Relationships Between Vocalization Forms and Functions in Infancy: Preliminary Implications for Early Communicative Assessment and Intervention

Suneeti Nathani Iyer and David J. Ertmer

Purpose: This preliminary study explored relationships between form and function in prelinguistic vocalizations to increase our understanding of early communicative development and to provide potential clinical implications for early communicative assessment and intervention.

Method: Twenty typically developing infants—5 infants in each of 4 age groups, from 3 to 20 months of age—were included. Vocalizations from these infants had previously been categorized for their form (Nathani, Ertmer, & Stark, 2006) and function (Stark, Bernstein, & Demorest, 1993) characteristics. In the present study, cross-classification tabulations between form and function were conducted to examine relationships between vocalization types and their apparent uses.

Results: As anticipated, earlier developing forms were mostly associated with earlier developing functions, and later developing forms were mostly associated with later developing functions. However, there were some exceptions such that some forms were associated with a variety of functions, and vice versa.

Conclusions: The results suggest that some forms are more tightly coupled to function than others in the prelinguistic and early linguistic period. Preliminary implications for developmental theory, future research, and clinical applications are discussed. Larger, longitudinal studies with typical and atypical populations and stricter methodological controls are needed to validate these findings.

Inquiries into the form (i.e., sound or vocal production characteristics) and function (i.e., social communication characteristics) of infant vocalizations have proceeded along mostly separate paths. Findings across three decades of research indicate that both form and function follow an orderly developmental progression during the prelinguistic period (Nathani, Ertmer, & Stark, 2006; Oller, 1980, 2000; Reddy, 1999; Stark, 1980; Stark, Bernstein, & Demorest, 1993). In the present study, we extend this body of work by examining the nature of the relationships between form and function in the prelinguistic and early linguistic period. If form and function are coupled together in infancy, it may suggest that we need to pay close attention in research and clinical practice not only to the forms themselves but also to the contexts in which these forms typically occur. On the other hand, if specific forms occur across a variety of circumstances, it may suggest that they need to be elicited and/or facilitated across several contexts in order to ensure their versatility. Although such careful attention to forms and functions would obviously be demanding, it may be important in helping children with communicative deficits develop typical relationships between various forms and functions.

Vocalization Forms in Infancy

Most researchers agree that typically developing infants appear to universally proceed through four or five developmental levels on their path from crying to the production of first words (for a review, see Oller, 2000). Researchers acknowledge, however, that some age-of-emergence estimates, operational definitions and classifications of vocalization forms, and terminology differ across models because of known individual variability and divergence in theoretical orientations. For example, Buder, Warlaumont, and...
Buder et al. might classify a vocalization as both a “canonical syllable” (well-formed, crisp adult-like consonant and vowel combination) to reflect the articulatory dimension of the vocalization and as a “squeal” (high-pitched sound) to reflect the phonatory dimension. On the other hand, although Nathani et al. (2006) acknowledge that a vocalization can be classified along multiple dimensions, they allow for only one classification per vocalization. Typically, the classification that is chosen reflects the dimension that is perceptually more salient and/or emerges later in development. Other prominent dimensions of the vocalizations are indicated via comments. Nathani et al. would therefore classify the previous vocalization as only a canonical syllable; the squeal-like phonatory dimension of the vocalization would be noted in a comment. These differences across models suggest that results obtained using any particular model should be cautiously interpreted.

For this preliminary effort, however, it was important to choose one model that had been shown to be comprehensive, internally consistent, and able to reliably detect differences in vocalization forms across ages. Nathani et al.’s (2006) model for the development of vocalization forms was chosen for the present study because it met these criteria and because the present study used archival data from Nathani et al. that had already been classified for vocalization forms. According to this model, there are five levels in the development of vocalization forms. At the first level, the infant produces crying, vegetative sounds, and quasi-resonant nuclei (primitive vowel-like sounds with muffled resonance). This level is expected to emerge approximately between 0 and 2 months of age. In the second level (~1–4 months of age), the infant produces fully resonant nuclei (vowel-like vocalizations that have full resonance but are not readily recognizable as adult-like vowels), chuckle or laughter, and adds some back and glottal consonants to his or her repertoire. They may also occasionally combine consonant-and vowel-like elements. In the third level (~3–8 months of age), mature and recognizable adult-like vowels, vocal play or expansion of the repertoire in the form of squeals, and marginal babbling (production of a series of consonant-like and vowel-like elements that sound “sloppy” or “slurred” to the adult listener) are seen. In the fourth level, which emerges approximately between 5 and 10 months of age, canonical syllables (well-formed, crisp adult-like consonant and vowel combination) and whispers are observed. The fifth level (~9–18 months of age) may overlap with the production of real words and is characterized by the production of complex canonical syllables (e.g., closed syllable shapes, clusters) and jargon (i.e., strings of canonical syllables with varying intonation and stress patterns) and well-formed diphthongs. The form of vocalizations at the fifth level may also be influenced by the language the infant is learning. Finally, infants within a given age group can produce vocalizations from more than one level, and infants across different age groups can produce vocalizations from the same level.

