Published in final edited form as:

J Speech Lang Hear Res. 2012 December ; 55(6): 1626–1639. doi:10.1044/1092-4388(2012/11-0250).

Identification of Prelinguistic Phonological Categories

Heather L. Ramsdell^a, D. Kimbrough Oller^b, Eugene H. Buder^b, Corinna A. Ethington^b, and Lesya Chorna^b

^aEast Carolina University, Greenville, NC

^bUniversity of Memphis, TN

Abstract

Purpose—The prelinguistic infant’s babbling repertoire of *syllables*—the phonological categories that form the basis for early word learning—is noticed by caregivers who interact with infants around them. Prior research on babbling has not explored the caregiver’s role in recognition of early vocal categories as foundations for word learning. In the present work, the authors begin to address this gap.

Method—The authors explored vocalizations produced by 8 infants at 3 ages (8, 10, and 12 months) in studies illustrating identification of phonological categories through caregiver report, laboratory procedures simulating the caregiver’s natural mode of listening, and the more traditional laboratory approach (phonetic transcription).

Results—Caregivers reported small repertoires of syllables for their infants. Repertoires of similar size and phonetic content were discerned in the laboratory by judges who simulated the caregiver’s natural mode of listening. However, phonetic transcription with repeated listening to infant recordings yielded repertoire sizes that vastly exceeded those reported by caregivers and naturalistic listeners.

Conclusions—The results suggest that caregiver report and naturalistic listening by laboratory staff can provide a new way to explore key characteristics of early infant vocal categories, a way that may provide insight into later speech and language development.

Keywords

infant vocal development; phonological templates; caregiver–infant interaction; prelinguistic development

Prior to production of words, infants develop *syllables*—potential components of words (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, Bernstein, & Demorest, 1993). When caregivers identify these babbled syllables, they intuitively begin to interact with the infant around them, treating them as potential words (Papoušek, 1994; Stoel-Gammon, 2011; Veneziano, 1988). As word learning begins, caregivers engage infants, recognizing babbled syllables and their potential relation with the ambient language.

We propose that the most important information to determine a prelinguistic infant’s repertoire of phonological categories (especially consonant–vowel [CV] syllables) should be derived from a caregiver report or from reports of individuals listening the way caregivers

do. This style of listening is very different from the laboratory style, where phonetic transcription is the traditional approach to acquire data to determine syllabic repertoires of canonical babbling (including well-formed syllables; e.g., [ba] or [na]) and early words (see, e.g., Locke, 1983; Stoel-Gammon, 1992; Vihman, 1986). Acoustic analysis is often also used in babbling research—but not to determine functional syllabic repertoires in infancy (e.g., Buder & Stoel-Gammon, 2002; Papaeliou, Minadakis, & Cavouras, 2002). Our focus is on *phonetic transcription*, the standard approach for acquiring data to determine canonical babbling repertoires.

When transcribing, trained listeners play infant sounds repeatedly and generate International Phonetics Association (IPA; 2010) descriptions that are often highly detailed (Ball, Müller, Klopfenstein, & Rutter, 2009; Pollock & Berni, 2001; Shriberg & Lof, 1991)—more detailed than seems reasonable. To date, restrictions on the level of detail considered in determining infant repertoires through transcription have been based primarily on practical criteria such as interobserver agreement, which is known to decrease with increasing transcription detail (Coussé, Gillis, Kloots, & Swerts, 2004; Cucchiari, 1996; Munson, Edwards, Schellinger, Beckman, & Meyer, 2010; Nathani & Oller, 2001; Oller & Ramsdell, 2006; Preston, Ramsdell, Oller, Edwards, & Tobin, 2011; Ramsdell & Oller, 2007; Shriberg & Lof, 1991; Stockman, Woods, & Tishman, 1981) rather than on a principled method aimed at determining the infant's *functional* repertoire of phonological categories. Caregiver judgment can help provide a principled method. We share the reasoning of Papoušek (1994) and Veneziano (1988), suggesting that the functional repertoire of infant syllables is best seen as that repertoire recognized by caregivers, around which interaction about the semantic meaning of words can begin. Accordingly, we propose that data should be drawn from caregivers to determine infants' functional prelinguistic phonological categories, or at least from information derived through a style of listening similar to that of caregivers.

Caregiver reports suggest their recall of infant phonological categories includes far less detail than found with transcription (Alford, 2006; Ramsdell, Oller, & Buder, 2007). We presume that memory and attention constraints focus caregivers on global patterns produced repetitively by infants (Kwon, Oller, & Buder, 2007), a presumption consistent with models of distributional learning (e.g., Goudbeek & Swingle, 2006; Gupta & Cohen, 2002) and memory (Cowan, 1995). Thus, caregivers may ignore phonetic detail and recognize only small repertoires of canonical syllables in babbling. Phonetic transcription, in contrast, can yield the impression that infants possess very large repertoires, consistent with the outmoded idea that infants produce all of the sounds of all the world's languages (Jakobson, 1941).

Modern researchers are clearly aware that infant functional repertoires are smaller than those portrayed directly in phonetic transcription. To limit overestimation, infant repertoires have sometimes been restricted to include only sounds occurring some minimal number of times in a transcribed sample (Rvachew, Creighton, Sauve, & Feldman, 2005; Stoel-Gammon, 1988), at some minimal proportion in a sample (Vihman, 1992), or abiding by presumed phonetic principles (MacNeilage, 2008). Still, the basis for determining repertoires has been derived from phonetic transcription rather than from caregiver report or from any attempt to simulate the caregiver's method of listening. Yet, the caregiver's listening style is part and parcel of the normal circumstance in which the infant's productive word learning occurs.

We proposed, then, a principled basis on which to determine the prelinguistic syllabic repertoires of infants, a basis invoking a more natural approach—termed here *naturalistic listening*. We targeted the interaction between phonological (syllabic) categories systematically produced by infants and recognition of those categories by caregivers, a recognition founded on caregivers' natural listening. We proposed to include in babbling research an explicit account of infant functional syllabic repertoires as recognized by

caregivers and/or by individuals simulating caregiver listening. Prior research has not included this systematic basis for assessing babbling repertoires, which, we argue, forms an essential foundation for word learning.

Much research suggests that caregiver assessments of infant speech capabilities may be as accurate as, or more accurate than, laboratory measures. For example, productive word repertoires up to 30 months (Dale, Dionne, Eley, & Plomin, 2000; Fenson et al., 1991) as well as onset of canonical babbling (Oller, Eilers, & Basinger, 2001) can be reliably assessed through parent report. Other research supports parent report of infant vocalizations as a predictor of speech development (Cameron, Livson, & Bayley, 1967; Lyytinen, Poikkeus, Leiwo, & Ahonen, 1996) and the use of parent diaries in developmental research (Adolph, Robinson, Young, & Gill-Alvarez, 2008; Naigles, Hoff, & Vear, 2009). In a similar way, caregivers may be informative about functional phonological categories in babbling.

Templates in Prelinguistic Phonological Categories

A foundation for our approach is the notion that variable utterances of each emergent word (or syllable) can be characterized by an underspecified phonological “template.” The idea, with deep roots in phonological theory (Firth, 1948; Keating, 1988; Kiparsky, 1985) and child phonology (Ferguson & Farwell, 1975; Waterson, 1987), has been elaborated recently by Vihman and colleagues (Velleman & Vihman, 2002; Vihman, 1992; Vihman & Croft, 2007). It suggests simple schemes to portray global categories. For example, the template *labial obstruent plus vowel* encompasses syllables containing fricatives, affricates, plosives, differing degrees of voicing, and a vowel of any height or frontness, with or without nasalization or lip rounding. A recorded sound perceived as “bababa,” for example, may be shown to include much syllabic variation (e.g., [baβɛbβə]) if played back repeatedly for transcription (Oller & Griebel, 2008). Early word production tends to vary similarly (Vihman, DePaolis, & Keren-Portnoy, 2009), and such variation has been observed by direct monitoring of child articulation (Smith & Zelaznik, 2004; Walsh, Smith, & Weber-Fox, 2006). In a similar way, we propose that templates may provide a convenient formal description for infant prelinguistic phonological categories, leaving many details indefinite and presumably preventing overestimation of the size of infants’ functional syllable repertoires, thus, avoiding attribution to infants of the ability to produce syllables with greater precision and consistency than is observed.

