Article:

Vihman, Marilyn orcid.org/0000-0001-8912-4840 (2017) Learning words and learning sounds:Advances in language development. British journal of psychology. pp. 1-27. ISSN: 0007-1269

https://doi.org/10.1111/bjop.12207

British Journal of Psychology

Advances in language development: Learning words and learning sounds

--Manuscript Draft--

Article Type:	Article
Manuscript Number:	
Full Title:	Advances in language development: Learning words and learning sounds
Order of Authors:	Marilyn May Vihman, PhD
Corresponding Author:	Marilyn May Vihman, PhD University of York York, Not Applicable UNITED KINGDOM
Corresponding Author E-Mail:	mv509@york.ac.uk
Keywords:	speech sounds; phonological development; word learning; exemplars; complementary systems model; perceptual narrowing; phonological template; vocal motor scheme
Abstract:	Phonological development is sometimes seen as a process of learning sounds, or forming phonological categories, and then combining sounds to build words, with the evidence taken largely from studies demonstrating 'perceptual narrowing' in infant speech perception over the first year of life. In contrast, studies of early word production have long provided evidence that holistic word learning may precede the learning of speech sounds. In that account, children begin by matching their existing vocal patterns to adult words, with knowledge of the phonological system emerging from the network of related word forms. Here we review the evidence from production and then consider how the implicit and explicit learning mechanisms assumed by the complementary memory systems model might be understood as reconciling the two approaches.
Additional Information:	
Question	**Response**
If you have any potentially competing interests to declare, please enter them in the box below. If you have no interests to declare, please enter 'none'.	n/a
Does this submission have any links or overlap with any other submitted or published manuscripts, for this or any other publication? (For example; as part of a long-term project, using a shared data set, a response to, or extension of, earlier work). If yes, please give brief details. If no, please enter 'none'. Any overlap not declared and later discovered will result in the manuscript being withdrawn from consideration.	n/a
Please specify the word count of your manuscript (excluding the abstract, tables, figures and references).	8057 words

Running head: **Advances in language development**

Advances in language development: Learning words and learning sounds

Marilyn M. Vihman

[1]University of York

Acknowlegement: Thanks are due to my colleagues Paul Foulkes and Amanda Cardoso, who provided useful feedback on a draft of the paper, as well as to audiences at the ESPP conference in Tartu, Estonia, and the Child Language Seminar in Warwick, England, where versions of this paper were presented in the summer of 2015.

Word count (exc. figures/tables): 8057

*Requests for reprints should be addressed to Marilyn Vihman, Language and Linguistic Science, University of York, Heslington, York YO10 5DD, England (e-mail: marilyn.vihman@york.ac.uk).

Abstract

Phonological development is sometimes seen as a process of learning sounds, or forming phonological categories, and then combining sounds to build words, with the evidence taken largely from studies demonstrating 'perceptual narrowing' in infant speech perception over the first year of life. In contrast, studies of early word production have long provided evidence that holistic word learning may precede the learning of speech sounds. In that account, children begin by matching their existing vocal patterns to adult words, with knowledge of the phonological system emerging from the network of related word forms. Here we review the evidence from production and then consider how the implicit and explicit learning mechanisms assumed by the complementary memory systems model might be understood as reconciling the two approaches.

In accounts of the nature of language acquisition the phonological aspect is often overlooked. Yet its role is necessarily foundational: Knowledge of a certain minimum of words and phrases is an essential basis for learning grammar, and learning words means gaining knowledge of speech forms and of the links between those forms and their meanings, which must be deduced from their situations of use. The rapid advances in phonological development of the first 18-24 months are thus an important element in early word learning. Here we discuss the nature of those advances, the relation between development in the production and processing of speech sounds and whole-word units and the mechanisms that underpin human learning over the lifespan. The central questions that will guide the overview concern the first units and how they are learned: Do infants begin by learning speech sounds and then combine them to recognize and produce words? Or do they begin by producing word-like vocalizations and retaining bits of the speech signal that match their production? Or do these processes occur in parallel?

Before beginning on a review of developmental studies we should briefly consider two prominent contrasting views of adult phonology. The formalist view takes the segment or phoneme (or the bundle of distinctive features that make up the segment or phoneme) to be basic to linguistic structure (Chomsky & Halle, 1968; Halle, 1971; Blevins, 2004), whereas the functionalist view sees units linked to meaning (whole word forms) as basic to both phonological structure (e.g., Pierrehumbert, 2001, 2003; Port, 2007; Vihman & Croft, 2007) and speech production (Redford, in press). A fundamental principle of structuralist linguistics was the key role of phonemic contrast, the minimal speech-sound opposition needed to distinguish meaningful units (Saussure, 1915/1959; Bloomfield, 1933; Jakobson, 1941/68). Contrast was taken to underlie the abstract representation of speech in terms of segments or distinctive features and was also part of the conceptualization of language that led Chomsky (1965) to argue that language was not learnable without innate knowledge of its organizational principles, foreknowledge that has come to be encapsulated in the idea of Universal Grammar (Chomsky, 1981). Models of speech production grounded in neurolinguistic evidence are sometimes misled into taking the formalist view as representing all linguistic approaches, missing the current controversies surrounding the once seldom disputed status of the phoneme or its surrogate, the distinctive feature bundle (e.g., Hickok, 2014).

Although phonemic oppositions undeniably play an important role in distinguishing linguistic units in any language, the emphasis on phonemes as 'minimal units of distinctive sound function, forming a unitary inventory within a language and concatenated with one another in an additive way to form words' (Anderson, 1985, p. 292) met with fundamental disagreement early on: In an important monograph Twaddell (1935) observed that 'words (not segments) are the minimal free forms of a language which stand in contrast' (Anderson, 1985, p. 292). More contemporary critiques emphasize the dynamic nature of language in use in contrast to its idealized conceptualization. In usage-based phonology, for example, Twaddell's promotion of words over individual sounds has its echo in the finding that the same phonological sequences may have different ranges of phonetic variation in connected speech in apparently homophonous pairs of words with differing frequencies of use (cf., e.g., *four/for, can* (auxiliary)/*can* (main verb) and even, though less dramatically, *time/thyme, right/write,* etc.: Bybee, 2001; Pierrehumbert, 2001; Johnson, 2007).

Similarly, sociolinguistic identification of sound changes in progress (Weinreich, Labov & Herzog, 1968; Labov, 2001) has long since cast doubt on such structuralist ideas as the maxim, 'once a phoneme, always a phoneme'. In fact, current work in sociophonetics has shown the importance, for explaining variability, of speakers' memory for word exemplars, which not only retain phonetic detail (rather than an abstract, minimally redundant sequence of phonemes) but also the socially relevant indices of the individual speaker that underlie accommodation and change (Foulkes & Docherty, 2006; Foulkes, 2010). Furthermore, both psycholinguistic and neurolinguistic research have made it abundantly clear that the avoidance of redundancy, another aspect of structuralist theory that was carried over intact into generative phonology (Chomsky & Halle, 1968), is not a psychological but a purely theoretical concern, given the extensive spare capacity of the human brain and the value, for robust processing, of multiple access to representations (Wedel, 2007; Menn, Schmidt & Nicholas, 2013). Also, by the 1990s psycholinguists had begun to question the absolute validity of categorical perception, noting that under the right circumstances listeners are able to hear within-category distinctions, although this is not typically available to conscious access (Miller, 1994; McMurray, Tanenhaus & Aslin, 2002). Thus an understanding of how infants begin to gain knowledge of language need not start from the assumption that phonemes or individual speech sounds are indisputably the first elements of language structure to be learned.

Infant speech perception and 'perceptual narrowing'

The study of infant speech perception was initiated with the finding that speech sound contrasts are discriminated categorically from the first months of life (Eimas, Siqueland, Jusczyk & Vigorito, 1971). Contrary to initial interpretations of this finding in terms of an innate human specialization for language, subsequent demonstration that chinchillas, macaques and dogs discriminate phonetic contrasts in the same categorical way suggested that special sensitivity to certain regions in the speech signal may be built into the auditory system that humans share with other mammals (Stevens, 1972, 1989, 1998); this led to the idea that categorical perception may reflect evolutionary auditory shaping of the phonology of human languages (Kuhl, 1986). Somewhat surprisingly, however, speech sound discrimination proved to be superior in infants, with their limited auditory experience, in comparison with the related function in adults, with their more narrowly circumscribed facility that discriminates just those contrasts that characterize their native language (Werker & Tees, 1983, 1984).