**Vocalization Functions in Infancy**

Several theoretical proposals have been made regarding the functions of infant vocalizations. Prelinguistic vocalizations were initially thought to be random and reflexive, with no meaning to the infant (Jakobson, 1962/1971, cited in Oller, 2000). Subsequently, it was proposed that infants used vocalizations to actively explore their articulatory motor and respiratory capabilities (Locke & Pearson, 1992). It was also suggested that infants’ vocalizations served as a source of self-stimulation and thereby promoted further vocal development (Locke & Pearson, 1992). Infant vocalizations were also proposed to convey information about infants’ affect and attention (Adamson, 1995). Still others proposed that infants’ vocalizations reflected their awareness of and information about the surrounding context (e.g., partner affect, objects) and therefore reflected an early form of social interaction and/or learning readiness (Adamson, 1995; Goldstein, Schwade, & Bornstein, 2009; Goldstein, Schwade, Briesch, & Syal, 2010; Hilke, 1988; Locke, 2002). Because adults often respond to infant vocalizations as if they were communicative signals, it was also suggested that infant vocalizations serve a social regulatory function or facilitate learning (Bloom, 1990; Goldstein et al., 2010). The range of theories suggests that various functions may be associated with prelinguistic vocal forms. Empirical support for these various functions has been provided by observing infants in various social contexts (e.g., playing alone, interacting with caregiver, still-face paradigm) and noting caregiver responses along with infant vocalization behaviors (e.g., Delack, 1978; Hsu & Fogel, 2001).

Unlike vocal form, a generally accepted developmental sequence for the functions of infant vocalizations has, however, not been established, in part because infants have been studied at only a few and discrete ages within the first year of life. The scheme proposed by Stark et al. (1993) to describe the development of vocalization functions was chosen for the present study because Stark et al. showed the scheme to be reliable and valid and because the present study used archival data from Stark et al. that had already been classified for vocalization functions. According to Stark et al., vocalizations that reflect biological states and are somewhat predictable—termed reflexive—are produced in the first function category and emerge approximately between 0 and 2 months of age. Vocalization types in this category include fussing, crying, and vegetative noises (e.g., burps). Next (~2–5 months of age), vocalizations are produced in face-to-face interaction with a caregiver or during visual regard of the environment or objects. These may occur with a neutral or positive facial expression on the part of the infant. In the third category (~6–9 months of age), vocalizations are produced as infants actively interact with objects (e.g., sounds made while banging objects). No interaction with an adult occurs when producing vocalizations in this category. In the fourth and highest category (~10–18 months of age), vocalizations appear to have communicative intent and are typically produced in social
interaction exchanges with objects and people (e.g., the infant looks at the adult for assistance and makes vocalizations indicating desire for or rejection of an object). Stark et al. divided this fourth category into four subcategories depending on whether the infant was producing vocalizations to request or reject objects or activities from the adult, to express emotional affect to the adult, to initiate or continue interaction with the adult, or to serve miscellaneous functions (e.g., making animal noises).

**Relationships Between Form and Function**

Some theoretical proposals and empirical results are available regarding relationships between form and function. Shatz (1983) proposed that communicative subsystems (e.g., form, function, cognition) initially develop independently of each other. Communicative growth occurs when these previously independent subsystems become integrated. Shatz reviewed the literature to show that children do not produce verbal requests right away even though they have the verbal means to do so. First, they request via non-verbal means, and only later do they begin to request using verbal means.