The template model has been successfully applied to early word learning, showing how “vocal-motor schemes” (templates) are related to phonetically similar words in adult language (Keren-Portnoy, Majorano, & Vihman, 2009; Vihman & Croft, 2007; Vihman et al., 2009). Such templates describe children’s early word forms. We argue that templates can be derived as foundations for first words even before first words occur, by referencing caregiver recognition of syllables in babbling. Just as children’s early words can be seen as consistent phonetic templates, infants’ *prelinguistic phonological categories* in babbling can be recognized by caregivers and interpreted as potential words.

Raw Phonological Material and Negotiable Phonological Product

Our focus hinges on the existence of at least two layers of perception of early infant syllables. We term the first layer *raw phonological material* (RM), which includes the many variable pronunciations that can be discerned with careful listening or acoustic analysis of infant vocalizations (see, e.g., Munson et al., 2010). We call the second layer *negotiable phonological product* (NP), which consists of the much smaller number of functional syllabic categories (templates) produced by infants, each including a large number of RM variants (Oller & Griebel, 2008; Ramsdell et al., 2007). Even though a phonemic–phonetic

distinction is perhaps the most fundamental distinction in linguistic phonology (Bloomfield, 1933; Hockett, 1942), researchers have generally avoided such a distinction in infant babbling. *Phonemes* are defined as the contrastive units from which words are composed, and each phoneme occurs with great phonetic variation within languages. Babbling does not include words, so by definition it cannot include phonemes. Researchers generally speak of phonetic elements (or *phones*) in babbling but do not assume the existence of a phoneme-like layer of description (see, e.g., Rvachew et al., 2005). We propose the NP–RM distinction as a precursor to the phonemic–phonetic distinction.

Babbling infants do not yet produce words, but in canonical babbling they do produce syllable categories with sufficient contrastivity that they can be recognized by caregivers as distinct units (NP) in spite of within-category variation (RM). These syllabic units, thus, have the potential to function as words or components of words. Each NP contrastive category (a prelinguistic template) can be viewed as a precursor to a phoneme sequence (a syllable), and its RM variants can be viewed as precursors to phonetic variants within that phoneme sequence.

Background on Negotiation in Infant Babbling and Word Learning

The NP/RM distinction hinges on the intended meaning of negotiation in babbling and early word learning. Our discussion of negotiation focuses on vocal categories infants produce. The emphasis is important because infants can discriminate many speech contrasts they do not produce (Jusczyk, 1997). However, it is production that limits the possible field of negotiation over meaning of child utterances, and, thus, our focus is on production. Both infant and caregiver may recognize aspects of the infant's vocalizations that the other ignores, but we reason that negotiation can only occur when the caregiver and infant interact over a phonological category produced systematically by the infant and recognized by the caregiver. Only then can they come to act as if they have agreed on meanings for those categories (Stoel-Gammon, 2011). The notion of negotiation in determining communicative function and meaning in early childhood and animal communication has been widely recognized (Fogel & Garvey, 2007; Hinde, 1985; Labov & Labov, 1978). Families often permanently adopt novel words produced by such negotiation, for example, “soda” was [toudi]; “thank you” was [taŋko]; “who’s that” was [hu də]; and “pacifier” was [pætsi], in the first author’s family, all adopted from infant forms. A recording from one of our longitudinal studies demonstrates negotiation: An infant produced a sequence of variable [ba]-like syllables, and the caregiver asked, “Are you saying ‘boot’?” Later, during which time the infant continued babbling [ba]-like syllables, the caregiver revised the interpretation, “Oh, I know, you mean ‘bubble’.”

It is important to recognize that negotiation occurs in explicit word learning and in pure babbling. Veneziano (1988) emphasized that in early “child-mother interaction,” the parties imitate each other, and have “joint activity” around the “vocal-verbal material itself,” an interaction focused on “dyadically shared sonorics elements” (p. 111). This view highlights negotiation with respect to the sounds of babbling before words and is supported in research reviewed by Stoel-Gammon (2011). The point is also highlighted in experiments illustrating that infants respond contingently to caregiver vocalization long before they talk (Bloom, Russell, & Wassenberg, 1987; Goldstein, King, & West, 2003) and tend to produce more canonical utterances with elicitation of canonical utterances by caregivers (Bloom, 1988; Goldstein & Schwade, 2008; Gros-Louis, West, Goldstein, & King, 2006).

Goals and Rationale

Naturalistic listening, such as that exemplified in caregivers, has not, to date, been used in studying pre-linguistic phonological category (NP) development. Yet negotiable syllabic categories form a primary basis for vocal interaction between caregivers and infants and, thus, we argue, provide foundations for word learning. In a similar way, in the past there has been no principled basis on which to differentiate RM from NP in babbling. Our longitudinal study brings the caregiver explicitly into the picture of identifying NP categories and attempts to relate caregiver reports to judgments obtained in new laboratory listening procedures simulating the caregiver's more natural style of listening. Laboratory procedures are expected to reflect NP or RM, depending on manipulation of procedures. For comparison we also assess infant RM through phonetic transcription. By manipulating the number of listening opportunities in transcription we hope to offer new perspectives on the level of RM detail observed through transcription. We review repertoire sizes determined through caregiver report, naturalistic listening by laboratory staff, and phonetic transcription. In addition, a set of prelinguistic phonological category templates is derived from aggregated babbling data across infants and ages based on caregiver report. Results based on infant syllables recognized through naturalistic listening and phonetic transcription are categorized in terms of the same templates, and the frequency of occurrence of syllables in laboratory staff report is correlated with the frequency of occurrence of syllables in caregiver report. This approach provides an initial perspective on the extent of match with respect to both the size and content of early phonological repertoires in caregiver report and laboratory methods (both naturalistic listening and transcribing). We presume that larger RM than NP inventories are inevitable. Our focus is on the magnitude of the differences between NP–RM derived inventories and on possible similarities in phonetic makeup of perceived templates across the different methods of assessing infant syllabic repertoire.

Method

We present three studies and a cross-study comparison. In Study 1, we report infant syllable repertoires identified by caregivers; in Study 2, by naturalistic listening; and in Study 3, by phonetic transcription. A cross-study summary compares repertoire sizes across studies, and correlation results show the degree to which measures from Studies 2 and 3 conform to the phonetic content of templates derived from the caregiver report.

Infants, Recordings, and Caregiver Interviews

Recordings for all three studies came from a 30-month longitudinal study of vocal development with eight infants, focusing here only on data from three ages: 8, 10, and 12 months (24 total recording sessions). All parents reported normal birth at term with no significant history of prenatal or perinatal problems. Six of the infants were female; one was African American, and one was White Hispanic; one was from a home where German, Spanish, and English were spoken; and one was from a home where Ukrainian and English were spoken. Potential middle-ear problems were queried during recording sessions (they were rare), and referrals were given as appropriate. All families were of middle socioeconomic status. Follow-up to 24 months with the infants and interviews with parents revealed no significant delays in onset of speech. One of the infants (Subject D, White, English learning) had been unexpectedly diagnosed with Asperger's syndrome at age 6. He is highly verbal and was not noticed as vocally atypical during the longitudinal research. A more detailed study of his data and current status, with cooperation of the parents, is underway.