The developmental course of perceptual discrimination contrasts sharply with that of the uniquely human ability to produce the core syllables basic to the phonology of the world's languages. Although that ability is absent at birth it develops rapidly, typically appearing in identifiably adult-like vocal production by 6-8 months (Oller, 2000). This key production milestone is followed, within a few months, by an emergent capacity to represent, recall and produce the forms of words, with word-form recognition often preceding full word comprehension (Hallé & Boysson-Bardies, 1994; Vihman, Nakai, DePaolis & Hallé, 2004; Swingley, 2009b). Furthermore, a range of different studies have shown that word production and use provide a more stable, more reliable, better-established representation than word recognition or comprehension alone, whether in adults or in children (Keren-Portnoy et al., 2010; MacLeod et al., 2010; Vihman, DePaolis & Keren-Portnoy, 2014; Icht & Mama, 2015; Zamuner, Morin-Lessard, Strahm & Page, in press). In addition, expressive vocabulary constitutes a strong predictor of lexical advance: What is known already affects the way the brain processes what is new. This has been shown indirectly in studies of processing speed (e.g., Swingley, 2009a; Fernald & Marchman, 2012) and eye-tracking (Horváth, Myers, Foster & Plunkett, 2015) as well as in direct measurement of brain function (Torkildsen et al., 2008, 2009). How are these various advances interrelated? What is the key function for word learning? And how does the emergent function of speech-like production relate to the ability to process speech sounds as phonological categories?

Infants' 'universal' discriminatory capacities in the first 6 months of life are well established: In the somewhat artificial conditions of repeated syllable presentation in a laboratory experiment infants readily discriminate consonantal contrasts and also vowels, whether they occur in the native language or not (see Vihman, 2014, ch. 3, for a review). The first perceptual 'advance', however, is an early *regression* in discrimination: Success, among groups of infants aged 6-8 months, in hearing differences between a variety of different speech sounds has been robustly shown to contrast with *failure* to discriminate phonological categories not distinguished by the ambient language by groups of infants only slightly older, aged 10-12 months. This shift in perceptual processing has been tested mainly in infants exposed to English (e.g., Werker & Tees, 1984; Best, 1994), but it applies equally well to infants exposed to other languages, including Japanese (Kuhl et al., 2006), Arabic (Segal, Hejli-Assi & Kishon-Rabin, 2016) and Urdu ([redacted], submitted).

No fully satisfactory explanation for this early (and rapid) loss of a generalized capacity to detect segmental distinctions has been provided as yet. Instead, various plausible accounts have been offered, based on developmental shifts that occur around the same time. These include infants' emerging capacity for voluntary attention (due to maturational changes in inhibitory control: Tipper, 1992; Ruff & Rothbart, 1996), their dawning responses to meaning (Huttenlocher, 1974; Bates et al., 1979; Benedict, 1979; Bergelson & Swingley, 2012, 2013), their increasing skill in adult-like syllable production (Vihman, 1992; Davis & MacNeilage, 1995, 2000) and their ongoing implicit learning of the distribution of speech sounds experienced in the input ('distributional learning': Maye, Werker & Gerken, 2002).

Speech sounds before words

Distributional learning is currently the most widely accepted source of the shift in perceptual discrimination. Maye et al. (2002) demonstrated 6- and 8-month-old infants' sensitivity to uni- vs. bimodal differences in distribution in a brief lab experiment: By editing and resynthesizing recorded tokens of English [da] and [ta] (excised from spoken *sta*) they created eight CV-stimuli evenly spanning the acoustic continuum from voiced to voiceless unaspirated alveolar stops (cf. Pegg & Werker, 1997). They familiarized two groups of infants with these stimuli, mixed in with four tokens of [ma] and [la]. One group heard more repeats of tokens in the middle of the range (4 and 5: 'unimodal exposure'), the other group more repeats of tokens toward the extremes of the range (2 and 7: 'bimodal exposure'). When tested with the extreme tokens, 1 and 8, which had been presented to the same extent in both cases, only the infants provided with bimodal exposure discriminated the stimuli, at either age. This suggests that separate categories were formed only in that condition.

The potential relevance of this experiment for the issue of perceptual narrowing in speech sound discrimination is clear. Contrasting sounds can be expected to cluster separately in any language, with minimal overlap, while similar sounds that do not contrast are likely to be more diffusely distributed. According to this model, infants will naturally form phonological categories from denser clusters of sounds; contrasts falling outside of these categories will no longer hold their attention. This could explain why, for example, Japanese infants no longer discriminate English /r/ from /l/ by the end of the first year (Kuhl et al., 2006), Arabic-learning infants no longer discriminate Hebrew /p/ from /b/ (Segal et al., in press), Urdu-learning infants no longer discriminate English /v/ from /w/ ([redacted] et al., submitted) and English-learning infants no longer discriminate the velar and uvular ejectives of the Interior Salishan (Native American) language, Nthlakapmx, also known as Thompson (Werker and Tees, 1984), or the voiceless unaspirated and voiced labial stops of Zulu (usually described as voiced vs. implosive labial stops: Best & McRoberts, 2003).

Does this mean that infants begin by learning sounds and contrasts, and are only subsequently able to begin to register and represent word forms? The fading of discriminatory attention to infrequent or non-occurring category contrasts is accompanied by a sharpening of frequently experienced category boundaries, according to Maye et al. (2002) and Kuhl et al. (2006, 2008). Thus the growing strength of representation of individual phonological categories (contrasting segments or phonemes) has been taken to provide the critical underpinnings for knowledge of word forms (Kuhl, 2004; Werker & Fennell, 2004).

On the other hand, unsupervised distributional learning is possible but demonstrably difficult for adults (Goudbeek, Swingley & Smits, 2009) and is insufficient in itself for inducing discrimination of some phonetic contrasts (Cristià, McGuire, Seidl & Francis, 2011). Furthermore, analysis of a good-sized corpus (700 single vowels produced by one mother to her 10-month-old) revealed a far greater extent of overlap in the distribution of distinct vowels than this model would predict (see Fig. 2, Swingley, 2009b). Both Swingley and Cristià et al. conclude that learning based on acoustic cue distributions alone is unlikely to be sufficient to account for infant learning of the phonetic categories of their language.

If we assume that children are learning speech sounds within lexical contexts, the problem becomes more tractable (Swingley, 2009b, Fig. 3). However, the issue of how infants first learn to recognize word forms in the early word-learning period remains unresolved. There is ongoing debate as to whether they begin by picking up statistically frequent sequences, independent of any meaning function, and gain knowledge of the accentual system of the language based on that learning (Thiessen & Saffran, 2003, 2007), for example, or whether some aspect of prosodic (accentual) structure is primary instead (Johnson & Jusczyk, 2001; Johnson & Tyler, 2010). An additional possibility, disfavored by most specialists in the area of infant word segmentation, is that the relatively small proportion of isolated word forms used in infant-directed speech (typically assessed at 9-10% of all words used, if 'non-syntactic' words such as *uh-oh, wow, yum-yum* are disregarded*:* e.g., Brent & Siskind, 2001), provides the infant with a 'wedge' into the speech stream, with access to highly familiar lexical units boosting attention to adjacent units (Bortfeld, Morgan, Golinkoff & Rathbun, 2005; Swingley, 2009b; [redacted], submitted).

Importantly, the phenomenon of perceptual narrowing has been found to occur, within the same time-frame, as part of category formation in a far broader range of cognitive domains, such as the discrimination of musical changes embodied in unfamiliar musical traditions or of individual monkey faces or human faces representing unfamiliar races (cf., e.g., Pascalis, DeHaan & Nelson, 2002; Scott, Pascalis & Nelson, 2007; Lewkowicz & Ghazanfar, 2009; Maurer & Werker, 2013). Within these quite different domains, maturational as well as experiential changes in attentional capacities must be relevant alongside any distributional factors.

Attentional shifts must be linked to developmental changes in what is meaningful for the infant, socially or affectively as well as referentially or semantically (i.e., in relation to word meanings). For example, infants fixate on faces for the first few months of life; their growing knowledge of faces, combined with the powerful emotional experiences associated with them, is a critical part of the process of widening social engagement, a foundational aspect of being human (Boysson-Bardies et al., 1993). Similarly, infants' advances in experience of 'action' or purposeful movement, which support their growing sensorimotor knowledge of the physical world, also support conceptual advances (Thelen &

Smith, 1994). All of this can be assumed to be involved as well in the apparent category formation that results in perceptual narrowing, inasmuch as the linked cycles of (self-)action and perception have been shown to underlie so much of cognitive and social as well as motoric development (see Campos et al., 2000). Thus the emergence of adult-like vocal production in the middle of the first year of life, in the form of the first CV syllables ([bababa, dadada, ŋaŋaŋa]), could be expected to affect infant speech processing as well, focusing infant attention on selected (matching or sufficiently similar) portions of the input speech stream and thus potentially playing a role in the fading of the early 'universal' capacity to discriminate phonetic differences.

