1 Introduction: the emergence of phonology: whole-word approaches, cross-linguistic evidence

Marilyn M. Vihman and Tamar Keren-Portnoy

Whole-word phonology is a particular approach to early phonological development. This volume is designed to bring together the classic papers which gave rise to it in the 1970s and current studies that build on and extend the model, which in essence took an emergentist and usage-based stance before its time; the book will make no attempt to cover other approaches to phonological development in any systematic way. Many of the papers, including Vihman and Croft (2007, this volume, Chapter 2),\(^1\) with which we begin, use the term “template” to refer to child-specific word patterns identifiable within the first year of word use. Templates, referred to sporadically in the earlier developmental literature (e.g., Menn 1983, this volume, Chapter 6) and given formal status for adult linguistic analyses in Prosodic Morphology (McCarthy and Prince 1995), are a more focused expression of the ideas formulated by Waterson (1971, this volume, Chapter 3), Ferguson and Farwell (1975, this volume, Chapter 4), and Macken (1979, this volume, Chapter 5), which provided the core of the whole-word phonology idea (see Vihman and Croft 2007, this volume, Chapter 2, for a summary of the basic arguments).

This volume is restricted to the study of early word production and the phonological patterning that can be seen in that domain. The year in which the first of our “setting papers” was written – Waterson (1971, this volume, Chapter 3) – also marks the year of publication of the first study of infant speech perception (Eimas, Siqueland, Jusczyk, and Vigorito 1971). Since then, perception studies have solidly documented infants’ remarkable early discriminatory capacities and the rapid advances in knowledge of the ambient language that follow over the first year of life (see Jusczyk 1997; Kuhl 2004; and Vihman forthcoming 2014 for reviews), while numerous studies demonstrating infant statistical learning (in language and other areas) from an early age have expanded our understanding of the learning mechanism that may underlie those advances (see Thiessen and Saffran 2007, and Johnson and Tyler 2010 for alternative positions on the role of statistical learning; Vihman forthcoming 2014: ch. 5 provides an overview). In addition, several distinct methodological procedures have been used to trace and explore the nature of early word-form learning over the first two years of life.\(^2\) The resultant studies are of evident relevance to phonological development but none are included here, as the addition of even a few would result in a far
longer and less focused volume (and the studies are readily available elsewhere). Nevertheless, what we have learned about perception and early learning capacities is critical to our understanding of the course of phonological development and clearly complements the “whole-word approach” presented here.

Whole-word phonology and templates in phonological development

As understood in the chapters that follow, templates involve (idiosyncratic) prosodic structures that appear to be generalized, in different ways by different children, from the forms of a child’s earlier babble vocalizations and first words. Templates typically lead to increased similarity in the forms of the child’s words at the expense of accuracy (i.e., of match to the adult target form). This corresponds to a sequence of, first, item (or exemplar) learning, then distributional learning, implicitly and automatically applied to repeatedly used child output forms – the presumed source of the “generalizing” of patterns to new targets. Taking an exemplar model perspective, this generalization can be thought of as the self-organization of the exemplar space, due to connections being formed between similarly shaped child forms; an alternative (but not necessarily incompatible), strictly sensorimotor perspective sees the “generalization” as no more than the automatization of one or more well-practiced procedures, namely, the child’s emergent neuromotor word-production routines (McCune, this volume, Chapter 16). The resultant patterns appear to constitute (unconscious or implicit) child responses to the phonological challenges posed by target word forms. In other words, the child’s existing resources (familiar production routines) are deployed to deal with what is novel and thus difficult to bring to mind, plan, and produce as needed. Although this understanding of the function of templates and of the mechanism underlying “generalization and analogy” (Macken 1979, this volume, Chapter 5, p. 144) is relatively recent, the core papers depict essentially the same learning sequences and the same conclusion as to the role of templatic patterns as a way of dealing with challenges by bringing familiar routines to bear on them.