The recording playroom resembled a nursery, connected by a one-way observation mirror to a control room. An infant vest housed a high-fidelity wireless microphone to control mouth-

to-microphone distance (Buder & Stoel-Gammon, 2002). Sixteen-bit quantization allowed a signal-to-noise ratio of up to 96 dB, with signals digitized at sampling rates of 44.1 or 48 kHz. Recordings were done in 20-min segments. An infant visit usually yielded two or three such recordings. Parent interviews were conducted during these recordings.

CV Syllable Program

In all three studies, possible syllables were drawn from CV combinations of 46 Cs and 16 Vs, for a total of 736 possible syllable types (see Table 1 of Online Supplementary Material). The Cs and Vs included all commonly occurring allophones of English plus additional IPA elements. For Studies 2 and 3, a program was developed in Logical International Phonetics Programs (LIPP; Oller & Delgado, 1999) to automatically tabulate all possible CV combinations, with principles taken in part from prior studies (Oller & Ramsdell, 2006; Preston et al., 2011; Ramsdell & Oller, 2007).

Study 1 Method: Infant Syllable Repertoires Determined by Caregiver Report

—To determine syllabic categories in infant inventories according to caregiver report, audio- and video-recorded questionnaires administered by laboratory staff during recordings were reviewed. The key question posed to caregivers at every recording starting at 3 months of age was, “What sounds has your infant produced since your last visit?” By the time infants had begun canonical babbling, caregivers consistently responded to this question with a list of sounds and syllables (e.g., squeals, *ah*, *ba*, *na*, etc.). At each age, the list of caregiver-reported CV syllables was taken to represent the negotiable CV categories. The primary purpose of Study 1 was to establish estimates of the number of syllables caregivers attribute to their infants. The design also offered the opportunity to assess a possible age effect. In the age range studied, infants had not developed more than a very few words, and consequently, it seemed likely that no reliable effect of age on syllable repertoire size would be seen. A repeated-measures analysis of variance (ANOVA) evaluated the dependent variable (number of different syllable types indicated by the caregiver) and the independent variable (infant age: 8, 10, and 12 months). Data from the caregiver report were not available for one infant at 12 months; that infant was excluded from the statistical test in Study 1.

Study 2 Method: Infant Syllable Repertoires Determined by Naturalistic

Listening—The naturalistic listening task was developed for this research in an attempt to simulate caregiver judgment in the laboratory. Listeners performed an utterance-level (UL) and a session-level (SL) task. At the UL, listeners heard an individual utterance, and then they verbally made a judgment about that utterance. At the SL, listeners heard all utterances from a single recording session, and then they verbally made a judgment about the group of utterances. Four members of our laboratory staff performed these naturalistic listening tasks. Each of the listeners was used in a project on infant vocalizations, early speech development, or both, and regularly participated in infant vocalization coding, but this task was unique, and no training was given.

The procedure for Study 2 began with utterance identification and extraction in acoustic analysis software TF32 (Milenkovic, 2001). Infant utterances were located by using a breath-group criterion (each vocalization occurred on a single egressive breath; Oller & Lynch, 1992). Vegetative or reflexive sounds and vocalizations with significant vocal overlay were not included. At least 70 utterances were extracted per infant per age, with a total of 2,269 utterances ($M = 94.54$, $SD = 28.53$; see Supplementary Table 2).

Two stimulus subsets were created, each including all 2,269 utterances: one for the UL task and one for the SL task. The UL subset consisted of 2,269 discrete utterance wave files. The

SL subset consisted of 24 randomly ordered wave files (one for each of 8 infants at three ages) made up of the 2,269 utterances. The 24 SL wave files were 4 to 10 min in duration, depending on the number of utterances in the session. Stimuli from both UL and SL tasks were combined in a single, complex, randomized set with utterances separated by 3 s of silence.

Prerecorded instructions were presented to the listeners, indicating whether a UL or SL stimulus was about to be heard. The listeners would click *play*, hear either an individual utterance (UL) or all of the utterances from a session (SL), and then speak their judgments into the microphone of a digital voice recorder (Olympus VN-4100 PC). The listeners imitated the sound or sounds they believed the infant to have produced, based exclusively on the utterance or group of utterances they heard. High-quality speakers (Logitech IHX) with a built-in amplifier were used to ensure audibility of even low-intensity utterances. Laboratory staff completed the naturalistic listening tasks in approximately seven sittings (a total of 5–6 hr of effort per listener). Each of the listeners heard a different semi-random ordering of the UL- and SL-stimulus subsets in an attempt to prevent them from realizing that each utterance occurred twice and to ensure blindness in regard to infant identity and age.

To determine the repertoire of syllables for each infant at each age, (a) the first author, an experienced transcriber, phonetically transcribed each audio-recorded judgment made by laboratory staff during the naturalistic listening tasks; (b) the transcribed judgments were sorted by infant, age, and task (UL or SL), resulting in 24 separate transcription files at the UL and 24 at the SL for each listener; then (c) the 48 transcription files were analyzed in the CV syllable program to tally syllables. The outcome measure was the mean number of CV syllables attributed to each infant at each age across listeners for the two naturalistic listening tasks. Data from the four listeners are collapsed in the main report. Supplementary Table 3 illustrates that all of the listeners showed much larger repertoires at the UL than at the SL for every age and every infant.

Observer variation was minimized because a single individual (the first author) transcribed judgments derived from the naturalistic listening tasks. However, the transcriber was not blind to the level of the listening task (UL or SL). To provide a reliability assessment and to test for any possible bias effect due to lack of blinding, a second transcriber, blind to the hypotheses of the study and to the level of the task (UL or SL), transcribed 10% of the recorded judgments. This subset of judgments transcribed a second time included a random assortment of utterances from all listeners, infants, infant ages, and naturalistic listening tasks.

For Study 2, the primary focus was listener judgment of syllable repertoires at the UL (estimating RM) versus the SL (estimating NP). The study assessed a role for memory in listening, consistent with theory presented in Cowan (1995). UL listening was expected to reveal good recall of phonetic detail (RM), as each utterance was judged immediately after it was presented. SL listening was expected to reveal poorer recall of detail (hence, NP), as judgments were delayed until an entire SL stimulus set was heard.

A repeated-measures ANOVA evaluated the 2×3 design: The dependent variable was number of syllable types perceived across naturalistic listeners; the within-subject independent variables were listening task (UL vs. SL) and infant age (8, 10, and 12 months). Only one primary prediction was made: Fewer syllable types would be perceived at the SL than the UL, reflecting the difference between NP and RM. Second, we anticipated that NP syllable inventories would increase with age and that RM inventories would either remain the same or increase.

Study 3 Method: Infant Syllable Repertoires Determined by Phonetic

Transcription—Phonetic transcription was used in this study as a way to reference the more traditional means of acquiring data on syllable repertoires. In LIPP, listeners transcribed a preselected set of utterances twice in two listening conditions. In the first, each utterance (randomly ordered) was transcribed after being presented once. In the second, each utterance (again randomly ordered) was transcribed while listening to six presentations, with edits allowed after each of the six. Four transcribers, all student workers employed on projects in infant vocalizations or early speech development (different individuals than in Study 2) performed the Study 3 transcriptions. They had all received ample training transcribing in LIPP through a graduate-level course in phonetic transcription and laboratory training. These individuals were trained and drilled on sounds and symbols of the IPA, including base symbols and diacritics, and met as a group to go through the steps with practice data to ensure understanding of the procedure before transcribing real data for the study.