Whole words before speech sounds

An alternative theoretical approach to phonological development is to assume that children do not learn speech sounds directly at all. Instead, they learn whole word patterns, with knowledge of those speech sounds frequently experienced in familiar words later emerging out of the representational network of known words of similar length and/or with similar onsets, rhymes and codas. This assumption derives primarily from production studies, which provide ample evidence that the first words are typically learned as whole items or sound-patterns (Vihman & Keren-Portnoy, 2013). However, Werker and Curtin (2005), whose PRIMIR model of phonological development makes virtually no mention of production, nevertheless similarly propose that 'once the infant has established a sufficient number and density of meaningful words, generalization of commonalities occurs, leading to the emergence of the Phoneme plane' (214).

A related conceptualization of a network of connections emerging from individually known lexical items underlies Bybee's Network Model of morphology (1985, 2001, 2010) as well as exemplar models of phonology (Bybee, 2001, 2010; Pierrehumbert, 2001; Wedel, 2007). Similarly, Munson, Beckman and Edwards (2011) see phonology as emerging from 'generalizations over the parametric phonetics and generalizations over the lexicon' (37; see also Beckman & Edwards, 2000a, b; Munson, Edwards & Beckman, 2012; Plummer & Beckman, in press).

A whole-word perspective on lexical and phonological knowledge can also be seen in studies of both toddlers and older children. The studies of both Storkel (2001) and Edwards, Beckman and Munson (2004) provide good evidence of the facilitative effect of familiar phonotactic sequences on novel word learning in children aged 3 to 6. This effect is indirectly echoed in the finding, with older children (8-12 years), that word-finding difficulties are most likely to affect words in low-density neighborhoods, as regards both access and errors in production; the difficulty of access is understood as resulting from a lack of sufficient use of those words and their concomitantly weaker network links: 'Those segments and segment combinations are not accessed frequently, and thus, may have relatively underdeveloped paths.' (German & Newman, 2004, 633)

The conceptualizations of phonological development as beginning with sounds or with whole words seem to clash, yet there is good reason to believe that each of these accounts is at least partially correct (see also Swingley, 2009b). How might evidence from perception studies supporting the early distributional learning of speech sounds be reconciled with the evidence from production studies for whole-word learning? We will review the evidence from production studies of the first 18 months. We will then consider what learning mechanisms might be able to account more satisfactorily for the evidence from both perception and production studies.

Word production, I: Item learning and 'pre-selection'

The idea that emergent control over vocal production might affect infants' processing of speech was first proposed as a way to account for the fact that infants' first words are surprisingly accurate (as noted by Ferguson & Farwell, 1975). That is, the first words may show some omission or substitution of consonants but, generally speaking, they constitute simple matches to comparably simple one- or two-syllable target words, as illustrated in Table 1 (see the first words of 48 children learning 10 languages, Appendix I, Menn & Vihman, 2011).

[Insert Table 1 about here.]

The phonetic repertoire seen in the first words is the same as that which characterizes babble, which may be considered unconscious practice for word production (Vihman et al., 1985). As seen in Table 1, the consonants are largely restricted to stops and nasals, glottals and glides (there are no fricatives or affricates here; exceptionally, one child produces [l] in three of his five words) and the forms rarely include more than a single supraglottal consonant type. Aside from Annalena's [data] for *das da* (with a change in the voicing of the stop, an aspect of production not typically controlled at this age: Macken, 1980), the only exceptions are Alice's [m̩ːɑnːə] for *mommy* and Kaia's 'progressive idiom', [kḭ:tɔ̥] for *kiisu,* the only word identified for this child for 6 months. (The classic example of a progressive idiom is Hildegard's whispered rendition, at 10 months, of *pretty,* which included the initial cluster and the change of consonant place across the word; the form was adjusted to fit her emergent phonological system, a year later, as [bɪdi]*:* Leopold, 1939; Ferguson & Farwell, 1975/2013, p. 106.) In the remaining words in Table 1 with more than a single 'true' (supraglottal) consonant in the target form we find consonant harmony, or full consonant agreement across the word, either in both adult target and child form (see Alice's *baby* and all of Annalena's and Kaia's remaining words, with – in Kaia's case – changes to the onset to match the medial consonant).

Menn and Vihman (2011) comment on the relatively unsystematic nature of the first word forms seen in almost half of the children whose data they present: Those children 'seem to have acquired a *word-length complex of gestures as an unanalyzed whole'* (p. 271; our italics). In other words, these children give evidence of learning words before learning speech sounds. Such early representations may be 'overdetailed' (in the sense of rich, segmentally unanalyzed exemplars: Houston & Jusczyk, 2000, 2003; see also Jusczyk, 1986, 1997, ch. 8), at least in part, and yet also sketchy in part (Hallé & Boysson-Bardies, 1996: Vihman et al., 2004). In general, infants' first words are similar to babble and build on that vocal practice (Vihman et al., 1985; McGillion et al., in press); the relative accuracy of these words strongly suggests 'pre-selection' (Ferguson, Peyser & Weeks, 1973; Ferguson & Farwell, 1975).

The relationship of emergent vocal production skills to speech processing*:* The articulatory filter hypothesis

The presumed phenomenon of pre-selection calls for explanation. To account for it Vihman (1991, 1993, 1996) proposed that, once children have begun to produce adult-like syllables on a regular basis, they may experience as particularly salient those frequently heard input forms that resemble whatever is most frequent in their own vocal output. Vihman referred to this concept as the 'articulatory filter', to express the idea that the child unconsciously filters what she hears in the input through her own production experience. (See Kleinschmidt & Jaeger, 2015, for parallel evidence, based on computational modeling, of the importance of familiarity in adult perceptual processing of speech – although neither production nor self-monitoring is considered in Kleinschmidt and Jaeger's 'ideal adapter' model.)

The proposal is justified in part on the grounds that the child's own vocalizations, themselves guided or primed by often-heard words or patterns in the speech stream (Boysson-Bardies & Vihman, 1991), will have a double effect on the child, being experienced as both an auditory and a proprioceptive stimulus (Vihman et al., 2014). This should strengthen the child's representation of speech forms that resemble her typical production (i.e., forms that constitute a rough match of input and output):

> The child may be seen as experiencing the flow of adult speech through an 'articulatory filter' which selectively enhances motoric recall of phonetically accessible words. (Vihman, 1996, p. 142; for recent experimental evidence of a specific articulatory effect on the processing of speech as early as 6 months, see Bruderer, Danielson, Kandhadai & Werker, 2015)

This hypothesis remained purely speculative until DePaolis, Vihman and Nakai (2013) ran an experimental study of 53 children acquiring either English or Welsh in North Wales. DePaolis et al. recorded infant vocalizations in the home four times over a two-month period, beginning at age 0;10.15 (10 months, 15 days). Two weeks after the last session, at about 1;0.15, they tested the infants in the lab, using the head-turn preference procedure (Kemler-Nelson et al., 1995). They presented the infants, in several trials, with two randomized lists of nonwords, each making repeated but varied use of one of two supraglottal consonants that are equally frequent in the input but that are expected to differ in extent of child use in production at this age ([t] vs. [s] for English, [b] vs. [g] for Welsh).

In the event, as expected, the 27 English-learning children made variable use of [t/d][1], ranging up to a mean of over 50 tokens in a single session, but seldom used [s], while the 26 Welsh children failed to show any real difference in use of [p/b] vs. [k/g], which were both produced a mean of about 15-20 times per session. (The Welsh children, like the English, produced [t/d] more often, with a mean of just over 30 tokens in a single session.) The results were consistent with the hypothesis that production affects speech processing, but not in the predicted way: The English children with the highest production of [t/d] in the final session showed greater interest in /s/, the speech sound they were producing only rarely, if at all, than in /t/, the speech sound they were most familiar with through production; only children with lower [t/d] production in that session showed greater interest in the stop than in the fricative. The Welsh

[1] No attempt was made to tally separately the voiced and voiceless members of a stop pair, as they are not reliably controlled at this stage, as noted above.

children, whose production was not well differentiated for the two speech sounds tested, showed roughly the same level of interest in both sets of stimuli.

This experiment demonstrated for the first time an effect of infant vocal production on speech processing, although the paradoxical nature of the effect, which had not been anticipated, meant that further research was needed. Two subsequent experiments made use of the individual differences consistently seen in infant vocal production to test the articulatory filter idea more directly. In order to more specifically test whether infants match their own patterns to input speech these studies adopted from McCune and Vihman (2001) a measure of consistency (or identifiability) and stability of vocal production, the 'vocal motor scheme' (VMS), which picks out recurrent and stable speech-sound use. McCune and Vihman found that, despite Jakobson's (1941/68) well-known claim that babble includes a wide variety of possible speech sounds, in reality infants only very gradually exhibit VMS use, which is taken to reflect speech-sound mastery (i.e., the voluntary control needed to allow word production), for increasing numbers of different consonants.