Vihman and Velleman (2000) introduced the terminological distinction between “selected” words, or child word forms that (roughly) match the form of their adult targets while conforming to a child’s preferred prosodic structure, and “adapted” words, or word forms based on adult targets that are less similar to the child’s pattern, which the child thus modifies more radically to arrive at an output that fits the template. Examples of both selected and adapted words can be found in many of the chapters of this book (see also Keren-Portnoy, Majorano, and Vihman, 2009). The earliest papers make no mention of the term “template,” let alone of “selecting” and “adapting,” yet the detailed data presented by Priestly (1977, this volume, Chapter 7), for example, make it easy to see that some words, such as lion, produced as [lajən], and whale [wεjəl] (“bisyllabic ordinary forms” in Priestly’s terms), are “selected” in our sense,
while others – the “bisyllabic experimental forms,” which Priestly found to be “not only amusing but systematic” (p. 217) – are “adapted”: e.g., berries [brjas], chocolate [kajak], peanut [ pijat], and tiger [ tajak].

From the child’s point of view, there is presumably no essential difference, except perhaps of degree, between the two kinds of words: things that are similar are treated similarly. The targets for “selected” words are similar to other “selected” word targets as well as to the child’s own forms of those words. The targets for “adapted” words are not as obviously similar to one another, yet they must sufficiently resemble other words rendered within the framework of that particular template to “attract” the child into associating them with the same type of “own” (child) form. It is typical of the forms used under the influence of a child’s dominant template that no attempt (by researchers) to separately trace or relate each segment to its presumed model in the target word will yield a satisfactory analysis (this is well exemplified by the data in both Waterson 1971 and Macken 1979, this volume Chapters 3 and 5 respectively, as both investigators emphasize). Instead, we see the child matching the overall shape of the adult word (CVC[C][V][C] in Priestly’s examples), often including the target syllable count, as here, and at least one of the consonants, while simplifying the overall structure through repetition of segments or syllables or through reordering to achieve a fixed output structure for multiple lexical items (here, CvVC). In short, the term “template” is used to formalize the notion of “whole-word learning” as the basis of a child’s phonology.

It is important to note that templates are not a lasting element in a child’s phonological system, even for children learning the classic “templatic” adult language, Arabic (see Khattab and Al-Tamimi, this volume, Chapter 14). Instead, templates typically gain increasing dominance over a period of days, weeks or months – often beginning toward the middle or end of the single-word period – but then fade thereafter, as the child comes to master (in terms of articulation, speech planning, and memory or representation) the more complex sequences of the adult language: see Priestly’s and Oliveira-Guimarães’ accounts of the rise and fall of templates in the phonological development of one English and two Brazilian children respectively (this volume, Chapters 7 and 10), as well as Macken (1979, this volume, Chapter 5), for the emergence of templates and the subsequent advance to accurate segmental sequences in the speech of a Spanish-learning child, and Vihman and Vihman (2011), a longitudinal account of the emergence, use, and fading of two templatic patterns in a diary study of an Estonian- and English-learning child’s first 500 words. Finally, note that the templatic shape itself is dynamic, changing in more or less subtle ways over the period of time in which it holds sway as the child’s phonological knowledge increases and stabilizes, often with a period of competition between variant “solutions” to the phonological challenge (see Priestly 1977, this volume, Chapter 7; Macken 1979, this volume, Chapter 5; Vihman and Velleman 1989, this volume, Chapter 8; Vihman, Velleman, and McCune 1994, this volume, Chapter 9; and Oliveira-Guimarães, this volume,
Chapter 10, as well as Menn and Matthei 1992, who discuss competition in child rules or patterns).

Universals vs. typological and individual differences: the role of rhythm

How does the child get started learning the phonetics and the phonology of the ambient language? What resources are available for “kick-starting” the process? It is worth considering the role of rhythm, in both perception and production, as a theoretical and developmental starting point for the child, and one which may go some way toward accounting for three separate aspects of child vocal production: its initial “universality,” its typological variability by language of exposure, and the individual differences found even within a single language group, all of which are amply illustrated in the chapters of this book.