Stimuli for Study 3 consisted of a subset of the 2,269 utterances extracted from all recording sessions. This subset included utterances where all four listeners in Study 2 reported at least one CV syllable. On average there were 29 utterances per session that met this criterion, for a total of 685 utterances ($M = 28.54$, $SD = 12.65$; see Supplementary Table 2) to be presented to transcribers in Study 3. The transcribers were instructed to transcribe every sound or syllable in each utterance, CV, or otherwise. Explicit instruction encouraged use of the range of IPA base and diacritic symbols to discourage bias associated with familiarity with the English phonemic system, a point previously emphasized in classroom and laboratory training. Each transcriber was required to work completely independently.

Each transcriber heard two counterbalanced, semi-random orderings of the CV utterances, with one and with six listening opportunities. Custom software played the stimuli, and transcribers were instructed at the beginning of each file to code utterances after one or six listens. Transcribers (always working in LIPP) were blinded to the fact that each utterance was presented twice at some point in the task, as well as to infant identity and age. The total procedure took 15 to 20 hr for each transcriber, across 10 to 15 sittings. After transcriptions were complete, LIPP analysis produced two compilations of CV syllables transcribed by each coder, one for utterances transcribed after one listening opportunity and one for utterances transcribed after six listening opportunities. The transcriptions were tabulated by the CV syllable program.

The primary outcome measure for Study 3 was the mean number of syllables attributed by the four transcribers to each infant at each age for each of the two listening conditions. More phonetic detail (corresponding with RM) was expected to be portrayed when transcribers listened to each utterance six times than one time. A repeated-measures ANOVA was used to evaluate the 2×3 design. The dependent variable was the mean number of syllable types recorded by the transcribers. The within-subject independent variables were listening opportunities (one vs. six) and age (8, 10, and 12 months). One primary prediction was made: Fewer syllable types would be attributed to infants when utterances were heard once as compared with six times. A secondary prediction was that syllable inventories would increase with age.

Results

Study 1: Infant Syllable Repertoires Determined by Caregiver Report

Table 1 shows that caregivers attributed a small number of syllable types to their infants at all ages (range = 3–14). Although age differences were not statistically significant, $F(2, 5) = 2.57$, $p = .171$, medium effect sizes suggested larger repertoires at 12 months (8 vs. 12, $d =$

0.78), a fact that can be observed through visual inspection of differences between judged repertoire sizes displayed in Figure 1. The number of syllables reported by caregivers was not significantly correlated with the total number of utterances extracted per recording, offering no indication of a positive relation between infant volubility and perceived syllable repertoire (mean $r = .33$ for the three ages and two listening tasks).

Study 2: Infant Syllable Repertoires Determined by Naturalistic Listening

The primary goal in Study 2 was to test for different sizes of estimated syllable inventories based on the two naturalistic listening tasks (SL vs. UL). The mean number of syllables attributed to infants at the UL was more than six times larger than at the SL (see Table 2, where values represent the mean number of syllables determined through naturalistic listening by four laboratory staff members; see Supplementary Table 3), with all 24 possible comparisons across SL and UL listening for the eight infants at three ages showing a much larger inventory at the UL. Consequently, the significant main effect for listening task, $F(1, 7) = 398.75, p < 10^{-6}$, was extremely large at 8, 10, and 12 months ($ds = 5.70, 6.01$, and 5.05 , respectively). All four listeners showed a similar major effect of listening task (Supplementary Table 3). Further, repertoire size estimated through naturalistic listening was not significantly correlated with the total number of utterances extracted per recording session (mean $r = .20$ for the three ages and two listening tasks), again providing no indication of positive relation between infant volubility and syllable repertoire size.

There was no statistically significant main effect for age, $F(2, 14) = 1.35, p > .05$, nor for the interaction between listening task and age, $F(2, 14) = 1.35, p > .05$, although most of the trends suggest increase (exception: 10 to 12 months at the UL). Between 8 and 12 months, the effect was medium at the UL ($d = 0.39$) and large at the SL ($d = 0.89$). Again, these differences are apparent in Figure 1.

Similar results were found for the subset of judgments made from naturalistic listening transcribed by two coders for reliability. All eight infants, three infant ages, and two listening tasks were represented in the reliability data, which included 10% of the total stimulus set. As transcribed by the original coder, the reliability subset showed a large and highly reliable main effect for listening task, $F(1, 7) = 1,443.8, p < .001, d = 5.2$; with no significant main effect for age, $F(2, 14) = 1.5, p > .05$; and no significant main effect for the interaction between task and age, $F(2, 14) = 0.2, p > .05$. As transcribed blind by the second coder, the reliability subset showed a similar large and highly reliable main effect for listening task, $F(1, 7) = 976.6, p < .001, d = 6.4$, with no significant main effect for age, $F(2, 4) = 3.1, p > .05$, and no significant main effect for the interaction between task and age, $F(2, 4) = 1.2, p > .05$. Paired samples t tests that compared transcribers for each level of task and age (e.g., the original transcriber at the SL for 8 months compared with the second transcriber at the SL for 8 months) showed no significant effects. Thus, lack of blindness of the original coder to the listening task did not appear to bias the results.

Study 3: Infant Syllable Repertoires Determined by Phonetic Transcription

The numbers of CV syllables ascribed to infants based on phonetic transcriptions for one versus six listening opportunities at each infant age are given in Table 3 (presenting mean number of syllables determined from phonetic transcription by four laboratory staff members; for transcriber data, see Supplementary Table 4). The primary goal was to test for effects of number of listening opportunities. As hypothesized, the main effect was statistically significant, with CV syllable types attributed to infants being higher with six listening opportunities ($M = 27.11$) compared with one listening opportunity ($M = 21.96$), $F(1, 7) = 62.33, p < .001$. The high significance level is consistent in that, in all 24 comparisons, of the eight infants at three ages, there was a higher number of syllables, on

average, across the four transcribers for six listening opportunities than for one (see Supplementary Table 4). Medium effect sizes for the difference between repertoires from one or six listening opportunities were seen at 8, 10, and 12 months ($d = 0.72, 0.53$, and 0.54). As with caregiver report and judgments made through naturalistic listening, the repertoire sizes estimated through phonetic transcription were not significantly correlated with the total number of utterances extracted per recording session (mean $r = .35$ for the three ages and two transcription conditions). This, again, provided no indication of positive relation between infant volubility and syllable repertoire size.

There was no statistically significant main effect for age, $F(2, 14) = 2.66, p > .05$, or for the interaction between listening opportunity and age, $F(2, 14) = 0.65, p > .05$. The numbers of CV syllable types at 8 months, however, were suggestively lower than at 10 and 12 months. Between 8 and 12 months, the effect was large with both one ($d = 0.89$) and six ($d = 0.81$) listening opportunities (see Figure 1).

Cross-Study Comparison

Comparison across the three studies is shown in Figure 1. With respect to the size of the repertoires perceived, results suggest that caregivers and listeners at the SL recognized similar numbers of syllables in infant repertoires (averaging fewer than seven). Listeners at the UL and phonetic transcribers with both one and six listening opportunities reported much larger numbers of different syllable types than caregivers (ranging from 17 to 40 per session). The data illustrate that transcription and listening in a more naturalistic way at the UL each yielded infant repertoire sizes greatly exceeding those found through caregiver reports and estimated from listening in a more naturalistic way at the SL.

To explore the phonetic makeup of perceived syllable types across groups of listeners, we developed under-specified phonological templates based on caregiver reports of infant syllable repertoires. We began with a list of each syllable attributed to each infant at each age by the caregivers. Templates were always CV in form. Thus, we ignored V onsets and C offsets. Also, because within each infant age V contrasts were rare in caregiver reports and showed limited co-occurrence patterns with Cs, we opted to collapse all Vs when generating templates. Further, templates were formulated such that all contrasts of Cs reported by caregivers for any session were accounted for in the following ways: (a) manner of C production was described as labial, coronal, lateral, or dorsal, and (b) place of C production was described as obstruent, nasal, or semivowel. With respect to coronal and labial Cs, the common existence in caregiver reports of both semivowels and nasals contrasting with obstruents provided a basis for positing templates to differentiate these types. For example, if a caregiver reported the infant possessed both [ba] and [da], these were assigned respectively to underspecified NP templates labial obstruent plus V and coronal obstruent plus V.