Oller (2000) proposed that although some prelinguistic speech-like vocalization forms (i.e., protophones) tend to occur more often in some contexts, infants can produce vocalization forms across a variety of social contexts very early in life. According to this proposal, the same protophone can occur across a variety of social contexts, and different protophones can occur within the same social context or even without any apparent social functions. For example, a 6-month-old may produce squeals when mouthing or banging objects, when visually staring at lights in the room, or when interacting with his or her caregiver. Oller referred to the flexibility in production of protophone vocalizations across various social contexts as contextual freedom (i.e., vocalizations are not tied to specific social functions) and hypothesized that these associations may change with age. He further argued that this contextual freedom of protophones is a critical building block for later symbolic language.

Recently, some support for the contextual freedom proposal has been obtained by Oller et al. (2013) and Hsu, Iyer, and Fogel (2014). Oller et al. showed that squeals, growls, and vowels were produced by infants throughout the first year of life to serve a variety of functions (e.g., continuing of “conversation” with caregiver, complaining) and were not restricted to specific functions. Similarly, Hsu et al. showed that identical vocalization types were produced in more than one context by 6- and 12-month-old infants (e.g., fully resonant nuclei were produced across two different games of tickle and peek-a-boo).

To date, however, studies investigating vocalization form-function relationships in infancy are scarce (D’Odoni & Franco, 1991; Hsu et al., 2014; Hsu & Fogel, 2001; Oller et al., 2013; Papaeliou, Minadakis, & Cavouras, 2002; Papoušek & Papoušek, 1989). In addition, there are three main difficulties in interpreting the findings from this literature. First, the available evidence has often categorized form in overly broad terms (e.g., a dichotomous syllabic-vocalic classification in which syllabic refers to all sounds that are resonant and are produced toward the front of the mouth and vocalic refers to all sounds that are less resonant and produced toward the back of the mouth; Hsu & Fogel, 2001). This broad dichotomy makes it difficult to map these sound types onto more detailed schemes for classifying form, such as the one used in the present study. Second, form-function relationships have often been examined in very restricted circumstances (e.g., only vocalizations that were produced in interaction with the caregiver were analyzed; other circumstances, such as playing alone, were not analyzed; Papaeliou et al., 2002). Third, this relationship has been investigated at only a few and discrete ages within the first year of life (e.g., 6 and 12 months; Hsu et al., 2014). Consequently, it is difficult to draw any definite conclusions regarding the nature, range, and development of form-function relationships. However, results from available investigations suggest that potentially both coupling and decoupling associations might exist between form and function in infancy and that shifts in form-function relationships may partly be a result of age. For example, Hsu and Fogel (2001) noted that infants ages 0;1–0;6 [years;months] produced more syllabic vocalizations when mutually active and engaged with their caregiver (symmetrical communication) than when only one partner was active (asymmetrical) or engaged (unilateral). At age 0;6, however, the number of syllabic vocalizations became similar across both symmetrical and unilateral conditions, suggesting a decoupling of form and function.

**Goals of and Motivation for the Present Study**

In this study, we further assessed the nature of form-function relationships in infancy by examining a broad range of infant vocalizations occurring across a wide variety of naturally occurring contexts in infants’ homes. An analysis of different vocal types across different circumstances in infancy should reveal whether there are any restrictions in usage of particular forms (Shatz, 1983) or whether they are free to vary across several circumstances (Oller, 2000). The present study was also the first, to our knowledge, to examine infants across the first 20 months of life within a variety of naturally occurring circumstances, unlike previous research, which has examined few and discrete ages only within the first year of life and often in the laboratory.

In addition to increasing understanding of the typical development of spoken communication in infants and toddlers, the findings of this study may have clinical implications for children at risk for communicative deficits. If associations between form and function are documented in the present study along with developmental changes in these relationships, the findings could be used to assess whether children with communicative deficits have delays and/or atypical patterns of association. It may be that children with developmental delays persist in using early-emerging vocalization forms for higher level functions well after ages at
which typically developing children make the shift to using higher level forms for higher level functions. For example, Roberts et al. (2007) showed that although older boys with Down syndrome (M = 9.4 years) demonstrated more adequate topic maintenance than mental-age-matched typically developing boys, they used less elaborate forms than the younger, typically developing boys. Other children may show atypical patterns of form–function relationships, such as the usage of higher level forms for lower level functions. For example, Tager-Flusberg (1999) described the use of negation and questions by children with autism. Although children with autism followed typical patterns of negation form development and were mastering advanced sentential negation forms, they used these utterances primarily to express early-developing negation functions of rejection and nonexistence. They rarely used negation forms for denial, a later-developing function. Similarly, questions were primarily used to express early-emerging functions (e.g., requests) rather than more advanced information-seeking questions. New insights into form–function relationships also have the potential to enhance clinical intervention practices by offering a continuum of vocal forms and appropriate, associated functions. With this knowledge, clinicians can apply language stimulation techniques (e.g., milieu teaching), such that the social functions associated with various forms are permitted to naturalistically occur, thereby taking advantage of our knowledge regarding form–function correspondences.