The earliest theoretical statements about the course of phonological development – those of Jakobson (1941/1968) – were based on diary studies, with their inevitable focus on the individual child and his or her early word production. Nevertheless, the conclusions of that highly influential first attempt at systematization, heavily shaped by the structuralist theoretical principles of the Prague School of linguistics of which Jakobson was a key member, were meant to serve as putative universals. Somewhat later, Brown (1958) provocatively hypothesized that babbling – thought by Jakobson to be unrelated to later phonological development – involved a phonetic “drift” in the direction of the ambient language. It was only later still, when the wide availability of first audio and then video recording devices made possible far more reliable and detailed phonetic observations of children’s speech and especially of their prelinguistic vocalizations, that the wide range of individual differences in pathways to language (even for children acquiring the same language) began to become evident from production studies (see, e.g., Vihman, Macken, Miller, Simmons, and Miller 1985; Vihman, Ferguson, and Elbert 1986; Menn and Vihman 2011). All three of these characteristics of phonological development must be encompassed in our understanding of this complex process: universals, or the commonalities to be found in the babble and first word production of children learning any language; ambient language effects and their implications for the mapping of what is perceived onto vocal production; and the variability due to the contribution of the individual child, within the constraints of perception, the neurophysiology of vocal production, and cognitive development.

Perceptual experience of the dominant rhythms of the ambient language can be taken to provide a phonological frame suitable for supporting first word forms (see Wauquier and Yamaguchi, this volume, Chapter 11, for evidence of the impact of rhythm on template formation in French). In other words, perceptual experience of the specific rhythms of the language will yield a typical one- or two-syllable unit, based on stress and syllable type and weight, which a child’s immature and inexperienced phonological memory will retain and
use, first in implicit segmentation (Nazzi, Iakimova, Bertoncini, Frédonie, and Alcantara 2006; Höhle, Bijeljac-Babic, Herold, Weissenborn, and Nazzi 2009; Pons and Bosch 2010), then in early attempts at production; this in turn will tend to strengthen the patterns that the child has tuned into, resulting in more ambitious targeting of adult words (i.e., of word targets beyond the child’s production abilities), which are thus adapted to a well-practiced pattern or template. (For evidence that phonological memory is constructed through use, see Keren-Portnoy et al. 2010.) The cross-linguistic data provided in this volume are largely consistent with this proposal, as child templates are shaped by target language affordances whose scope is typically a lexical unit (a word or a short phrase) in interaction with the child’s own babbling practice and first word production experience (through selecting and adapting).

The evidence from templates suggests that rhythm is critical here, providing a perceptual “envelope” into which the child’s individual production patterns can be fitted. As Brown anticipated, individual children’s vocal practice (babble) gradually “drifts” toward (or is shaped by) the rhythms of input speech (Boysson-Bardies, Hallé, Sagart, and Durand 1989; Boysson-Bardies and Vihman 1991); this implicit sensorimotor experience of babbling is a critical mechanism for transforming heard speech patterns into the production base for word learning – a different base for different children, despite broadly similar input and neuromotor constraints. It is this prosodic framing of speech sequences that eventually leads to the individual but ambient-language-influenced phonological templates.

In contrast to the implicit shaping of babble by perceived input speech, the integration of what is heard with what can be produced as learned word forms is neither automatic nor effortless. Furthermore, because this integration will depend on such individually variable factors as the particular characteristics of a child’s babble, emergent representational ability, and volubility or sociability, among other things, we should not be surprised at the wide variability identified in production even among children learning the same language. In general, the patterns that we find described in this volume, for one or more children per language, broadly reflect the prosody of the individual language and support the notion that rhythm is an important starting point for phonology (for further discussion, see Wauquier and Yamaguchi, this volume, Chapter 11).

Whole-word learning from the perspective of an exemplar model

In what sense is the phenomenon that we have been describing “whole-word learning”? The child’s rendition of the word shows sensitivity to some of the segments or phonetic features that occur in it. However, it does not necessarily maintain the order in which the segments or features appear, but may instead redistribute, merge or spread some of those features. This is seen as whole-word learning because it is within the lexical unit (a word or a short but often repeated
The phrase: *all gone, in there, what's this?* (that this lack of conformity to the identity or ordering of parts is observed. Within the lexical unit there may be no clear evidence that the child has registered information about the identity and number of all of the segments (as perceived by adult speakers) or their relative order. Based on evidence from production, then, children seem to have a representation or memory trace of the adult form, but that representation is not constructed out of an ordered sequence of segments.