Aggregated caregiver reports yielded eight templates accounting for all contrastive CV syllable types across infants and ages: coronal obstruent, labial nasal, labial obstruent, dorsal, coronal nasal, coronal semivowel, lateral, and labial semivowel plus V templates. Coronal obstruent plus V templates, such as /da/, were reported by caregivers most often, a total of 30 times. Labial semivowel plus V templates, such as /wa/, were reported least often, a total of 4 times.

Table 4 presents the numbers of syllables aggregated across infants and ages assigned to each template for caregivers, listeners at the SL and UL, and transcribers with one and with six listening opportunities. For each laboratory condition, template counts were correlated with caregiver-based template counts. The correlations indicate strong positive prediction of relative frequencies for templates based on caregiver report with those for listeners at the SL

($r = .87, p < .01$) and for transcribers with both one and six listening opportunities ($r = .76, p < .05$; and $r = .75, p < .05$, respectively). The predictiveness was lower but still accounted for more than 20% of variance for listeners at the UL ($r = .45, p > .05$).

Discussion

Our proposal suggests a theoretical and a methodological enhancement with regard to the evaluation of babbling as a foundation for word learning. On the theoretical side, we emphasize interaction between caregiver and infant as a fulcrum for early phonological category development even before infant productive words exist. Prior efforts to determine syllable repertoires in babbling have largely been based on data derived exclusively from phonetic transcriptions and collapsed according to frequencies of occurrence or phonetic principles independent of caregiver recognition of categories (see, e.g., MacNeilage, 2008). We have argued that the caregiver is the critical judge of the infant repertoire, being a key participant in negotiation over word learning.

On the methodological side, our approach provides some first suggestions about how to acquire both caregiver estimates of the infant's prelinguistic phonological categories and a simulated estimate based on naturalistic listening in the laboratory. The naturalistic listening procedure can be conducted rapidly and appears to replicate the parent report in terms of repertoire size fairly well, although the initial methods used here may underestimate the repertoires in both cases. In addition, naturalistic listening yielded similar prelinguistic templates to the caregiver report. We have, thus, pointed toward ways to assess babbling repertoire without reference to words, emphasizing natural listening. Future study of caregiver report and naturalistic listening procedures, with tracking of word development into the second year and beyond, will enable us to determine the extent to which these methods may predict later speech and language outcomes.

Estimating the Size of NP and RM Repertoires

Our results demonstrate that early NP (negotiable phonological product, recognized in the natural listening circumstance) appears to consist of only a handful of syllabic categories (which can be characterized as underspecified phonological templates), whereas RM (raw phonological material, as reported by transcribers and laboratory staff listening repeatedly to utterances) appears to consist of a much larger number of phonetic variants pertaining to each NP category. We argue that in the absence of recognizing the distinction between NP and RM, we run the risk of overestimating the size of the syllabic repertoire that is available for negotiation in early word learning. The present work provides rough estimates of the magnitude of NP repertoires in the latter part of the first year (see Figure 1). Caregiver responses to the simple question, "What sounds does your infant produce?" yielded an estimate of NP repertoire size (five to seven syllables) similar to that obtained when listeners at the SL (listening in a naturalistic way) heard 90 to 100 utterances from an infant and then verbally reported the syllable repertoires they thought the infants controlled. Compared with our estimated NP inventory sizes, the typical RM inventory size was larger by a factor of at least five when gauged from UL naturalistic listening and by a factor of at least three when based on phonetic transcription with repeated listening.

There is in fact no obvious limit on the size of RM repertoires. Every vocalization appears to be unique when viewed spectrographically, and consequently, the RM for any NP category and for any infant can be viewed as indefinitely large. This is not to say that infant NP categories are undefined. Each tends to occupy a limited portion of the vocal space, appearing to consist of many RM variants with similar acoustic features. Further, each NP category of an infant is identifiable by caregivers or laboratory observers and may also be

characterized by a phonological template, thus, differentiating it from other NP categories of the infant.

The Possibility of Better Outcome Prediction

Characteristics of infant vocal development have long been researched as a potential basis for predicting outcomes in language and cognition (Cameron et al., 1967; Lyytinen et al., 1996; Menyuk, Liebergott, & Schultz, 1986; Obenchain, Menn, & Yoshinaga-Itano, 1998; Roe, 1975). Procedures to make such predictions, however, have never involved estimates of infant pre-linguistic phonological categories derived through a principled basis with reference to the caregiver or the caregiver's natural mode of listening. It seems reasonable to expect that this principled method could help to better predict language outcomes than more traditional methods that ignore natural listening. Analysis of pre-linguistic templates derived from caregiver report and templates derived from laboratory methods suggests on the one hand that one-time listening to each utterance in transcription appears to produce *no less concordance* with caregiver judgment for frequency of occurrence of templates than listening six times (see Table 4). Even more notably, templates from naturalistic listening at the SL provided no less concordance to caregiver templates than phonetic transcription did. In fact, the results suggest that naturalistic listening at the SL may provide a better match to caregiver report than phonetic transcription or naturalistic listening at the UL. The data imply that with further improvement of caregiver report and naturalistic listening protocols, it may be possible to develop better models of early phonology and better predictors of future speech–language outcomes and anomalies in development.

The Role of Recording Sample Size in Repertoire Size Estimates

Naturalistic listening and transcription procedures probably needed larger sample sizes in order to produce optimal estimates of both repertoire size and content. We speculate that estimates of repertoire sizes based on phonetic transcriptions (with both one and six listening opportunities) were lower than those for naturalistic listening at the UL due to sample size differences between the two studies; listeners in Study 2 judged all 2,269 utterances, whereas transcribers in Study 3 only judged 685 utterances derived from Study 2 as including CV syllables. Thus, transcribers made judgments based on a subsample of less than one third of the utterances judged in naturalistic listening tasks. A larger sample for transcribers would likely have included more RM variants to transcribe and, consequently, larger repertoires to portray. Thus, repertoire sizes estimated from transcription and naturalistic listening at the UL are not fully comparable.

To our knowledge no one has ever investigated sample-size requirements for estimates of syllable repertoires in laboratory-based research on infant vocal development. However, new work (Molemans, Van den Berg, Van Severen, & Gillis, 2012) estimates sample-size requirements for reliably determining onset of canonical babbling. For 95% confidence, 300 to 500 utterances were shown to be required, depending on age. It seems likely that estimating syllable repertoire sizes within canonical babbling and obtaining a reliable portrayal of the phonetic templates of each infant will require samples at least as large. Our study did not include such large samples, and yet there was fairly good agreement between both repertoire sizes and repertoire content (in terms of templates) across aggregated caregiver report data and aggregated naturalistic listening data at the SL. Note that the small size of the syllable repertoire reported by caregivers and SL listeners is plausible from a practical standpoint at the ages we studied—a very small lexicon can easily be transmitted by a very small syllable repertoire, and no child in this study had more than a few words during the period of the research.

Without more extensive investigation, it will not be possible to decisively interpret the match between caregiver and naturalistic listening-based report. There are good reasons to expect both caregiver-based and naturalistic listening-based reports of repertoires to vary depending on several factors. An important limitation of our work can be seen in its reliance on caregiver responses to a single, simple interview question, posed only at periodic visits of participants to the laboratory. It is easy to imagine more reliable ways to acquire information from caregivers using such methods as computerized parent diaries (currently being implemented for more systematic report of infant vocal inventories and their functional uses).

