welsh children and one child acquiring British English showed particular facility with fricatives in coda position, while none of the Finnish children showed a production bias in favour of fricatives.

Finally, simple input frequency does not account for all of the ambient language effects that we can observe in the shapes of individual word templates. The most striking case is that of word shapes that reflect initial consonant omission, which can be taken to contradict the ‘universal’ dominance or ‘lack of markedness’ of CV syllables, which children are accordingly expected to retain, especially when the onset consonant is within repertoire. In English the onset consonant indeed tends to be omitted only in very early word production, or in the case of unusually difficult segments, but in Finnish and Italian – both languages with geminate consonants – adaptation of disyllabic words with an initial consonant to the <VCV> pattern is quite common (cf. also Hindi: Bhaya-Nair, 1991). Vihman and Croft (2007) argued that it is the perceptual salience of the geminate for the child that leads to disregard of the initial consonant. The <VCV> pattern also occurs in French templates, although the language lacks geminates entirely: Here there is reason to believe that the iambic accentual pattern draws attention away from the onset consonant (Vihman et al., 2004); data from Brazilian Portuguese appear to support this interpretation (Oliveira, in prep.). Finally, in Welsh both the medial consonant and the final vowel are lengthened under accent in most disyllables; as we could predict, the <VCV> pattern occurs here as well (see the description of Fflur, Vihman & Kunnari, 2006). In all of these cases it appears that perceptual salience is at least partially responsible for the shape of the template – but this again must interact with particular production proclivities of the individual child, as not all children adhere to the pattern in any language group.

The strong evidence that ambient language influence, mediated both by input frequency and by prosodic salience, is filtered by the output preferences of individual children may seem obvious, but it contradicts the tendency of some investigators to conclude that the patterns they observe must be ‘universal’. Only the most basic perceptual and production factors – such as the prevalence in early words of the CV syllable, of one- and two-syllable words, and of low vowels, stops and nasals (the production repertoire practiced in babbling) – appears to have any likelihood of proving ‘universal’ when data have been collected from a far larger number of children and, in particular, from many more languages.

**Representation for perception vs. production**

‘Phonological representation’ – or lexical representation – in infancy is clearly complex and as yet poorly understood, despite the widespread use of the term (see also the discussion in Vihman et al., 1994). In an earlier chapter we saw that different inferences are drawn from experimental studies depending on differences in accentual pattern (English vs. French) and task demands (word learning vs. word recognition). The differences in task demands are important: In the case of word learning, as Werker and her colleagues have argued, significant attentional resources must be allocated to the problem of retaining the arbitrary sound-meaning link, making the task difficult, we would argue, for children who lack a stock of existing well-practiced production patterns or routines to support memory of the new word (see Keren-Portnoy et al., in prep, for strong evidence of the role of production in the construction of phonological memory). In the case of
word recognition, on the other hand, both the word form and the contextual situation or image may be expected to prime memory for the word and its associations.

Word production itself is more or less demanding depending on the context. In the case of imitation, particularly of a known word form, priming is again available. In the case of spontaneous word production, context-based word use is clearly less challenging than context-flexible use, in which a child may request an object not currently in view, for example, or refer to an event that occurred earlier in the day (Bates et al., 1979; Vihman & McCune, 1994). This latter use, which is generally accepted as reflecting the onset of referential or symbolic word use, is the most demanding of all. To reconcile the contradiction between perception experiments and the implications of production, then, we would argue that word production requires memory and planning as well as perception and articulatory capacity, leading to the use of holistic, temporarily activated representations for production which show regression in accuracy as the child comes to make use of more word types and which are clearly not characterized by ‘fine phonetic detail’.

According to Edelman’s Theory of Neuronal Group Selection (1987, summarized at length in Thelen & Smith, 1994, ch. 5), the ‘storage’ of representations (experienced connections) consists only of potential patterns of brain activity. ‘Internal representations’ can be conceptualized as potential connections that are activated in various ways - e.g., by auditory or visual priming, situational context, or purely mental communicative intention. The level of ‘detail’ will depend on the nature of priming or source and strength of representation. In dynamic systems terms (Thelen & Smith, 1994), ‘knowledge’ is not all or none but strengthens with practice or use – so it is not surprising that the most difficult task, that of accessing representations for spontaneous word production, results in the most holistic or imperfectly specified word forms in the period in which children are just beginning to build a lexicon.