DePaolis, Vihman and Keren-Portnoy (2011)[2] recorded infants in their homes, beginning at 9-11 months, and transcribed the sessions as quickly as possible in order to obtain evidence of consistent, stable use of a single speech sound (VMS) as soon as the skill emerged, to allow for timely testing. Operationally, VMS use was defined as 10 uses of the same consonant, disregarding voicing differences in the case of stops, in three recording sessions (following McCune & Vihman, 2001), or 50 or more uses in a single session (as this level of use seemed

[2] The 2013 study was actually run earlier but was published only two years after this study, run a few years later.

sufficient to attest to VMS mastery); vowel use was not assessed because it is too variable in babble, and transcription of infant vowels too unreliable, to allow repeated use to be identified with confidence. Eighteen infants met the criteria in the home recordings. The experiment was subsequently replicated with 26 Italian children, first seen at around 6 months and then recorded in the home longitudinally from the onset of canonical babbling (between 7 and 11 months) until at least one VMS was identified (Majorano, Vihman & DePaolis, 2014). In the English study infants were tested on short passages featuring nonwords with either a VMS the child was using ('own VMS'), a different possible VMS that the child was not using ('other VMS'), or a labiodental fricative, to control for the effect of a speech sound none of the children were likely to be using with any frequency ('non-VMS'). In the Italian study infants were tested with word lists (as in DePaolis et al., 2013), similarly contrasting 'own-', 'other-' and 'non-VMS'. As the findings of the British and Italian studies are similar we report them together here.

The results were consistent with the English/Welsh study, in that the infants fell into two groups, depending on their level of VMS knowledge. Those with more than one VMS in repertoire were significantly more interested in 'other-' than in 'own-VMS' (no group differences were found in relation to 'non-VMS'; that condition is not further discussed here), while in the larger Italian study infants with a single VMS were significantly more interested in their own VMS (in the British study the same effect was only a trend): See Figure 1. Note that in both groups the Italian infants looked longer at both sets of stimuli: This presumably reflects the difference in presentation, with the individual VMS being more readily accessed in word lists (Italian) than in passages from which the VMS-rich words had to be segmented (British). In addition, the Italian study found in a separate experiment that in the pre-VMS period (at 6 mos.) there was no difference in attention to the different stimuli, which were distinguished only by their subsequent VMS status for the child and which were thus not expected to affect processing at an earlier developmental point.

[Insert Figure 1 about here.]

The findings of the three studies are in good accord. When a child first begins producing one consonant stably and consistently, as established by VMS identification, he or she is particularly attentive to that speech sound in input word forms (as shown in Majorano et al., 2014). When the child has advanced to production at VMS level of more than a single consonant the known (VMS) consonants no longer hold his or her attention; instead, the child seems to discover a world of varied stimuli and to begin to attend more to what is novel or unfamiliar (as seen for infants producing multiple VMS in both of these studies and also in DePaolis et al., 2013, in which the high-producers of [t/d] showed greater interest in the speech sound they had not yet mastered in production; for further discussion see Vihman et al., 2014 and [redacted], submitted). This developmental profile fits with Hunter and Ames' (1988) general model of child attention, which predicts a steady increase in initial focus on what is familiar, followed by satiation, which in turn leads to a period of greater focus on novel exemplars of a similar kind, a way of exploring what remains relatively unfamiliar or unknown once what is familiar has been sufficiently well internalized. The issue of experimentally identifying and interpreting attentional shifts between what is familiar and what is novel is complex, however, and remains controversial (see [redacted]., submitted). Nevertheless, the fact that the child's own level of production of speech sounds affects the way she processes or represents those sounds appears to be solidly established by this series of studies.

Word production, II: From holistic matches to reorganization and systematicity

The experimental evidence generally supports the pathway from vocal practice to first words that we have proposed based on observational findings (Vihman, 1993, 1996). First, babbling, which is maturationally timed but requires experience of input speech to be maintained (as deaf infants do not reliably babble in the first year: Oller & Eilers, 1988), lends salience to aspects of the input. In exemplar theory terms, similarity of form between heard word-form and existing child vocal pattern creates an 'echo' or resonance (Goldinger, 1996, 1998). As a result, secondly, frequently heard word forms come to be represented more robustly in the child's mind than forms for which the child lacks a possible vocal match. (For similar effects in L2 learners see Ellis & Sinclair, 1996, for example, as well as the studies of the 'production effect' referenced above.) Thirdly, that production-based salience in the speech stream facilitates formation of a form-meaning link in relevant and frequently repeated contexts, which can in turn result in early identifiable word production, under priming from a familiar situation of use. (Note that most early word use is primed by a highly familiar or routine situational context: Bates et al., 1979;

Vihman & McCune, 1994.)[3] This account, which sees in typical early word production individual, unrelated instances of 'item learning', would explain why the first words tend to be accurate, similar to the particular child's babbling repertoire and, in most cases, not phonologically related to one another in any systematic way (Menn & Vihman, 2011).

The idea that the early words are 'holistically' represented is also consistent with the finding that infant long-term memory for word forms may be insufficient to block word-form recognition when certain aspects are experimentally changed or 'mispronounced' in the period of first word use (i.e., in largely pre-linguistic 11-month-olds, but not in more verbally advanced 15- to 18-month-olds: Swingley, 2009a). Thus, final consonants of monosyllables in Dutch (Swingley, 2005) or onset consonants in unaccented syllables in French (Hallé & Boysson-Bardies, 1996) or English (Vihman et al., 2004) appear to be weakly represented in the pre-linguistic period; similarly, Vihman & Majorano (in press) demonstrate the perceptual neglect of word-initial consonants in Italian words with medial geminates (phonologically contrastive long consonants), though not in those with medial singletons. All of these instances of less-well represented elements of early word forms provide evidence that the form of the word as a whole affects infant processing. In fact, word production studies provide ample evidence of a difference between strongly represented sounds, such as the word-

[3] For example, the 'progressive idiom' shown for Kaia in Table 1, some months before any other words could be identified, was a repeatedly whispered response to the sight of one of the kittens born when she was 10 months old. This precocious verbal expression – surprisingly advanced in form – did not appear to be communicative in intent but rather served as a marker for self of a strongly experienced visual event (for discussion of the emergent use of communicative expression in this period see also Vihman & Miller, 1988; McCune et al., 1996).

initial consonants of trochaic (strong-weak) words in English, which are rarely omitted, and weakly represented sounds, such as the word-initial consonants of iambic (weak-strong) words in French (Vihman & Kunnari, 2006) or Hebrew (Keren-Portnoy & Segal, in press) or the onsets of trochaic words with geminates in Estonian, Finnish and Hindi as well as Italian (Vihman, 2016; Vihman & Croft, 2007; Vihman & Majorano, in press), all of which are commonly omitted in child word forms.

First word use leads to a small expressive lexicon. Once a few different words are being used with some regularity, children are typically found to generalize, or to begin to overuse, one or more of their production patterns, with two effects on their word forms taken as a whole: (i) they become more similar to one another and (ii) they become less accurate. In other words, we see regression in match to the adult model as a concomitant of the advance in systematicity evidenced by the increased similarity of the child's forms, which begin to fall into a small number of often used prosodic structures (or word structure in terms of length in syllables and of consonant and vowel [C-V] sequences). Such favored child word patterns are termed phonological templates, idiosyncratic child patterns found to apply both to 'selected' words, which exemplify the pattern, and 'adapted' words, more challenging adult word forms that are assimilated to it (see Table 2).

[Insert Table 2 about here.]

Table 2 illustrates the templates of the four children whose first words we saw in Table 1. For three of these children something of the later template can, with hindsight, be identified in their very first words: Annalena is extending a preference for reduplicated forms to more complex targets (Vihman & Croft, 2007); Alice is building on an affinity for producing words with palatal consonants and final [i] (Vihman et al., 1994); and Laurent has systematized his use of [l], now producing words in which it serves as the onset to the accented syllable, regardless of the actual structure of the word or phrase he is targeting (Vihman, 1993). In the case of Kaia, however, only one of the first words includes a medial geminate, which is the basis for her template at 16 months (Vihman, 2016).

The children's use of templates reflects their generalization of production patterns as their word learning advances. This can be conceptualized in at least two different ways: (i) as a purely procedural or motoric extension of existing production routines; (ii) as 'secondary' distributional learning, based on each child's individual database of early words. In either case the template is necessarily shaped by the ambient language target forms as well as by the child's individual production patterns. The choice of theoretical conceptualization is independent of the data themselves, which are robust: Evidence of template formation, at varying levels of lexical development and for varying periods of use, is available for a range of different languages – all those, in fact, for which individual cases of phonological development have received close linguistic analysis, although not all children provide evidence for such patterns (see the seven languages represented in Vihman & Keren-Portnoy, 2013, and the overview of 13 languages for which data are available in Vihman & Wauquier, in press).