The vocal form and function types used in the present study have been objectively and operationally defined in two prior studies and shown to be both valid and reliable (Nathani et al., 2006; Stark et al., 1993). The first study using this database (Stark et al., 1993) observed the functions of vocalizations produced by 51 typically developing infants in the first 20 months of life. Using the same database, the second study (Nathani et al., 2006) evaluated the form of the vocalizations produced by a randomly selected subset of 30 infants whose data were still available. Functions of vocalizations were not known to the judges in the Nathani et al. (2006) study, and vocalization forms were unknown to the judges in the Stark et al. (1993) study. Results from the two studies showed that infants within an age group produced a range of forms and functions. In addition, most infants showed a fairly orderly developmental progression in the forms and functions they used in the first 20 months of life.

The present study sought to extend this work by providing a preliminary perspective on two questions of fundamental interest. First, do associations exist between the vocalization forms and vocalization functions that had been identified by Nathani et al. (2006) and Stark et al. (1993), respectively? If so, do these form–function relationships change with development? We hypothesized that early-emerging forms would be associated more often with early-emerging function categories and later-emerging forms with later-emerging function categories, regardless of age, because of the known disparate frequency in occurrence of these forms and functions in infants of different ages.

### Method

#### Participants

Twenty infants, whose vocalization forms had been classified in Nathani et al.’s (2006) investigation, were used in the present study. The 20 participants included five different infants in each of four age groups: (a) 12–22 weeks (0;3–0;5), (b) 26–36 weeks (0;6–0;8), (c) 40–55 weeks (0;9–1;1), and (d) 72–88 weeks (1;4–1;8). These 20 infants had been randomly selected by Nathani et al. from a larger archival pool of infants whose vocalization functions had been examined by Stark et al. (1993). Data from an additional five infants in the 1;1–1;3 age group analyzed by Nathani et al. could not be examined in the present study because vocalization functions had not been coded by Stark et al. (1993) due to missing videotapes for these infants. Gender had been balanced across all age groups, except for the 0;6–0;8 age group, which had more boys than girls. Hearing and developmental status had been ascertained to be within normal limits for all participants (Stark et al., 1993). Some of the participant information from the Stark et al. database was unavailable. For example, the reason for unequal age spans from 3 months in the two younger age groups to 5 months in the two older age groups had not been stated. Similarly, although it had been stated that the race and socioeconomic status of the infants were proportionally sampled, specific information about these participant characteristics were not provided by Stark et al. Despite these shortcomings, the Stark et al. database provides a preliminary opportunity to cross-sectionally examine relationships between forms and functions in the vocalizations produced by 20 infants across a variety of natural circumstances within children’s own homes during prelinguistic and early linguistic development.

#### Data Collection

All 20 infants included for the present study had been investigated using a mixed cross-sectional and longitudinal research design. Each infant had been audio- and video-recorded in their homes five times within the age range for each age group. A standard protocol had been followed for each of the five recording sessions that included 10–20 min of mother–infant interaction, 10–20 min of solitary play, and a 5-min feeding episode, for a total of approximately 40 min per session. The objective of these activities was to elicit a range of vocalization forms and functions. Data from Session 1 and Session 5 had been coded for form and function by Nathani et al. (2006) and Stark et al. (1993), respectively. For this study, form and function data from Session 1 and Session 5 were combined for each infant in order to provide a more stable and representative estimate of their vocal behaviors.

#### Procedure for Classifying Infant Vocalizations

In the original database, an utterance had been operationally defined as a vocalization or series of vocalizations...
subcategories had primarily been produced by only infants aged 1.4–1.8, thereby resulting in limited data within each subcategory, these subcategories were not separately analyzed in the present study. Instead, the Communicative category was analyzed as a whole. A brief description of the categories included in Stark et al.’s study and associated vocalization types are provided in Appendix B. Complete information about the categories, vocalization types within each category, and subcategories within the highest category can be found in Stark et al.