The emergence of segments
The psychological reality of segments for adult speakers is generally taken for granted, although the fact that the principle behind alphabetic writing, with its one segment-one letter correspondences, has apparently been discovered or ‘invented’ only once raises some doubts as to the validity of this assumption (‘the unique historical status of the alphabetic system…must reflect the relative inaccessibility of phonetic segments as compared with syllabic units’, Vihman, 1996, p. 175). As we noted on opening this chapter, Ferguson and Farwell saw their proposal regarding ‘lexical primacy’, or what we have been calling here ‘whole word phonology’, as pertinent to adult as well as child representations, but they believed that ‘phonological elements and relations’ were also represented in parallel. The psychological reality of segments has been most strongly supported on the basis of data from slips of the tongue. MacNeilage and Davis (1990) open their paper by stating categorically that ‘serial ordering errors in normal adult speech that involve single speech segments moving around in an otherwise correct utterance tell us beyond doubt that the individual segment is an independent unit in the control of adult speech’, p. 55, our italics). Thus, the issue is not only one of representational reality. Regardless of the status of lexical or phonological representations, there is a question of motor control. In fact, the fundamental
question addressed by the Frame/Content theory of phonological development is one that follows from that statement: ‘How do infants achieve…segmental independence?’ (p. 55) – by which MacNeilage and Davis mean, essentially, how do infants break free of the reduplicative cycles of babbling and the variegated patterning that quickly follows, with changes of manner of consonants and of vowel height within single utterances but stringent C-V associations – to achieve the relative freedom of combination seen in adult languages?

This question of the move to more flexible production, which must be rooted in part in a move to more fully analysed representation, deserves to be addressed, whether or not one subscribes to the C-V association account in relation to the period that follows the onset of word use. But what longitudinal accounts are there, sufficient to provide insight into the shift from use of templates, or routinised production plans, to more varied and adult-like sequences, with more specifically targetted substitutions of individual consonants or vowels? In 1996 Vihman commented of Macken’s 1979 case study that ‘although there are other detailed longitudinal accounts of one child’s phonological development…no other study has so clearly traced the path from the ‘word as a prosodic unit’…to a more adult-like system based on the principle of phonemic contrast’ (p. 221). More than a decade later this remains the case. In fact, there have been, to our knowledge, no new attempts to address this question on the basis of the observation of children’s word production. A new empirical departure can be reported, however, in the form of a massive compilation of ‘kids’ slips’, or developmental speech errors, together with a box-and-arrows developmental account of speech production and representation (Jaeger, 2005).

**Learning mechanisms**

In discussing the polarization of researchers concerned with language development into empiricists and nativists Braine (1994) came to the logical conclusion that any scientist wishing to account for language learning without positing specifically linguistic innate knowledge (i.e., without recourse to the concept of Universal Grammar) would have to posit, instead, powerful ‘innate learning mechanisms’. Studies such as Morgan, Shi and Alloppenna (1996) and Saffran, Aslin and Newport (1996), among others, have now shown ‘statistical’ or distributional learning by infants to be one such mechanism (see also Kelly & Martin, 1994, whose review of the literature and experimental studies anticipate those findings).

**Distributional or statistical learning**

The conceptual distinction between explicit and implicit learning is not new (see Reber, 1967; Ellis, 1994), but only in the past few years have experimental findings made it clear that children, like adults, automatically tally distributional regularities in the environment (Saffran et al., 1996; Kirkham, Slemmer & Johnson, 2002). In experimental studies adults and older children have been shown to ‘pick up’ sequential patterning to which they are incidentally exposed while attending to a completely different task (Saffran, Newport, Aslin, Tunick & Barrueco, 1997). Other studies have demonstrated implicit learning for segmentation of the speech stream and for learning syntax, in research examining both adults exposed to artificial languages (Reber, 1993; Saffran, Newport &
Aslin, 1996) and infants exposed to sequences of syllables produced in a monotone (Saffran, Aslin & Newport, 1996) or, more naturally, with variations in stress and in coarticulation (Johnson & Jusczyk, 2001). These kinds of study reveal probabilistic (statistical, distributional) rather than categorical learning or ‘symbol manipulation’ (see the review of infant artificial language learning in Gómez & Gerken, 2000).