Regardless of the mechanism, it is clear that these favoured routines or templates facilitate production – including articulation, planning and memory, or access to an emergent, still unstable representation. McGregor and Johnson (1997) put it succinctly:

> Template application allows the child to fit a production to a well-practiced routine, thereby reducing the demand on resources. Templates may aid the memory for the sound system as well as the planning and execution of motoric gestures. (p. 1220)

We can see templates as mediating between input- and output-based learning: First, over the first several months of life, the child becomes familiar with input speech, which comes to include his or her own adult-like vocalizations or output forms. This familiarization process itself can be understood as involving two processes occurring in parallel: Statistical or distributional learning, which operates as early as 6-9 months (Saffran, Aslin & Newport, 1996; Thiessen & Saffran, 2007), provides growing familiarity with the overall ambient language structure. Beginning at about the same time the very first words are recognized, with regards to both form and meaning (Tincoff & Jusczyk, 1999; Bergelson & Swingley, 2012). This emergent knowledge can be understood as corresponding to clouds of exemplars of similar forms for frequently heard words or short phrases (Johnson, 1997; Jusczyk, 1997; Pierrehumbert, 2001, 2003). Note that although babble is produced in strings of varying lengths, target-based word production is commonly limited to one or two syllables, regardless of ambient language structure (Vihman & Wauquier, in press). This limitation is most likely due to infants' untutored phonological memory, which will come to retain longer and more complex input elements as a concomitant of growth in lexical experience and use (Keren-Portnoy et al., 2010). Thus, while a general sense of the prosodic, phonotactic and coarticulatory regularities of the ambient language is gained by the end of the first year, as shown in segmentation studies (see Vihman, 2014, Ch. 5), lasting traces of individual lexical forms (exemplars) can be expected to accumulate more slowly and with strong representational constraints on word length and complexity.

Once the child begins to produce his or her own word forms with specific targets, the same implicit mechanisms (distributional learning of sequences and patterns, self-organization of exemplars) can be assumed to operate in combination on the *new database formed from the child's own words*. The increasing numbers of representations of forms the child is producing, albeit with a good deal of variability, will at some point generally become robust enough – in combination with the relatively slow pace of advances in neuromotor control and speech-planning – to give rise to one or more templates. As the child shifts from a primarily outward- to a primarily inward-oriented model for production we see the regression in accuracy described above along with an increase in the numbers of different word types produced. (Menn, 1971, among others, has reported an increase in the child's rate of word-learning at this point, although this has yet to be reliably demonstrated.)

This account points to the lexicon as the source of longer-term, robust phonological knowledge of individual segments (presumably this is the Phoneme plane of Werker and Curtin's PRIMIR model). The representations of production units, or units of form that have a link with meaning (words and short formulaic phrases, such as *all gone, what's this?*), can be expected to self-organize into networks based on similarity. Evidence of infant reliance on such networks can be seen in child lexical selection errors or 'mini-malapropisms', which tend to be based more often on holistic word-form similarities such as length in syllables and accentual pattern than on agreement in the initial sound (or letter), the most common basis for adult errors of this kind (Aitchison, 1972; Fay & Cutler, 1977; Vihman, 1981). The relations between sub-units in different words, whether word-initial consonants, rhymes, accentual patterns or other repeatedly represented elements, are subsequently analysed implicitly (for accounts of longitudinal change that suggest such a process of reorganization, see Priestly, 1977; Macken, 1979; Vihman & Vihman, 2011). As Edwards, Munson and Beckman (2011) put it, 'phonemes do not exist in nature, to be "discovered" by children. Rather, they emerge gradually as children make increasingly robust abstractions over the words that they learn' (38). In short, self-organization and implicit analysis mean systematization and integration into networks of phonological similarity. Those networks provide multiple access paths to shared 'positional variants' (Pierrehumbert, 2003) or phonemes, strengthening the representation of speech sounds with every instance of language use, whether receptive or expressive.

Learning mechanisms: The complementary systems model

How does the initial attunement to the native language described in the first part of this paper, the decline in attention to non-native contrasts based on passive exposure to speech, relate to the attention-based item learning that we have discussed and illustrated with children's first words? Infant knowledge of speech sounds based on distributional learning of the acoustic manifestations of input phonological structure cannot account for the production of identifiable word forms. In contrast, experience of word use *can* give rise to implicit knowledge, for the purposes of perceptual processing as well as production, of the phonological categories of the ambient language. Based on the complementary learning or memory systems model (McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Norman, 2002; Lindsay & Gaskell, 2010; McClelland et al., 2010; McClelland, 2013), knowledge of both words and sounds can be understood as being the byproduct of the integration, in active word learning and use, of implicit and explicit learning mechanisms*.* (See Ellis, 2005, for a similar account of L2 learning.)

Very few experimental studies have directly addressed memory functions in relation to infant word learning (but see now Friedrich, Wilhelm, Born & Friederici, 2015; Horváth et al., 2015). However, studies of word learning in adults and older children (e.g., Gais & Born, 2004; Dumay & Gaskell, 2007, 2012; Backhaus et al., 2008; Henderson, Weighall, Brown & Gaskell, 2013; Brown & Gaskell, 2014; Gaskell et al., 2014; Takashima et al., 2014; Henderson, Devine, Weighall & Gaskell, 2015) demonstrate the applicability to this domain of the principles of the complementary systems model (McClelland et al., 1995), which developed out of animal studies, neuroscience and computational modeling.

The memory system must be plastic enough to allow new learning, yet new learning must not be allowed to overwrite existing knowledge (the 'stability – plasticity dilemma'). The proposed solution is learning supported by two independent brain systems (McClelland et al., 1995; O'Reilly & Norman, 2002;

Kumaran & McClelland, 2012; McClelland, 2013): (i) The *neocortex* gives rise to 'incidental', implicit (including distributional or statistical) or procedural learning, with no need for focused attention; (ii) the *hippocampus and the prefrontal lobes* together support learning *with* attention (Wilhelm, Prehn-Kristensen & Born, 2012). Note, however, that neocortical activity is always present, whether focal attention is also engaged or not; this is one of the many difficulties involved in assessing the independent contribution of each of the two systems to subsequent access, (implicit) recognition and (explicit) recall (Jacoby, 1991).

Implicit (distributional or statistical and procedural) learning

The sensorimotor areas of the neocortex learn slowly from repeated experiences, gradually gaining automaticity in motor skills (procedural learning, such as balancing on a bicycle or producing a particular consonant at will), tallying statistical co-occurrences and, crucially, *categorizing the new in terms of what is already known*; only minimal attention, if any, is required for this incidental experiential learning. Implicit learning of any kind supports unconscious, involuntary recognition and a 'feeling of familiarity' when previously experienced items or events – or items or events that closely resemble what was previously experienced – are encountered anew (Jacoby, 1991). Access to such implicit or procedural memories is possible *only* with close contextual matching, however; it is not available to consciousness and cannot be called up at will.

Explicit or declarative learning

One function of the prefrontal lobes is to focus attention on aspects of experience and inhibit attention when it is no longer required, permitting the kind of flexible selection of points of focus that begins to appear in infants only from the second half of the first year (Ruff & Rothbart, 1996). In conjunction with focal attention – which strictly channels experience, permitting only a single focus – the hippocampus serves to bind the experienced event together with all of its unique spatiotemporal features; it is the key mechanism for retaining in memory the conjunction of separate (multimodal) aspects of experience. This notably includes the most essential characteristic of human language, the (typically arbitrary) link between a speech form and its situational context or meaning. These hippocampal snapshots of episodes experienced with attention underlie spontaneous (conscious, voluntary) recall; this is item learning.

However, the hippocampus supports rich but sparsely distributed neural codes, which are resistant to interference between similar experiences; the neocortex, in contrast, abstracts the structure underlying related experiences through its use of overlapping codes (Kumaran & McClelland, 2012). A key function of sleep appears to be the deeper processing of experiences, with active intercommunication between the hippocampus and the neocortex (e.g., Walker & Stickgold, 2004), in children as well as adults (Backhaus et al., 2008). In sleep, through neural reactivation of elements of experience, attention-based memory traces are restructured and consolidated (Lindsay & Gaskell, 2010), resulting in categorization into networks of sound and meaning. The process of selective strengthening of associations involved in this restructuring may be a key factor in the 'discovery of a shared structure' in representations (Drosopoulos, Schulze, Fischer & Born, 2007). In other words, the process of integrating new experiences with what is already known may constitute the critical basis for the generalization of knowledge that yields phonological categories and systems.