Cross-classification tabulations. Forms and functions derived from Nathani et al. (2006) and Stark et al. (1993), respectively, were cross-tabulated for each utterance. The number of utterances at each combination of form and function (e.g., utterances that received both a Control of Phonation form level and a Reactive function category) were added up for each infant. These combinations were then divided by the total number of utterances for that particular form to obtain proportions of utterances at each function for that form. For example, a total of 461 Control of Phonation form level utterances were produced across all 20 infants. Of these 461 utterances, 39 were produced with a Reflexive function category designation. Thus, approximately 8% of all Control of Phonation form level utterances across all infants were used for Reflexive functions. Because the number of vocalizations for each form level and function category combination was different, the proportions at each combination were weighted (or multiplied) by the total number of vocalizations that infants produced for that particular combination.

Reliability

Approximately 10% of the data had been analyzed for interjudge agreement on assignment of form levels in Nathani et al.’s (2006) study. Cohen’s κ had been found to be .64. Similarly, interjudge agreement on approximately 10% of the data for function category assignments had been found to range from 76% to 89% across two observers in Stark et al.’s (1993) study.

Statistical Analysis

The data in the present study were analyzed using a generalized linear mixed-effects model of the multinomial regression type with SAS procedure PROC GLIMMIX (Fienberg, 1994). This procedure is used to model response variables that are not continuous. Because form and function were ordered categorical variables, this model was chosen for analysis. Form was considered the predictor variable, and function was considered the response variable. Infants were included as random effects. Because the multinomial-based model explicitly does not assume normality, it was able to handle nonnormally distributed proportional data precisely of the sort used here. To examine the strength of the associations between form and functions, Cramer’s coefficients (Cramer’s V), a type of correlation measure for ordered categorical variables, were also
computed (Conover, 1999). Values for Cramer’s coefficients can range from 0 to 1, with 1 indicating perfect association and 0 indicating no association.

The two younger (0;3–0;5 and 0;6–0;8) and the two older (0;9–1;1 and 1;4–1;8) age groups were consolidated because the sample size of five infants per age group was too small for any meaningful statistical analysis. This consolidation meant that the age range for the younger age group spanned 6 months and the range for the older age group spanned 13 months. The consolidated younger and older age groups were analyzed separately from each other because not all age groups produced vocalizations from all form levels and function categories. The younger age groups rarely produced form Levels 4 (Basic Canonical Syllables) and 5 (Advanced Forms) and function Category 4 (Communicative). Thus, there was essentially very little variance on those levels and categories, and they could not be included in the analysis. Only form Levels 1–3 (Reflexive, Control of Phonation, and Expansion) and function Categories 1–3 (Reflexive, Reactive, and Activity) could be included in the multinomial regression analysis and Cramer’s coefficient computations with younger infants. On the other hand, older infants (0;9–1;1 and 1;4–1;8) produced all form levels and function categories. Therefore, all four function categories (Reflexive, Reactive, Activity, and Communicative) and all five form levels (Reflexive, Control of Phonation, Expansion, Basic Canonical Syllables, and Advanced Forms) were included in the multinomial regression analysis and Cramer’s coefficient computations for the older age groups.

In an effort to determine whether there was some consistency across the results between younger and older age groups, Cramer’s coefficients were also computed wherever there was some overlap in the data between the younger and older age groups. If the values of Cramer’s coefficients were similar when data from the younger age group were compared to the entire data set for overlapping levels and categories, it would suggest at least some consistency across age groups even though different groups of infants were included because of the cross-sectional design of the study.

Results

A total of 2,308 utterances for the 20 infants were analyzed for associations between form and function. Figure 1 shows the proportions of utterances across the four categories of function (Reflexive, Reactive, Activity, and Communicative) at each of the five form levels (Reflexive, Control of Phonation, Expansion, Basic Canonical Syllables, and Advanced Forms) for all infants across all age groups. A 0.2 criterion, as used by Oller and Eilers (1988), was informally used to identify form levels that significantly co-occurred with function categories. Oller and Eilers chose this criterion because parents were reliably able to identify the emergence of canonical babbling at this rate of production. It was expected that associations between forms and functions at this criterion would be similarly notable to observers.

As seen in Figure 1, the Reflexive form level was predominantly associated with the Reflexive function category. The Reflexive form level was also associated, albeit to a lesser extent, with the Reactive function category. The Control of Phonation form level was most frequently associated with the Reactive function category. The Expansion form level was frequently associated with all but the Reflexive category of function. The later-emerging Basic Canonical Syllables and Advanced Forms levels appeared to be most often associated with the later-emerging Communicative function category.