If we generalise these findings to other experimental studies of prelinguistic responses to speech in the ambient language (e.g., Jusczyk, Cutler & Redanz, 1993; Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1994; Jusczyk, Luce & Charles-Luce, 1994; Mattys, Jusczyk, Luce & Morgan, 1999; Mattys & Jusczyk, 2001), we can conclude that infants gradually gain a sense of input language patterning as regards sequences at any level of linguistic organization – segments, syllables, accentual patterns, words, phrases, clauses. The studies demonstrate a powerful capacity for distributional learning, or sensitivity to statistical regularities in the environment, including a capacity to induce the purely formal linguistic patterns that constitute input speech. This learning capacity is not restricted to speech (i.e., is not ‘domain specific’), however, but has been shown to be ‘domain general’ in the sense that the learning automatically applies to any regularly recurring sequence in the infants’ environment (Kirkham et al., 2002).

Based on the studies of adults, it is clear that this implicit learning occurs in the absence of any specific intent to learn or even of (conscious or focussed) attention to linguistic patterning as such. Studies of infant responses to speech have revealed emergent sensitivity to prosodic coherence in §ever smaller prosodic units over the course of the first year (clauses, in infant directed speech, as early as 4.5 months [Kemler Nelson, Hirsh-Pasek, Jusczyk & Wright Cassidy, 1989], then phrases at 9 mos. [Jusczyk, Kemler Nelson, Hirsh-Pasek, Kennedy, Woodward & Piwoz, 1992], and finally words at 11 mos. [Myers, Jusczyk, Kemler Nelson, Charles-Luce, Woodward & Hirsh-Pasek, 1996]). Such developmental effects can also be taken to be the result of implicit learning. Although this kind of learning is not strictly speaking ‘statistical’ or ‘distributional’, it is not arbitrary, symbolic, or based on attention either. It can safely be termed implicit learning, or learning in the absence of voluntary or focussed attention to the stimuli, intention to learn, or conscious awareness of learning. While the criterion of ‘consciousness’ is not helpful in relation to infants (but is intuitively the most easily grasped characteristic of adult ‘explicit’ experience), the related notions of learning without voluntary attention or intention to learn, can be applied to observations of infant behaviour and can be experimentally manipulated.

The effect of implicit perceptual learning in infants can also be seen in their vocal production. Specifically, implicit learning is the evident source of the subtle ambient language effects on vowel production revealed by acoustic analyses of the babbling of infants exposed to British English, French, Arabic or Cantonese (Boysson-Bardies, Hallé, Sagart & Durand, 1989; see also Whalen, Levitt &

---

17 Another characteristic of this kind of learning is its relative inflexibility of access (Squire & Kandel, 1999). That is, what has been learned can be retrieved only under closely matching conditions; there is no generalisation.
Wang, 1991, who found that rising pitches were more common in the babbling of French than American infants in the age range 6-12 months, and Boysson-Bardies & Vihman, 1991, who reported a larger proportion of labials in the vocalizations of 10-month-olds exposed to English and French than in those exposed to Japanese or Swedish; in each case the ambient language pattern is reflected in these early infant differences in production). The mechanism is less direct here, as production is only a secondary reflection of the child’s perception of adult speech. To account for the effect we must assume that infants are biased to selectively reproduce with greatest frequency the vocalizations that they perceive as better matches to what they are hearing with greatest frequency or consistency in input speech. The effect should then be seen only in patterns that have a solid grounding in the infants’ production repertoire, i.e., that are also produced with sufficient frequency to be subject to such a ‘pruning’ effect of the perception/production match – and indeed this is the case, as phonetic categories less commonly produced by most infants in the babbling period (such as syllable-final consonants: Vihman & Boysson-Bardies, 1994) are influenced by the distributional frequencies in the adult language at a later stage of phonological learning than are phonetic categories that come under infant control earlier (such as medial geminates: See Vihman & Velleman, 2000, for the similarity between children exposed to English and French, on the one hand, and Finnish, on the other, at the earliest stage of word production, and for the sharp differentiation of Finnish from the other two groups as infants exposed to English and French restrict the length of their medial consonants by the time that they are producing 50 or more words while Finnish children increase theirs).

Lexical or symbolic (categorical) learning (‘explicit’ or ‘declarative’)

Lexical learning is not the same as statistical learning, however. Only a few years ago one could have gained the impression that learning of any aspect of language occurred only in conjunction with attention. The phenomenon of ‘joint attention’ (between caretaker and child) has received a good deal of experimental study and is widely accepted as constituting one of the foundations for language development (Tomasello, 2003). Evidence that attention in early infancy can predict later levels of intelligence or word learning have also long been sought (e.g., Ruddy & Bornstein, 1982; Rose, Slater & Perry, 1986), and some promising correlations have been reported between infant patterns of visual habituation (an index of attention) and later word learning (Colombo, Shaddy, Richman, Maikranz & Blaga, 2004).