This claim, that the child’s representation lacks a clear structure made up of neatly ordered parts, has often been misunderstood. The failure to appreciate what is meant here has led some researchers to ascribe to the proponents of “whole-word” phonology, or holistic representations, the claim that such representations are “vague” or “underspecified” (Gerken, Murphy, and Aslin 1995; Swingley and Aslin 2002, 2007; Storkel and Maekawa 2005); holistic representations are contrasted here with segmentally detailed representations (Storkel and Maekawa 2005) that are characterized by phonetic specification (Swingley and Aslin 2002). Gerken et al. (1995) aptly present this viewpoint: “Children represent early words in terms of holistic properties, such as prosodic structure and acoustic shape, or in terms of phonetic features that are not bundled into individual segments” (p. 476). In fact, as we understand them, these child representations include abundant detail—much more than is apparent in phonemic or even broad phonetic description.

Taking an exemplar model perspective on whole-word learning, let us consider what whole-word learning might be like. As suggested by Pierrehumbert (2003), the perceptual input for speech is “an auditory coding of the speech signal. A covering map provides an analog representation of the phonetic space, with the dimensions being the many phonetic parameters which are relevant to speech perception” (p. 132; see also Edwards, Munson, and Beckman 2011). Thus, for infants the representation is highly detailed, perhaps hyper-detailed, or even overly detailed in some aspects but less so in others. In addition, since infants need not at first know which acoustic parameters are relevant for speech perception, they may assign weights to parameters differently than adults would.

Something Pierrehumbert does not mention, but which may also affect information processing in the young child, is salience: parts of the acoustic signal which are less readily perceived (shorter, lower pitch, quieter—typically, unstressed) may be processed less successfully, with more error or more loss of information (as shown in Vihman, Nakai, DePaolis, and Hallé 2004). Since unstressed parts of words also tend to be produced with increased motoric variability (Goffman, Gerken, and Lucchesi 2007), the unstressed parts of different exemplars of the same word would differ more, leading those parts to be less coherently represented; that is, their representations would contain more variability or noise. In a noisy or variable exemplar space the treatment of a newly encountered exemplar as belonging, or not, to the particular category will be less consistent.
Note that our claim is that young children’s exemplar space is sparser and more variable than adults’, with less clearly defined clumps or categories, and that it therefore functions with less clearly defined boundaries for what does or does not fall within each category. The more frequently a child encounters—or produces—the exemplars of a given lexical type or structure, the sharper will be the organization of the corresponding portion of exemplar space (see Ota, this volume, Chapter 15).

There is no vagueness or lack of detail in this scenario. What is lacking is segmental organization, or a tidy organization into sequentially ordered time-bound units, each built of a unique co-occurring set of features. It is this abstract level of categorization that is missing, not fine detail. In this sense the child’s representations, based on the evidence of phonological templates, is both richer and poorer than what is implied by standard phonetic transcription: It is rich in featural texture but poor in sequential organization. Nor does the interpretation of child phonological representations as lacking segmental units constitute a problem for “continuity” between child and adult phonological knowledge, to the extent that some theoretical models similarly deny any such organization for adult representations (Browman and Goldstein 1989, 1991, 1992; Pierrehumbert 2003; Edwards, Beckman, and Munson 2004; Edwards, Munson, and Beckman 2011; Munson, Edwards, and Beckman 2012).