Integrating the findings

The characteristic profile for growth in word comprehension is a slow start (by 6-9 months: Bergelson & Swingley, 2012) followed by a rapidly rising curve, with a first inflection being observed only at about 14-18 months (Oviatt, 1980; Bergelson & Swingley, 2012, 2013), despite the fact that word-form recognition is reliably seen, cross-linguistically, by 11 months, as indicated above. The gap between initial word comprehension and word-form recognition and the more rapid, steadier advances in lexical learning that follow presumably reflects the benefit, for novel word learning, of a growing reference sample of familiar forms to which the novel items can be connected. Phonological memory, which develops through emergent use of word forms in production (Keren-Portnoy et al., 2010), can be taken to be a key element here. In addition, the sleep studies offer an account of how novel experiences are restructured through assimilation to existing patterns, which clarifies the importance of existing knowledge for consolidating new advances. This would support the idea that the onset of word production plays a pivotal role in integrating (i) emergent infant familiarity with the phonological categories of the ambient language with (ii) the infant's growing receptive lexicon of form-meaning pairs.

The first referential (symbolic) word use – which reflects the generalisation of semantic representations of arbitrary form-meaning relations across different particular instances – is also typically observed only from about 14 months on

(Bates et al., 1979; Vihman & McCune, 1994; McCune, 1992, 1995, 2008). And the phonological template use that we have illustrated here, involving generalization or schema formation rooted in a learned database, is observed from about the same age. (Template emergence, as indicated above, is closely tied to lexical growth, but not in any predictable or mechanistic way; individual differences prevail here, not universal rules or stages.)

In short, as infants begin to gain knowledge of a small number of often heard words (the first attention-based item-learning) and to register (implicitly) differences in the distribution of phonological categories in input speech they are laying the foundation for first word production, which additionally requires babbling practice. Production of some 50 to 100 different word types, in turn, prepares the ground for more rapid learning of new words along with the generalization of form patterns (as initially seen in templates) and meanings (as seen in symbolic word use). Given this conceptualization of knowledge and learning there is no real clash of sounds-before-words vs. words-before-sounds, as the learning of sounds and words necessarily proceeds in parallel.

References

Aitchison, J. (1972). Mini-malapropisms. *British Journal of Communication Disorders, 7,* 38-43.

Anderson, S. (1985). *Phonology in the Twentieth Century: Theories of rules and theories of representations.* Chicago: University of Chicago.

Backhaus, J., Hoeckesfeld, R., Born, J., Hohagen, F. & Junghanns, K. (2008). Immediate as well as delayed post learning sleep but not wakefulness enhances declarative memory consolidation in children. *Neurobiology of Learning and Memory, 89,* 7-80*.*

Bates, E., Benigni, L., Bretherton, I., Camaioni, L. & Volterra, V. (1979). *The Emergence of Symbols: Cognition and communication in infancy.* New York: Academic Press.

Beckman, M. E., & Edwards, J. (2000a). Lexical frequency effects on young children's imitative productions. In M. B. Broe & J. B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge: Cambridge University Press.

Beckman, M. E., & Edwards, J. (2000b). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development, 71*, 240-249.

Benedict, H. (1979). Early lexical development: Comprehension and Production. *Journal of Child Language*, *6*, 183-200

Bergelson, E. & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA,* 109, 3253-3258.

Bergelson, E. & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition, 127,* 391-397.

Best, C. T. (1994). The emergence of language-specific phonemic influences in infant speech percpetion. In J. C. Goodman & H. C. Nusbaum (eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge, MA: MIT Press.

Best, C. C. & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech, 46,* 183-216.

Blevins, Juliette (2004). *Evolutionary Phonology: The emergence of sound patterns.* Cambridge: Cambridge University Press.

Bloomfield, L. (1933). *Language.* New York: Holt.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science, 16*, 298-304.

Boysson-Bardies, B. de, Schonen, S. de, Jusczyk, P., MacNeilage, P. F. & Morton, J. (eds.). (1993). *Developmental Neurocognition: Speech and face processing in the first year of life* (411-419). Dordrecht: Kluwer.

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

Brent, M. R. & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33-B44.

Brown, H. & Gaskell, M. G. (2014). The time-course of talker-specificity and lexical competition effects during word learning. *Language, Cognition and Neuroscience, 29,* 1163-1179.

Bruderer, A. G., Danielson, D. K., Kandhadai, P. & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences, 112,* 13531-1356.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

Bybee, J. (2001). *Phonology and Language Use.* Cambridge: Cambridge University Press.

Bybee, J. (2010). *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J. & Witherington, D. (2000). Travel broadens the mind. *Infancy, 1,* 149-219.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*, Dordrecht, Holland: Foris.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Cristià, A., McGuire, G. L., Seidl, A. & Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *Journal of Phonetics, 39,* 388-402.

[redacted]. (submitted). An order effect in English infants' discrimination of an Urdu affricate contrast. *Journal of Phonetics.*

Davis, B. L. & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 38*, 1199-1211.

Davis, B. L. & MacNeilage, P. F. (2000). An embodiment perspective on the acquisition of speech perception. *Phonetica, 57,* 229-241.

[redacted] (submitted). Making sense of infant familiarity and novelty responses to words at lexical onset. *Frontiers in Psychology.*

DePaolis, R., Vihman, M. M. & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development, 34,* 590-601.

DePaolis, R., Vihman, M. M. & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development, 36*, 642-649.

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General, 136,* 169-183.

Dumay, N. & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18,* 35-39.

Dumay, N. & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition, 123,* 119-132.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47, 421-436.*

Edwards, J., Munson, B. & Beckman, M. E. (2011). Lexicon-phonology relationships and dynamics of early language development – a commentary on Stoel-Gammon's 'Relationships between lexical and phonological development in young children'. *Journal of Child Language, 38*, 35-40.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27,* 305-352.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49a,* 234–250.

Elsen, H. (1996). Two routes to language. *First Language, 16,* 141-158.

Fay, D. & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8,* 505-520.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419-439. Reprinted in Vihman & Keren-Portnoy (2013).

Ferguson, C. A., Peizer, D. B., & Weeks, T. A. (1973). Model-and-replica phonological grammar of a child's first words. *Lingua, 31*, 35-65.

Fernald, A. & Marchman, V. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83,* 203-222.

Foulkes, P. (2010). Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology, 1*, 5-39.

Foulkes, P. & Docherty, G. J. (2006). The social life of phonetics and phonology, *Journal of Phonetics, 34*,409-438.

Friedrich, M., Wilhelm, I., Born, J. & Friederici, A. (2015). Generalisation of word meanings during infant sleep. *Nature Communications, 6,* 6004 [DOI: 10.1038/ncomms7004].

Gais, S. & Born, J. (2004). Declarative memory consolidation: Mechanisms acting during human sleep. *Learning & Memory, 11,* 679-685.

Gaskell, M. G., Warker, J., Lindsay, S., Frost, R., Guest, J., Snowdon, R. & Stackhouse, A. (2014). Sleep underpins the plasticity of language production. *Psychological Science, 25*, 1457-1465.

German, D. J. & Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language and Hearing Research, 47,* 624-636.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251-279.

Goudbeek, M., Swingley, D. & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1913-1933.

Halle, M. (1971). *The Sound Pattern of Russian*: *A linguistic and acoustical investigation.* Berlin: Walter de Gruyter.

Hallé, P., & Boysson-Bardies, B. de. (1994). Emergence of an early lexicon: Infants' recognition of words. *Infant Behavior and Development, 17,* 119-129.

Hallé, P. & Boysson-Bardies, B. de. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development, 19,* 435-451*.*

Henderson, L. M., Weighall, A., Brown, H., & Gaskell, M. G. (2013). On-line lexical competition during spoken word recognition and word learning in children and adults. *Child Development, 84*, 1668–1685.

Henderson, L. M., Devine, K., Weighall, A. & Gaskell, G. (2015). When the daffodat flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories. *Developmental Psychology, 51,* 406-417.

Hickok, G. (2014) The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience, 29,* 2-20.

Horváth, K., Myers, K., Foster, R. & Plunkett, K. (2015). Napping facilitates word learning in early lexical development. *Journal of Sleep Research*. DOI: 10.1111/jsr.12306

Houston, D. M. & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1570-1582.

Houston, D. M. & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1143-1154.

Hunter, M. A. & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt, eds., *Advances in Infancy Research*, vol. 5, pp. 69-95.

Huttenlocher, J. (1974). The origins of language comprehension. In R. L. Solso (ed.), *Theories in Cognitive Psychology*. Hillsdale, NJ: Lawrence Erlbaum.