In fact, there can be no doubt but that word learning does depend at least in part on infant attention. Specifically, before the first referential or symbolic word learning is demonstrated – in which a child comprehends or produces an adult-based word pattern in novel situations, showing an ability to extend the form-meaning pairing beyond the situation in which it has already been heard – that child must have attended to adult use of the word in a situation in which both

---

18 This is not to say that implicit learning of aspects of both form, the phonetic basis of word learning, and meaning, the semantic basis, is not relevant to word learning. It seems clear that implicit learning is an ongoing part of an infant’s life from the time that he or she is capable of receiving sensory information, i.e., in the case of auditory signals, from the last trimester before birth.
word form and use were clear and salient. Indeed, as indicated in Chapter 2, before word use becomes well established children have been shown to point, ‘show’, and grunt in intentional communication. By the first half of the second year infants typically demonstrate an interest in language and an intention to learn as well as a capacity for explicit attention to, and memory for, word use (see Ch. 2).

Furthermore, in mothers’ speech to their one-year-old infants, words used repeatedly in isolation – that is, words on which the children are more likely to have focussed their attention – have been shown to correlate significantly with the first word production (Brent & Siskind, 2001; see also Ninio, 1992). Thus, despite the well-established abilities of prelinguistic infants to make use of prosodic, phonotactic, and coarticulatory cues to segmentation (Jusczyk, 1997), words or phrases that are readily available to the child’s attention without the need for segmentation of the speech stream seem to be more readily incorporated into the emergent production lexicon. It is plausible, then, to make a distinction between learning with and without attention: This would correspond roughly to the distinction between explicit and implicit learning in adults – although we should also emphasize that even in adults it is difficult to make a sharp or categorical distinction between the two types of learning or memory on the basis of behavioural evidence, as neither task nor test can be guaranteed to be ‘pure’ or free of complementary memory effects in a laboratory experiment (Jacoby, 1991).

Current thinking in neuroscience supports the idea of a dual memory system. It is widely accepted (based on studies of amnesia as well as on animal studies: e.g., Squire, 1992; Baddeley et al., 2001) that the hippocampus is required to consolidate detailed, multimodal episodic memories, which are the basis of any one-off learning from unique experiences. Furthermore, the registering and recall of arbitrary form-meaning pairs depends on processing in both the frontal lobes (known to be involved in the selection of percepts for focussed attention) and the hippocampus, which alone is capable, in adults, of rapidly learning conjunctions of associated elements of experience (McClelland, McNaughton, and O’Reilly, 1995; Squire & Kandel, 1995; Ullman, 2004; for a review, see Ellis, 2005). In contrast, the registration of regularities – the essence of distributional learning – occurs even in the face of hippocampal damage, permitting patients to abstract structure from a set of related items, for example (Knowlton & Squire, 1993). Slow skill learning (based on sufficient motoric practice) occurs without involvement of the hippocampus (e.g., Wilson, Maruff & Lum, 2003) as does the gradual learning of repeatedly occurring perceptual consistencies (e.g., Nissen & Bullemer, 1987; Willingham, Nissen & Bullemer, 1989). Based on connectionist modelling, McClelland et al. argue that two distinct learning or memory systems evolved for a good reason, namely, to permit rapidly learned novel patterns to be added to an existing system without catastrophic interference (p. 432f.).

The evidence thus solidly supports a distinction between two types of learning – one with, the other without voluntary attention or ‘intention’ to learn, one categorical, symbolic, the basis for a stable lexicon or store of words and phrases, the other probabilistic, statistical, sensitive to distributional properties such as frequency of occurrence and sequential patterning (Ellis, 2002). In the processing of longer linguistic units exceptions to regular patterns (e.g., irregular morphology) may also depend on declarative memory (Ullman, 2001, but for
differing views see McClelland xxx [add refs]). But what are the implications of all this for language development?

We suggest that three types of learning should be distinguished for first language learning. Two of these three are the product of ‘implicit’ processing by the ‘non-declarative’ system (Squire & Kandel, 1999); the third requires processing by the declarative system. The three types of learning are the following:

1. Procedural tallying of regularities perceived in sensory data of any kind (e.g., of frequencies of occurrence, sequencing, or rhythmic patterning),
2. Declarative (categorical, symbolic) registering of arbitrary form-meaning co-occurrences or associations, leading to a mental lexicon of linguistic items, and
3. ‘Secondary’ procedural induction (a kind of ‘tallying’) of the regularities inherent in the linguistic items registered in the mental lexicon, ultimately resulting in abstract knowledge of linguistic system.