The orientation of this volume

In the 1970s three papers appeared that have since become classics: Waterson (1971, this volume, Chapter 3) took a Firthian approach to one child’s phonology and introduced the notion of “schemas,” or child-specific word patterns; Ferguson and Farwell (1975, this volume, Chapter 4) argued for whole-word or lexical patterns as the core of adult as well as child phonological knowledge; and Macken (1979, this volume, Chapter 5) demonstrated the unusual adult-to-child-form mappings that can be found in early phonology to meet the child’s constraints on output forms. These papers all stood outside of phonological theory as it was understood at the time, shortly after publication of the definitive statement of generative phonology, Chomsky and Halle (1968). As it happened, that formalization was about to be superseded by the range of new perspectives that emerged in response to the perceived limitations of Chomsky and Halle’s approach (see Van der Hulst and Smith 1982; Anderson 1985; Goldsmith 1995; and Scheer 2013). This period in the study of phonological development culminated in the widely cited paper by Menn (1983, this volume, Chapter 6), who adopted a psycholinguistic perspective and formulated the “two-lexicon model” (for a rethinking of this model, see Menn and Matthei 1992; Menn, Schmidt, and Nicholas 2009, as well as this volume, Chapter 17).

In the period that followed, phonological theory blossomed and expanded, diversifying into a range of distinct theories, including CV phonology (Clements...
and Keyser 1983), Lexical Phonology (Mohanan 1986), Autosegmental and Metrical Phonology (Goldsmith 1990), Dependency Phonology (Durand 1990), Government Phonology (Kaye, Lowenstamm, and Vergnaud 1990), Declarative Phonology (Coleman 1998), and most recently CVCV phonology (Scheer 2004). However, during the 1990s one of the new models, Optimality Theory (OT: Prince and Smolensky 1992/2004), began to dominate the field, to the point that it came often to be the only theoretical perspective presented to linguistics students. A number of attempts have been made to cast phonological development in terms of OT (see Boersma and Levelt 2003); the studies collected in Kager, Pater, and Zonneveld (2004) are dedicated to the presentation of acquisition data from an OT perspective. Yet no extensive OT treatment of data from one or more children has appeared to date. The present volume returns to the whole-word phonology approach, which has much in common with the early work of McCarthy and Prince (1986, 1993, 1995) but which, on the basis of extensive cross-linguistic studies of child data, diverges sharply from OT, with its reliance on Universal Grammar and markedness theory and the tendency of its advocates to expect linear advances along with set stages of development and across-the-board changes in child forms.

As Menn and her colleagues point out in their recent efforts to model what we know about how children learn phonology (Menn et al. 2009; see also Menn et al. this volume, Chapter 17), any adequate theory of phonological development must be able to account for three key findings, all solidly grounded in forty years of empirical research:
1. individual differences across children,
2. lexical variation within a given child,
3. the phenomenon of regression (nonlinear advance, or the U-shaped curve), in which early accuracy is succeeded by less accurate, more child-specific word forms – only to be followed, much later, by a return to adult-like forms, or relative accuracy.

No theoretical approach that sees phonological development as the automatic suppression of innate processes (Stampe 1969), across-the-board changes in rule application (Smith 1973), triggering of parameters (Fikkert 1994) or reordering of constraints (various chapters in Kager et al. 2004) can account for these core characteristics in any straightforward way.

**Vihman and Croft** (2007, this volume, Chapter 2) propose a new way of thinking about phonology, based on the ideas of Charles Ferguson and the evidence from child data, with specific reference to the three characteristics listed above. *Menn and her colleagues* (this volume, Chapter 17), who focused on OT in their 2009 critique of what has been missing in theories of phonological acquisition to date, now propose to extend Vihman and Croft’s “exemplar model” by including some key missing elements – namely, (i) a role for representation of the adult target form; (ii) mappings from input to output (corresponding to the influential “rules” or “processes” of earlier generative models such as Stampe 1969 and Smith 1973); and (iii) mappings from output
to input, this latter an important characteristic not yet incorporated in any purely phonological theory.