On the Lack of Robust Age Effects

Secondary predictions of these studies (that NP would increase with age) were not statistically significant. However, effect sizes were medium to large, effects that might prove statistically significant with larger sample sizes. Vocal repertoire sizes must increase at some stage, but perhaps growth beyond the numbers of NP types seen here occurs for most children only beyond 12 months. Indeed, 12 months is a typical age of onset of first words (Fenson et al., 1994; Huttenlocher & Smiley, 1987)—only after 12 months, then, might children need additional syllables for word formation. Also, 12-month repertoires could appear to be small because of frequent production of a limited number of first words.

A Note on Possible Effects of Culture

Our proposal regarding negotiation between caregivers and infants over potential meanings to be assigned to babbled syllables is not intended to imply that all families or all cultures engage in negotiation over babbling in identical ways, nor to the same extents. Indeed, it has been widely documented that baby talk does not occur in the same ways across all cultures (Heath, 1983; Ochs & Schieffelin, 1984) and that some parents interact vocally far less with their infants than is observed in the American middle class (Hart & Risley, 1995; Hoff, 2003; Kulick & Schieffelin, 2004; Pesco & Crago, 2007). Cross-linguistically, caregivers respond in different ways to the same infant input. Still, this cultural variation does not suggest that caregiver–infant interaction plays no role in word learning. Even if parents do not interact vocally with infants very much, other caregivers (e.g., older siblings, playmates, aunts or uncles, etc.) may, and at some point some caregiver must interact in this way for the child's vocal communications to be functional. We take seriously the argument that differences across cultures may be more quantitative than qualitative with respect to negotiation in babbling and early speech (Lieven, 1994, 1997).

We contend that, in spite of familial–cultural differences, negotiation and the distinction between NP and RM will be relevant in all cases (although to varying extents). When infant babble is heard by any caregiver, principles of distributional learning will apply, whether the caregivers are paying attention or not. In accord with those principles, we reason caregivers will tend to recognize NP and to a large extent ignore RM. The intensity and frequency of occurrence of events of negotiation (e.g., imitation of babbling, elicitation of vocalization, proposal of meanings, etc.) may differ widely across families and cultures, but it seems likely they occur universally at some point in development. Further, whether or not negotiation occurs over infant babbling in any individual family, it is clear that it occurs very frequently in middle-class American families, who are the focus of the present research.

Another issue related to possible cultural effects in our study concerns the possibility that listeners may have been influenced by their varying language or dialect backgrounds or infants may have differed in vocal production of syllables because of their variable language or dialect backgrounds. Indeed it has been reported that listeners are affected in phonetic

transcription by their language background (Coussé et al., 2004). However, such effects have been focused at the segment level and have not been focused on infancy nor have they been shown to affect syllable inventories determined by transcription. In a similar way, there have been numerous reports of ambient language effects on phonetically transcribed utterances of infancy (de Boysson-Bardies, Halle, Sagart, & Durand, 1989; Levitt & Wang, 1991; Thevenin, Eilers, Oller, & LaVoie, 1985), but other studies have reported no such effects of ambient language (Atkinson, MacWhinney, & Stoel, 1970), and again the work in both cases has not directly addressed possible effects on syllable inventory size. As a consequence, at this point it is uncertain what role cultural–linguistic differences may play in our results or what role transcription at a raw material or negotiable product level might play.

Conclusions

Our approach suggests a principled basis for determining infant syllable repertoires utilizing information provided by caregivers or obtained in naturalistic listening. Exploration of caregiver perception and naturalistic listening enables us to discern global features of syllable categories even in the midst of variation in how the infant produces each category moment to moment. Our proposal adds to prior laboratory methods that have assessed infant vocal repertoires primarily through phonetic transcription but have not directly considered the most important speech perceiver in the infant's life, the caregiver.

We have argued for the importance of caregiver-based repertoire reports, along with a laboratory method to simulate repertoire size estimates within a range comparable to that reported by caregivers. Our results have illustrated that estimated repertoire sizes can be at least several times larger when observers depart from the styles of listening that are natural to the caregiver in the home. In our proposal, the caregiver's judgments become the centerpoint of repertoire assessments, rather than an afterthought.

At least in part, these studies suggest that caregiver recognition of only a small number of prelinguistic phonological categories (NP), each composed of a large number of variable productions (RM), appears to be due to the caregiver's natural memory constraints. Perhaps paradoxically, these constraints aid the caregiver in recognizing an essential infant capability for word learning, the ability to produce a small number of consolidated vocal patterns at will. Infant vocal categories are, thus, seen to emerge from a combination of systematic though complex infant exploration and caregiver recognition of global patterns within that exploration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Plough Foundation and by National Institute on Deafness and Other Communication Disorders Grants R01 DC006099 and R01 DC011027, both awarded to the second author.

References

- Adolph KE, Robinson SR, Young JW, Gill-Alvarez F. What is the shape of developmental change? *Psychological Review*. 2008; 115:527–543. [PubMed: 18729590]
- Alford, AR. Unpublished master's thesis. University of Memphis; TN: 2006. Procedures for discerning syllabic templates in babbling: How infants form phonological categories.

- Atkinson K, MacWhinney B, Stoel C. An experiment in the recognition of babbling. *Papers and Reports on Child Language Development*. 1970; 5:1–8.
- Ball M, Müller N, Klopfenstein M, Rutter B. The importance of narrow phonetic transcription for highly unintelligible speech: Some examples. *Logopedics, Phoniatrics, Vocology*. 2009; 34:84–90.
- Bloom K. Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*. 1988; 15:469–480. [PubMed: 3198716]
- Bloom K, Russell A, Wassenberg K. Turn taking affects the quality of infant vocalizations. *Journal of Child Language*. 1987; 15:211–227. [PubMed: 3611239]
- Bloomfield, L. *Language*. New York, NY: Holt; 1933.
- Buder EH, Stoel-Gammon C. Young children's acquisition of vowel duration as influenced by language: Tense/lax and final stop consonant voicing effects. *The Journal of the Acoustical Society of America*. 2002; 111:1854–1864. [PubMed: 12002869]
- Cameron J, Livson N, Bayley N. Infant vocalizations and their relationship to mature intelligence. *Science*. 1967 Jul 21; 157:331–333. [PubMed: 6028407]
- Coussé, E.; Gillis, S.; Kloots, H.; Swerts, M. In: Lino, M.; Xavier, M.; Ferreira, F.; Costa, R.; Silva, R., editors. *The influence of the labeller's regional background on phonetic transcriptions: Implications for the evaluation of spoken language resources*; Proceedings of the 4th International Conference on Language Resources and Evaluation; Paris, France: ELRA; 2004. p. 1447-1450.
- Cowan, N. *Attention and memory: An integrated framework*. New York, NY: Oxford University Press; 1995.
- Cucchiari C. Assessing transcription agreement: Methodological aspects. *Clinical Linguistics & Phonetics*. 1996; 10:131–155.
- Dale PS, Dionne G, Eley TC, Plomin R. Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language*. 2000; 27:619–642. [PubMed: 11089341]
- de Boysson-Bardies B, Halle P, Sagart L, Durand C. A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*. 1989; 16:1–17. [PubMed: 2925806]
- Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. *Monographs of the Society for Research in Child Development*. 1994; 59(5 Serial No 242)
- Fenson, L.; Dale, P.; Reznick, JS.; Thal, D.; Bates, E.; Hartung, J.; Reilly, JS. *The MacArthur Communicative Development Inventories*. San Diego, CA: Singular; 1991.
- Ferguson CA, Farwell CB. Words and sounds in early language acquisition: English initial consonants in the first fifty words. *Language*. 1975; 51:419–439.
- Firth JR. Sounds and prosodies. *Transactions of the Philological Society*. 1948:127–152.
- Fogel A, Garvey A. Alive communication. *Infant Behavior & Development*. 2007; 30:251–257. [PubMed: 17376534]
- Goldstein MH, King AP, West MJ. Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, USA*. 2003; 100:8030–8035.
- Goldstein MH, Schwade JA. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*. 2008; 19:515–522. [PubMed: 18466414]
- Goudbeek, M.; Swingle, D. Saliency effects in distributional learning. Paper presented at the Proceedings of the 11th Australian International Conference on Speech Science & Technology; Auckland, New Zealand. 2006 Dec.
- Gros-Louis J, West MJ, Goldstein MH, King AP. Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*. 2006; 30:509–516.
- Gupta RS, Cohen NJ. Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*. 2002; 109:401–448. [PubMed: 11990324]
- Hart, B.; Risley, TR. *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes; 1995.
- Heath, SB. *Ways with words: Language, life, and work in communities and classrooms*. New York, NY: Cambridge University Press; 1983.