Icht, M. & Mama, Y. (2015). The production effect in memory: A prominent mnemonic in children. *Journal of Child Language, 42,* 1102-1124.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

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

Johnson, E. & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44,* 548-567.

Johnson, E. K. & Tyler, M. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science, 13*, 339-345.

Johnson, K. A. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullenix (eds.), *Talker Variability in Speech Processing* (pp. 145-165). San Diego: Academic Press.

Johnson, K. A. (2007). Decisions and mechanisms in exemplar-based phonology. In M-J. Solé, P. S. Beddor & M. Ohala (eds.), *Experimental Approaches to Phonology* (pp. 25-40). Oxford: Oxford University Press.

Jusczyk, P. W. (1986). Toward a model of the development of speech perception. In J. S. Perkell & D. H. Klatt (eds.), *Invariance and Variability in Speech Processes* (pp. 1-19). Hillsdale, NJ: Erlbaum.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language.* Cambridge, MA: MIT Press.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development, 18,* 111-116.

Keren-Portnoy, T. & Segal, O. (in press). Phonological development in Hebrew-learning infants and toddlers: perception and production. To appear in R.

Berman (ed.), *Acquisition of Hebrew from Infancy to Adolescence.* TiLAR. Amsterdam: John Benjamins.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R., Whitaker, C. & Williams, N. A. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language, and Hearing Research*, *53*, 1280-1293.

[redacted]. (submitted). Do infants learn from isolated words? An ecological study.

Kleinschmidt, D. F. & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122,* 148-203.

Kuhl, P. K. (1986). Reflections on infants' perception and representation of speech. In J.S. Perkell & D. H. Klatt, *Invariance and Variability in Speech Processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831-843.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9,* F13-F21.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). *Philosophical Transactions of the Royal Society, B: Biological Sciences, 363,* 979-1000.

Kumaran, D. & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review, 119,* 573-616.

Labov, W. (2001). *Principles of Linguistic Change. Volume 2: Social Factors*. Oxford: Blackwell.

Lewkowicz, D. & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13,* 470-478.

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

Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning, 60,* S2, 45-63.

Macken, M. A. (1980). Aspects of the acquisition of stop systems: A cross-linguistic perspective. In G. Yeni-komshian, J. F. Kavanagh & C. A. Ferguson (eds.), Child Phonology, I: Production (pp. 143-168). New York: Academic Press.

MacLeod, C., Gopie, N., Hourihan, K. L., Neary, K. R. & Ozubko, K. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 671-685.

Majorano, M., Vihman, M. M. & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development, 10*, 179-204.

Maurer, D. & Werker, J. F. (2013). Perceptual narrowing during infancy: A comparison of language and faces. *Development Psychobiology, 56,* 154-178.

Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101–B111.

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal* *of Experimental Psychology: General, 142*, 1190-1210. doi: 10.1037/a0033812.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C. Rogers, T. T., Seidenberg, M. S. & Smith, L. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences, 14,* 348-356.

McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419-517.

McCune, L. (1992). First words: A dynamic systems view. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press.

McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology, 31,* 198-206.

McCune, L. (2008). *How Children Learn to Learn Language: Productive skills in a dynamic system.* Oxford: Oxford University Press.

McCune, L. & Vihman, M. M. (2001). Early phonetic and lexical development. *Journal of Speech, Language and Hearing Research, 44,* 670-684.

McCune, L., Vihman, M. M., Roug-Hellichius, L., Delery, D. B. & Gogate, L. (1996). Grunt communication in human infants (*Homo sapiens*). *Journal of Comparative Psychology, 110,* 27-37.

McGregor, K. & Johnson, A. (1997). Trochaic template use in early words and phrases. *Journal of Speech, Hearing and Language Research, 40,* 1220-1231.

McGillion, M. M., Matthews, D., Herbert, J., Pine, J., Vihman, M. M., Keren-Portnoy, T. & DePaolis, R. A. (in press). What paves the way to conventional language? The predictive value of babble, pointing and SES. *Child Development.*

McMurray, B., Tanenhaus, M., & Aslin, R. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86*, B33–B42.

Menn, L. (1971). Phonotactic rules in beginning speech: A study in the development of English discourse. *Lingua, 26,* 225-251.

Menn, L., Schmidt, E. & Nicholas, B. (2013). Challenges to theories, charges to a model: The Linked-Attractor model of phonological development. In M. M. Vihman & T. Keren-Portnoy (eds.), *The Emergence of Phonology: Whole word approaches, cross-linguistic evidence.* Cambridge: Cambridge University Press.

Menn, L. & Vihman, M. M. (2011). Features in child phonology: inherent, emergent, or artefacts of analysis? In N. Clements & R. Ridouane (eds.), *Where Do Phonological Features Come From? The nature and sources of phonological primitives* (pp. 261-301). Amsterdam: John Benjamins.

Miller, J. L. (1994). On the internal structure of phonetic categories: A progress report. *Cognition, 50,* 271-285.

Munson, B., Edwards, J. & Beckman, M. E. (2012). Phonological representation in language acquisition: Climbing the ladder of abstraction. In A. C. Cohn, C. Fougeron & M. K. Huffman (eds.), *Oxford Handbook of Laboratory Phonology.* Oxford: Oxford University Press.

Oller, D. K. (2000). *The Emergence of the Speech Capacity.* Mahwah, NJ: Lawrence Erlbaum.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59,* 441-449.

O'Reilly, R. C. & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advancesin the complementary learning systems framework. *Trends in Coognitive Sciences, 6,* 505-510.

Oviatt, S. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, *50*, 97-106.

Pascalis, O., De Haan, M. & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science, 296,* 1321-1323.

Pegg, J. E. & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America, 102,* 3742-3753.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (eds.), *Frequency Effects and the Emergence of Linguistic Structure.* Amsterdam: John Benjamins.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46,* 115-154.

Plummer, A. R. & Beckman, M. E. (in press). Framing a socio-indexical basis for the emergence and cultural transmission of phonological systems. *Journal of Phonetics.*

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology, 25*, 143–170.

Redford, M. A. (in press). Unifying speech and language in a developmentally sensitive model of production. *Journal of Phonetics.*

Ruff, H. A. & Rothbart, M. K. (1996). *Attention in Early Development*. Oxford: Oxford University Press.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Saussure, F. de (1959). *Course in General Linguistics.* W. Baskin, tr. New York: Philosophical Library. First edition, 1915. C. Bally & A. Sechehaye, eds., in collaboration with A. Reidlinger.

Scott, L. S., Pascalis, O. & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16,* 197-201.

Segal, O., Hejli-Assi, S. & Kishon-Rabin, L. (2016). The effect of listening experience on the discrimination of /ba/ and /pa/ in Hebrew-learning and Arabic-learning infants. *Infant Behavior and Development, 42,* 86-99.

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E.E. David, Jr. & P.B. Denes (eds.), *Human Communication: A Unified View.* New York: McGraw-Hill.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3-45.

Stevens, K. N. (1998). *Acoustic Phonetics.* Cambridge, MA: MIT Press.

Storkel, H. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321-1337.

Swingley, D. (2005). Eleven-month-olds' knowledge of how familiar words sound. *Developmental Science, 8,* 432-443.

Swingley, D. (2009a). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language, 60,* 252-269.

Swingley, D. (2009b). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society*, *B, 364*, 3617–3632.

Takashima, A., Bakker, I., van Hell, J. G., Janzen, G. & McQueen, J. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *Neuroimage, 84,* 265-278.

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

Thiessen, E. D. & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- and 9-month-old infants. *Developmental Psychology, 39,* 706-716.

Thiessen E.D., & Saffran, J.R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning & Development, 3,* 73-100.

Tincoff, R. & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science, 10,* 172–175.

Tipper, S. (1992). Selection for action: The role of inhibitory mechanisms. *Current Directions in Psychological Science, 1,* 105-109.

Torkildsen, J. V. K., Svangstu, J. M., Hansen, H. F., Smith, L., Simonsen, H. G., Moen, I. & Lindgren, M. (2008). Productive vocabulary size predicts Event-Related Potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience, 20,* 1266-1282.

Torkildsen, J. V. K., Hansen, H. F., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I. & Lindgren, M. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language, 108,* 73-88.

Twaddell, W. F. (1935). On deifning the phoneme. *Language Monograph,* no. 16. Reprinted in Joos, 1957, 55-80.

Vihman, M. M. (1981). Phonology and the development of the lexicon: evidence from children's errors. *Journal of Child Language*, *8*, 239-264.

Vihman, M. M. (1991). Ontogeny of phonetic gestures: Speech production. In I.G. Mattingly & M. Studdert-Kennedy (eds.), *Modularity and the Motor Theory of Speech Perception*. NY: Lawrence Erlbaum Associates.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*. Parkton, MD: York Press.

Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, *21*, 61-82.