The concluding section of this volume includes both Menn et al.’s future-oriented proposals for extension and revision of Vihman and Croft’s model and McCune’s developmentally oriented thoughts on the importance of the study of phonological development as a whole, and of early word templates in particular, for child language development. These discussion chapters clash on a number of specific points, which can only be a healthy symptom of the liveliness of creative thinking in our field, even within the scope of a broadly similar theoretical (here, functionalist) inclination. Beneath the evident differences—in defining the notion of “representation,” in particular—the editors of this volume find some deeply rooted similarities. Specifically, the interrelatedness of representations in a network of potential associations, both formal and meaning-related, within a broadly neurological framework seems to us to emerge from both of these chapters. Based on this idea we can conceptualize the “potentiality” or transitory nature of representations, such that they only come into existence in moments of what, in adults, could be termed consciousness (a better expression, for developmental purposes, might be “moments of use,” for either speaker or listener). To return to the exemplar metaphor, individual instances (or their subparts) are activated to differing degrees in different situations of use, when activation “lights up” differing elements in the network of associations. This would mean that the “representation” of a given item has no essential stability over time, especially in the early period of phonological development. It is our hope that the contrasting perspectives provided will lead to discussion, debate, and further empirical research.

The contents of this volume

Three basic considerations guided the choice of papers for this volume. First, in Part II (Setting papers) we sought to provide a fair representation of the core papers that gave rise to the whole-word approach. Secondly, in Part III (Cross-linguistic studies) we included empirical papers that work through the implications of this approach. We limited ourselves to data-oriented papers in which the child’s word forms are presented in sufficient numbers to give the reader a clear understanding of the shape of his or her emerging phonology, excluding papers with only anecdotal mention of specific forms to support a rule or constraint. Most of the chapters in this volume also take an overtly “whole-word approach,” but that is not the case with all of them (neither Ota, Chapter 15, nor Priestly, Chapter 7, have any such explicit orientation, for example).

Thirdly, this volume is specifically designed to provide data and analysis exemplifying the ways in which these fundamental properties of phonological development are manifested in a range of different languages and child learners. The diversity of languages included makes it possible to document, for
example, the exceptional salience of geminates for children acquiring languages that have them – here, Finnish (Savinainen-Makkonen, Chapter 13) and Arabic (Khattab and Al-Tamimi, Chapter 14) – or the unavoidability of learning clusters early in a language in which they are particularly common (and one child’s solution to that problem: see Szreder, Chapter 12). (Of the chapters in Part II, only Macken’s, Chapter 5, concerns a language other than English.) This will enable a serious student of child phonology to see just what kinds of data need to be accommodated, not in one child or one language alone but in a broad sample – and also to see the lines of similarity, patterns, and limitations or constraints that recur in one child and one language after another, although not always in the form predicted by adult-theory-based notions of markedness (see also Vihman and Kunnari 2008). However, the typological study of child phonological development is certainly still in its infancy. We hope that this volume will stimulate empirical studies of children learning a far wider range of languages.

We believe that whole-word phonology (in Ferguson and Farwell’s terms), or Templatic Phonology (Vihman and Croft 2007), provides a model that, while still limited in many ways, is at a minimum faithful to the evidence afforded by large quantities of child data. This is one of the key legacies of Charles Ferguson’s approach: the empirical data are allowed to speak – although the interpretation will necessarily be influenced by the investigator’s training and habits of mind. We hope that this edited volume, with its mix of classics and both old and new data-based papers as well as contemporary re-evaluations from the points of view of both linguistics and developmental psychology, will bring the whole-word phonology approach to the attention of a new generation of linguists, psychologists, psycholinguists, and speech scientists.

Notes

1. Papers included in this volume are indicated in bold face; those published here for the first time are cited without mention of year of publication.

2. Some of these latter studies are specifically designed to challenge the notion of “whole-word phonology”; we discuss below some of the ways in which this idea has been interpreted (or misinterpreted). Although we cannot here enter into a discussion of the differences between experimental responses to a limited number of stimuli and spontaneous speech production, Vihman, DePaolis, and Keren-Portnoy (2009) discuss the issue briefly, while Vihman (forthcoming 2014: ch. 7) is devoted to “Experimental studies of word form learning.”

3. Smith (1973), still the most extensive analysis of any one child’s phonology, came out in the same period. Smith’s study held closely to the mainstream formalization represented by Chomsky and Halle and purported to demonstrate – in direct contradiction to the studies included in this volume – that there is no basis for assuming that the child’s word forms reflect an independent system.

4. Some of these models overlap with or subsume others.