- Hinde, RA. Expression and negotiation. In: Zivin, G., editor. *The development of expressive behavior: Biology–environment interactions*. New York, NY: Academic Press; 1985. p. 103–116.
- Hockett CF. A system of descriptive phonology. *Language*. 1942; 18:3–21.
- Hoff, E. Causes and consequences of SES-related differences in parent-to-child speech. In: Bornstein, MH.; Bradley, RH., editors. *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Erlbaum; 2003. p. 147–160.
- Huttenlocher J, Smiley P. Early word meanings: The case of object names. *Cognitive Psychology*. 1987; 27:63–89.
- International Phonetics Association. The principles of the International Phonetic Association (1949). *Journal of the International Phonetic Association*. 2010; 40:299–358.
- Jakobson, R. *Kindersprache, Aphasie, und allgemeine Lautgesetze* [Child language, aphasia, and phonological universal]. Uppsala, Sweden: Almqvist and Wiksell; 1941.
- Jusczyk, PW. *The discovery of spoken language*. Cambridge, MA: MIT Press; 1997.
- Keating PA. Underspecification in phonetics. *Phonology*. 1988; 5:275–292.
- Keren-Portnoy T, Majorano M, Vihman MM. From phonetics to phonology: The emergence of first words in Italian. *Journal of Child Language*. 2009; 36:235–267. [PubMed: 18789180]
- Kiparsky P. Some consequences of lexical phonology. *Phonology Yearbook*. 1985; 2:85–138.
- Koopmans-van Beinum, FJ.; van der Stelt, JM. Early stages in the development of speech movements. In: Lindblom, B.; Zetterstrom, R., editors. *Precursors of early speech*. New York, NY: Stockton Press; 1986. p. 37–50.
- Kulick, D.; Schieffelin, BB. Language socialization. In: Duranti, A., editor. *A companion to linguistic anthropology*. Malden, MA: Blackwell; 2004. p. 349–368.
- Kwon, K.; Oller, DK.; Buder, EH. Evidence of systematic repetition in infant vocalizations. Paper presented at the American Speech-Language-Hearing Association Convention; Boston, MA. 2007 Nov.
- Labov W, Labov T. The phonetics of cat and mama. *Language*. 1978; 54:816–852.
- Levitt AG, Wang Q. Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech*. 1991; 34:235–249. [PubMed: 1843525]
- Lieven, EVM. Cross-linguistic and cross-cultural aspects of language addressed to children. In: Gallaway, C.; Richards, BJ., editors. *Input and interaction in language acquisition*. Cambridge, England: Cambridge University Press; 1994. p. 56–73.
- Lieven, EVM. Variation in a crosslinguistic context. In: Slobin, DI., editor. *The cross-linguistic study of language acquisition: Expanding the contexts*. Vol. 5. Mahwah, NJ: Erlbaum; 1997. p. 199–263.
- Locke, JL. *Phonological acquisition and change*. New York, NY: Academic Press; 1983.
- Lyytinen P, Poikkeus AM, Leiwo M, Ahonen T. Parents as informants of their child's vocal and early language development. *Early Child Development & Care*. 1996; 126:15–25.
- MacNeilage, PF. *The origin of speech*. Oxford, England: Oxford University Press; 2008.
- Menyuk, P.; Liebergott, J.; Schultz, M. Predicting phonological development. In: Lindblom, B.; Zetterstrom, R., editors. *Precursors of early speech*. New York, NY: Stockton Press; 1986. p. 79–93.
- Milenkovic, P. TF32 [Computer program]. Madison WI: University of Wisconsin—Madison; 2001.
- Molemans I, Van den Berg R, Van Severen L, Gillis S. How to measure the onset of babbling reliably. *Journal of Child Language*. 2012; 39:523–552. [PubMed: 21892989]
- Munson B, Edwards J, Schellinger SK, Beckman ME, Meyer MK. Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical Linguistics & Phonetics*. 2010; 24:245–260. [PubMed: 20345255]
- Naigles, LR.; Hoff, E.; Vear, D. Flexibility in early verb use: Evidence from a multiple-N diary study. Hoboken, NJ: Wiley; 2009.
- Nathani S, Oller DK. Beyond ba-ba and gu-gu: Challenges and strategies in coding infant vocalizations. *Behavior Research Methods, Instruments & Computers*. 2001; 33:321–330.

- Obenchain P, Menn L, Yoshinaga-Itano C. Can speech development at 36 months in children with hearing loss be predicted from information available in the second year of life? *Volta Review*. 1998; 100:149–180.
- Ochs, E.; Schieffelin, BB. Language acquisition and socialization: Three developmental stories. In: Levine, RSR., editor. *Culture theory: Essays on mind, self, and emotion*. Cambridge, England: Cambridge University Press; 1984. p. 276-320.
- Oller, DK. The emergence of the sounds of speech in infancy. In: Yeni-Komshian, G.; Kavanagh, J.; Ferguson, C., editors. *Child phonology: Production*. Vol. 1. New York, NY: Academic Press; 1980. p. 93-112.
- Oller, DK.; Delgado, RE. *Logical International Phonetics Programs (Windows version)* [Computer software]. Miami, FL: Intelligent Hearing Systems; 1999.
- Oller DK, Eilers RE, Basinger D. Intuitive identification of infant vocal sounds by parents. *Developmental Science*. 2001; 4:49–60.
- Oller, DK.; Griebel, U. The origins of syllabification in human infancy and in human evolution. In: Davis, B.; Zajdo, K., editors. *Syllable development: The frame/content theory and beyond*. Mahwah, NJ: Erlbaum; 2008. p. 368-386.
- Oller, DK.; Lynch, MP. Infant vocalizations and innovations in infraphonology: Toward a broader theory of development and disorders. In: Ferguson, C.; Menn, L.; Stoel-Gammon, C., editors. *Phonological development*. Parkton, MD: York Press; 1992. p. 509-538.
- Oller DK, Ramsdell HL. A weighted reliability measure for phonetic transcription. *Journal of Speech, Language, and Hearing Research*. 2006; 49:1391–1411.
- Papaeliou C, Minadakis G, Cavouras D. Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech, Language, and Hearing Research*. 2002; 45:311–317.
- Papoušek, M. Vom ersten Schrei zum ersten Wort: Anfänge der Sprachentwicklung in der vorsprachlichen Kommunikation [From the first cry to the first word: The beginning of linguistic development in preverbal communication]. Bern, Switzerland: Verlag Hans Huber; 1994.
- Pesco, D.; Crago, M. Language socialization in Canadian Aboriginal communities. In: Duff, PA.; Hornberg, NH., editors. *Encyclopedia of language and education*. 2. Vol. 8. New York, NY: Kluwer Academic; 2007. p. 273-285.
- Pollock KE, Berni MC. Transcription of vowels. *Topics in Language Disorders*. 2001; 21:22–40.
- Preston JL, Ramsdell HL, Oller DK, Edwards ML, Tobin SJ. Developing a weighted measure of speech sound accuracy. *Journal of Speech, Language, and Hearing Research*. 2011; 54:1–18.
- Ramsdell HL, Oller DK. Predicting phonetic transcription agreement: Insights from research in infant vocalizations. *Clinical Linguistics & Phonetics*. 2007; 21:793–831. [PubMed: 17882695]
- Ramsdell, HL.; Oller, DK.; Buder, EH. Early empirical results on the emergence of phonological categories: Raw phonological material and negotiable phonological product. Paper presented at the International Child Phonology Conference; Seattle, WA. 2007 Jun.
- Roe KV. Amount of infant vocalization as a function of age: Some cognitive implications. *Child Development*. 1975; 46:936–941.
- Rvachew S, Creighton D, Sauve R, Feldman N. Vocal development of infants with very low birth weight. *Clinical Linguistics & Phonetics*. 2005; 19:275–294. [PubMed: 16019776]
- Shriberg LD, Lof GL. Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics & Phonetics*. 1991; 5:225–279.
- Smith A, Zelaznik H. The development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*. 2004; 45:22–33. [PubMed: 15229873]
- Stark RE, Bernstein LE, Demorest ME. Vocal communication in the first 18 months of life. *Journal of Speech and Hearing Research*. 1993; 36:548–558. [PubMed: 8331912]
- Stockman IJ, Woods DR, Tishman A. Listener agreement on phonetic segments in early infant vocalizations. *Journal of Psycholinguistic Research*. 1981; 10:593–617. [PubMed: 7343665]
- Stoel-Gammon C. Prelinguistic vocalizations of hearing-impaired and normally hearing subjects: A comparison of consonantal inventories. *Journal of Speech and Hearing Disorders*. 1988; 53:302–315. [PubMed: 3398483]