Vihman, M. M. (1996). *Phonological Development: The origins of language in the child*. Oxford: Basil Blackwell.

Vihman, M. M. (2014). *Phonological Development: The first two years*. (2nd ed.) Malden, MA: Wiley-Blackwell.

Vihman, M. M. (2016). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition, 19,* 69-88*.*

Vihman, M. M. & Croft, W. (2007). Phonological development: Toward a 'radical' templatic phonology. *Linguistics, 45,* 683-725. Reprinted in. Vihman & Keren-Portnoy (2013).

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2014). The role of production in infant word learning. *Language Learning, 64: Suppl. 2, 121-140.*

Vihman, M. M. & Keren-Portnoy, T. (eds.) (2013). *The Emergence of Phonology: Whole word approaches, cross-linguistic evidence.* Cambridge: Cambridge University Press.

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

Vihman, M. M., Macken, M.A., Miller, R., Simmons, H. & Miller, J. (1985). From babbling to speech: a reassessment of the continuity issue. *Language*, *61*, 397-445.

Vihman, M. M. & Majorano, M. (in press). The role of geminates in infants' early words and word-form recognition. *Journal of Child Language.*

Vihman, M. M. & McCune L. (1994). When is a word a word*? Journal of Child Language, 21,* 517-542.

Vihman, M. M. & Miller, R. (1988). Words and babble at the threshold of language acquisition. In M. D. Smith & J. L. Locke (eds.), *The Emergent Lexicon,* pp. 151-183. N.Y.: Academic Press.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50,* 336-353.

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

Vihman, M.M. & Wauquier, S. (in press). Templates in child language. In M. Hickmann, E. Veneziano & H. Jisa (eds.), *Sources of variation in first language acquisition: Languages, contexts, and learners.* Trends in language acquisition research (TiLAR). Amsterdam: John Benjamins.

Walker, M. P. & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron, 44,* 121-133.

Wedel, A. B. (2007). Feedback and regularity in the lexicon. *Phonology, 24*, 147-185.

Werker, J. F. & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language, Learning and Development, 1,* 197-234.

Werker, J. F. & Fennell, C. T. (2004). Listening to sounds versus listening to words: Early steps in word learning. In D. G. Hall & S. Waxman (eds.), *Weaving a Lexicon* (pp. 79-109). Cambridge, MA: MIT Press.

Weinreich, U., Labov, W. & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics: A Symposium* (pp. 95-195). Austin: University of Texas Press.

Werker, J. F. & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37,* 278-296.

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

Wilhelm, I., Prehn-Kristensen, A. & Born, J. (2012a). Sleep-dependent memory consolidation - What can be learnt from children? *Neuroscience and Behavioral Reviews, 36,* 1718-1728.

Zamuner, T. S., Morin-Lessard, E., Strahm, S. & Page, M. P. A. (in press). Spoken word recognition of novel words, either produced or only heard during learning. *Journal of Memory and Language.*

Table 1. First words in four languages (based on observational research studies [English, French] or diary studies [Estonian, German])

GERMAN: Annalena, 8-10 mos. (Elsen, 1996)	
das da ‘that one there’ /das da/	[data]
Mama 'mama' /mama/	[mama]
Papa 'papa' /papa/	[baba]
pieppiep ‘peeppeep’ /pipip/	[pɪpɪ]
Teddy /tɛdi/	[dɛdɛ]
ENGLISH: Alice, 9-10 mos. (Vihman, Velleman & McCune, 1994)	
baby	[pɛpɛ:], [tɛɪti:]
daddy	[dæ]
hi	[hɑ:i:], [ʔɑ:jɛ], [hɑɪje] [haɪjʌ]...
mommy	[m̩:ɑn:ə]
no	[njæ̃]
FRENCH: Laurent, 10 mos. (Vihman, 1993)	
allo ‘hello’ [alo]	[hailo], [ailo], [haljo], [aljo], [alo]
donne (le) ‘give (it)’ [dʌnlø]	[dlə], [də], [ldɛ], [heldɔ]
l'eau-l'eau ‘bottle (nursery word)’ [lolo]	[ljoljo]
non [nɔ̃] ‘no’	[ne]
tiens [tjɛ̃] ‘here, take it’	[ta]
ESTONIAN-ENGLISH: Kaia, 11-15 mos. (Vihman, 2016)	
anna ‘give’ /an:a/	[an:an:a]
head'aega ‘byebye’ /heat'aeka/	[dada]
kiisu ‘kitty’ /ki:su/	[ki̥:tɔ̥]

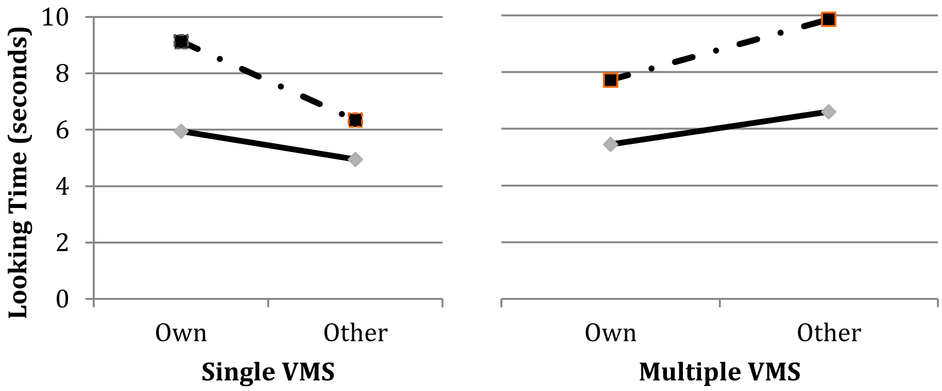
mõmmi ‘teddybear’ /mɤm:i/	[mʌm]
naba ‘belly button’ /napa/	[baba]
nämma ‘yum’ /næm:a/	[mæm:]
* ‘progressive idiom’: A word form well in advance of the child’s current production (Ferguson & Farwell, 1975).	

Table 2. Phonological templates in later words in four languages. ‘Selected’ words are close to the target, suggesting possible sources for the template; ‘adapted’ words show changes to the target that assimilate it to the template. < > = schematic template form; C = consonant, V = vowel, C_0 = optional consonant slot

select		*adapt*	
target word	*child form*	*target word*	*child form*
GERMAN*:* Annalena, 10-12 mos. < σ_1 σ_1 >, i.e., reduplicated syllables (Elsen,			
Pipi /'pipi/ 'peepee'	[pipi:]	*Annalena* /'analena/	[nana]
wauwau /'vauvau/ 'bowwow'	[vava]	*kikeriki* /kikeri'ki::/ 'cock-a-doodle-do'	[ki:ki:]
		Bauch /baux/ 'belly'	[baba]
		Tag /tak/ '(good)day'	[dada]
		Zahn(bürste) /'tsa:nbürʃtə/ 'tooth(brush)'	[nana]
ENGLISH*:* Alice, 14 mos. <CVCi> (Vihman, Velleman & McCune, 1994)			
baby	[bebi]	*bottle*	[baʤi, ba:tʃi, batji]
daddy	[tæɟi]	*hiya*	[ha:ji]
lady	[jɛiji]		
mommy	[maɲi]		
FRENCH*:* Laurent, 15 mos. < C_0VlV > (Vihman & Kunnari, 2006)			
allo ‘hello’ /alo/	[alo]	*canard* ‘duck’ /kanaʁ/	[kɔla]
dans l'eau, de l’eau 'in/some water’ /dɑ̃lo/, /dəlo/	[dəlo]	*chapeau* ‘hat’/ʃapo/	[bɔlo]

ballon 'big ball' /balɔ̃/	[palɔ]	*la brosse* 'the brush' /labrʌs/	[bəla]
pas là 'not there' /pala/	[pala]	*la cuillère* 'the spoon' /lakɥijɛʁ/	[kola]
		voilà 'there you are' /vwala/	[lala]
ESTONIAN: Kaia, 16 mos. < aC:V> (Vihman, 2016)			
anna /'an:na/ 'give'	[an:a]	*õue* /'ɤʊ:we/ 'to outside'	[au:a]
		juua /'ju:wa/ 'to drink'	[au:a]
		auto /'au:to/ 'car'	[at:o]
		lutti /'lut:ti/ 'pacifier'	[at:i]
		lahti /'lah:ti/ 'open, unstuck'	[at:i]

Figure 1. Infants with single VMS compared with infants with two or more VMS in their response to a passage (solid line, British study, DePaolis et al., 2011) or a list of isolated words (dashed line, Italian study, Majorano et al., 2014), each featuring a particular VMS.


Looking Time (seconds)
10
8
6
4
2
0
Own
Other
Single VMS
Own
Other
Multiple VMS