Procedural, distributional or statistical learning, based on the gradual registration of recurrent regularities that is characteristic of neocortical processing (i.e., direct sensory processing, in the absence of the consolidation of independent sensory sources of information for which hippocampal processing is required), is sufficient to account for infant advances in knowledge of the ambient language that have been reported for the first year of life (1). The units over which such learning occurs will change as the child gains increasing knowledge, first of prosody (especially rhythmic patterning: Nazzi, Bertoncini & Mehler, 1998; Nazzi, Jusczyk & Johnson, 2000; Ramus, 2000), then of segmental sequences, leading to a capacity for more detailed processing of the ambient language. This kind of learning requires no preestablished ‘knowledge base’ (Murphy, McKone & Slee, 2003) and can thus begin to ‘inform the child about the world’ just as soon as the infant’s sensory organs are complete.

The declarative learning of linguistic units, or referential (symbolic, generalized) word learning – based on the second memory system and thus the second type of processing and learning, is not typically found until the first half of the second year (2). Unprimed or ‘flexible’ (‘explicit’-like) access to lexical items requires, at a minimum, symbolic understanding, or the ‘nominal insight’ (individual word tokens or exemplars are understood to refer to word categories or ‘types’), so that each new encounter with a given word form in an identifiably related situation is taken to belong to the same lexical category. (The live dog barking next door can be referred to using the same word form as the stylized doggy found in the picture book or on the side of a cup, say.) This understanding, and thus this kind of access to the mental representation of experiences, cannot be expected to emerge until a stable base of frequently heard forms of words or phrases has developed, along with a sufficiently rapid and stable capacity for

---

19 Following Squire & Kandel and Ullman, we will generally use the terms ‘procedural’ and ‘declarative’ here, in lieu of the more widely used terms ‘implicit’ and ‘explicit’, to avoid the difficulties of applying the notion of ‘explicit’ or ‘conscious awareness’ to infants for whom no such awareness can be clearly established.
retaining both phonetic and semantic representations. This is necessary to free up the attentional resources required for declarative learning: Attention is needed to enable the child to relate new forms to new referents (Werker et al., 2002). Once such voluntary access to lexical representations becomes possible, a lexical knowledge base will begin to be established, somewhat different for each child but nevertheless, within a given linguistic group, sampling from the same adult patterns and thus converging, over a period of years, on a lexicon very similar to that of other members of the speech community.

Thirdly, once a child has begun to establish a lexicon of words or phrases with both phonological form and semantic content, ‘secondary’ procedural or distributional learning will automatically occur, as the neocortex again goes to work on recurrent regularities (3). The input to that implicit or procedural processing is now no longer at the level of direct perceptual input but is rather a representational derivative of the cognitive processing that created the lexical entries – hence the term ‘secondary’. As Karmiloff-Smith (1992) has put it, ‘a specifically human way to gain knowledge is for the mind to exploit internally the information that it has already stored…. by redescribing its representations or, more precisely, by iteratively re-representing in different representational formats what its internal representations represent’ (p. 15). Beyond that difference in ‘raw material’ or input to the processor, however, the learning process itself may be assumed to be the same as in (1). This secondary procedural learning can now be understood as gradually building up the abstract knowledge of system or structure to which we generally apply the term ‘grammar’ (Pierrehumbert, 2003; Beckman & Pierrehumbert, 2003). This kind of pattern induction may be sufficient to account for the learning of the abstract categories that are the basis for phonological and syntactic systems without the need for innate ‘foreknowledge’ of linguistic structure in the form of Universal Grammar.

In this chapter we have elaborated on and illustrated a conception of phonological development that assumes that ‘primary’ (1, above) and ‘secondary’ (3, above) procedural learning are separate contributions, with declarative learning (2) serving as the mediator between the two. Figure X – adapted from Vihman and Kunnari, 2006 – illustrates the ‘spiral model’, by which procedural or implicit learning ‘sets the stage’ while declarative or explicit learning adds concrete lexical items to the mix. Once the process has functioned repeatedly to this point, yielding a small lexicon, procedural or implicit memory is triggered again, resulting in new levels of phonological knowledge. The process may be supposed to function over the life-span, although new lexical learning becomes less frequent in the native language once an adult-like level has been achieved, typically in the teenage years, perhaps.
References


Edwards, 1996


Kõrgvee, K. (2001). Lapse sõnavara areng vanuses 1;8-2;1. [A child’s lexical development, aged 1;3-2;1] Undergraduate thesis, Tartu University.


Salo, A. (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.