- Stoel-Gammon, C. Prelinguistic vocal development: Measurement and prediction. In: Ferguson, C.; Menn, L.; Stoel-Gammon, C., editors. *Phonological development: Models, research, implications*. Timonium, MD: York Press; 1992. p. 439-456.
- Stoel-Gammon C. Relationships between lexical and phonological development in young children. *Journal of Child Language*. 2011; 38:1–34. [PubMed: 20950495]
- Thevenin D, Eilers RE, Oller DK, LaVoie L. Where's the drift in babbling drift? A cross-linguistic study. *Applied Psycholinguistics*. 1985; 6:3–15.
- Velleman SL, Vihman M. Whole-word phonology and templates: Trap, bootstrap, or some of each? *Language and Speech*. 2002; 33:9–23.
- Veneziano, E. Vocal-verbal interaction and the construction of early lexical knowledge. In: Smith, MD.; Locke, JL., editors. *The emergent lexicon: The child's development of a linguistic vocabulary*. San Diego, CA: Academic Press; 1988. p. 109-147.
- Vihman, MM. Individual differences in babbling and early speech. In: Lindblom, B.; Zetterstrom, R., editors. *Precursors of early speech*. New York, NY: Stockton Press; 1986. p. 21-35.
- Vihman, M. Early syllables and the construction of phonology. In: Ferguson, CA.; Menn, L.; Stoel-Gammon, C., editors. *Phonological development: Models, research, implications*. Timonium, MD: York Press; 1992. p. 393-422.
- Vihman M, Croft W. Phonological development: Toward a 'radical' templatic phonology. *Linguistics*. 2007; 45:683–725.
- Vihman, MM.; DePaolis, RA.; Keren-Portnoy, T. Babbling and words: A dynamic systems perspective on phonological development. In: Bavin, EL., editor. *The Cambridge handbook of child language*. Cambridge, England: Cambridge University Press; 2009. p. 163-182.
- Walsh B, Smith A, Weber-Fox C. Short-term plasticity in children's speech motor systems. *Developmental Psychobiology*. 2006; 48:660–674. [PubMed: 17111401]
- Waterson, N. *Prosodic phonology: The theory and its application to language acquisition and speech processing*. Newcastle upon Tyne, England: Grevatt & Grevatt; 1987.

Figure 1. Number of different syllable types as determined by reports from caregivers, naturalistic listening (NL), and phonetic transcription (T). This is a summary of the inventory results from the three studies.

Table 1

Number of different syllable types determined by caregiver report.

Infant	No. of different syllable types		
	At age 8 months	At age 10 months	At age 12 months
A	9	8	14
B	6	8	6
C	5	5	11
D	4	3	5
E	4	4	3
F	6	7	6
G*	3*	7*	
H	4	4	8
<i>M (SD)</i>	5.43 (1.81)	5.57 (2.07)	7.57 (3.78)

Note. Values marked with an asterisk were not included in statistical significance analysis because of missing data at 12 months but were included in computation of means and standard deviations in the table.

Table 2

Number of different syllable types determined by naturalistic listening.

Level	No. of different syllable types			
	Infant	At age 8 months	At age 10 months	At age 12 months
Utterance level	A	38.25	51.80	51.00
	B	45.50	38.25	29.50
	C	40.00	51.50	37.25
	D	23.25	38.50	38.00
	E	35.50	32.00	46.50
	F	35.75	44.75	31.75
	G	35.25	30.75	42.00
	H	26.75	38.25	25.75
	<i>M (SD)</i>	35.03 (7.10)	41.08 (8.57)	37.72 (8.61)
Session level	A	3.50	7.75	9.50
	B	7.00	7.50	6.25
	C	8.00	7.00	6.00
	D	1.00	4.50	6.75
	E	4.00	5.00	6.25
	F	6.00	6.25	6.25
	G	5.25	3.50	6.75
	H	6.00	7.25	5.75
	<i>M (SD)</i>	5.09 (2.21)	6.09 (1.58)	6.69 (1.19)

Table 3

Number of different syllable types determined by phonetic transcription.

Listening opportunities	Infant	No. of different syllable types		
		At age 8 months	At age 10 months	At age 12 months
1 listen	A	22.25	39.00	36.75
	B	27.50	26.75	16.50
	C	18.25	19.50	29.50
	D	18.00	23.50	12.50
	E	13.00	11.00	28.50
	F	5.50	19.50	23.75
	G	13.75	27.50	14.00
	H	17.75	32.25	30.50
	<i>M (SD)</i>	17.00 (6.54)	24.88 (8.59)	24.00 (8.82)
6 listens	A	27.00	50.00	44.75
	B	32.75	29.00	19.75
	C	22.75	26.50	37.25
	D	22.75	30.50	17.50
	E	21.50	13.25	32.75
	F	8.00	23.25	30.00
	G	17.75	28.25	15.50
	H	23.25	40.00	36.75
	<i>M (SD)</i>	21.97 (7.15)	30.09 (10.98)	29.28 (10.63)

Table 4

Underspecified phonological templates derived from caregivers (C), naturalistic listening (NL), and transcribing (T) across infant and infant age.

Underspecified phonological template	C	NL		T	
		Session level	Utterance level	1 listen	6 listens
Coronal obstruent + vowel	30	182	1,757	1,214	1,473
Labial nasal + vowel	21	82	591	401	452
Labial obstruent + vowel	17	90	1,084	553	701
Dorsal + vowel	13	76	605	513	669
Coronal nasal + vowel	8	42	614	499	524
Coronal semivowel + vowel	8	86	1,341	608	763
Lateral + vowel	8	12	253	161	246
Labial semivowel + vowel	4	41	1,099	289	414
Total	109	611	7,344	4,238	5,242