Form–Function Relationships in the Younger Age Groups (0;3–0;5 and 0;6–0;8)

The results showed a significant association between form and function, $F(2, 70) = 2,732.59, p < .001$, in the younger age groups. The strength of the form–function association was also significant, Cramer’s $V = .363$, $\chi^2 (4, N = 10) = 278.48, p < .001$. Recall that for statistical analysis, the two younger age groups (0;3–0;5 and 0;6–0;8) had to be consolidated and that only the first three out of the five form levels and the first three out of the four function categories could be included for these two age groups. Analysis of the means revealed that early-emerging form levels were more often associated with early-emerging than later-emerging function categories, and vice versa. Specific age effects could not be identified because of small sample sizes within each of the two age groups that were combined for this analysis.

Figure 2 shows the proportions of utterances across the three categories of function (Reflexive, Reactive, and Activity) at the three form levels (Reflexive, Control of Phonation, and Expansion) analyzed for the two younger age groups (0;3–0;5 and 0;6–0;8). Again, using the informal 0.2 criterion, some trends can be noted. The Reflexive form level was most often associated with the Reflexive category of function. The Reflexive form was also associated, to a somewhat lesser extent, with the Reactive function category. The Control of Phonation form level was most often associated with the Reactive category of function. The Expansion form level was most frequently associated with the Reactive function category.

Form–Function Relationships in the Older Age Groups (0;9–1;1 and 1;4–1;8)

Results from the statistical analysis for the older age groups (0;9–1;1 and 1;4–1;8) also revealed a significant association between form and function, $F(4, 133) = 240.73, p < .001$. In addition, although not as strong as that seen for the younger age groups, the strength of the association between form and function was significant for the older age groups as well, Cramer’s $V = .251$, $\chi^2 (12, N = 10) = 221.49, p < .001$. Analysis of the means revealed a pattern similar to that found for the younger age groups: Early-emerging form levels were more often associated with early-emerging...
than later-emerging function categories, and vice versa. Specific age effects could not be identified again because of small sample sizes within each of the two age groups combined for this analysis.

Figure 3 shows the proportions of utterances across the four categories of function (Reflexive, Reactive, Activity, and Communicative) at the five form levels (Reflexive, Control of Phonation, Expansion, Canonical, and Advanced) for the two older age groups (0;9–1;1 and 1;4–1;8). Again, a 0.2 criterion was informally used to identify trends in the data. As is evident in Figure 3, Communicative functions were most prominent in these older infants. However, Communicative functions were not restricted to a specific form level; they were instead associated with all form levels. Early-emerging form levels (Reflexive and Control of Phonation) were also associated with several function categories,
unlike that seen for younger infants. Similar to that seen with younger infants, Expansion level vocalizations continued to be associated with several function categories. The two later-emerging form levels, namely Basic Canonical Syllables and Advanced Forms, were predominantly associated with Communicative functions.

Because older infants also produced form levels and function categories 1–3 similar to younger infants, it was possible to compare the combined data from both younger and older infants for these levels and categories to the data obtained when only the younger age groups were included. The strength of the association between form levels and function categories 1–3 for all age groups together, Cramer’s $V = .341$, $\chi^2(4, N = 20) = 350.39$, $p < .001$, was similar to that previously reported for the younger infants alone, Cramer’s $V = .363$, $\chi^2(4, N = 10) = 278.48$, $p < .001$, suggesting some consistency to the results.

Discussion

The results indicate that form and function are significantly associated throughout infancy. Earlier-emerging form levels (Reflexive and Control of Phonation) were often associated with earlier-emerging function categories (Reflexive and Reactive), and later-emerging form levels (Basic Canonical Syllables and Advanced Forms) were often associated with later-emerging function categories (Communicative). These trends suggest that form and function co-emerge in infancy. Thus, infants appear to start by using simple forms to express simple functions. As infants begin to use more complex forms, they become more communicative in function and appear to demand increasingly elaborate responses from the caregiver.

It may be that age is the common underlying substrate that unites the two domains. Because younger infants (<0;9) primarily produced vocalizations from Levels 1–3 for form and Categories 1–3 for function, linkages between early-emerging levels and early-emerging categories were exceedingly likely. Similarly, older infants (>0;9) frequently produced vocalizations from later-emerging form levels and function categories, leading to a greater probability of linkages between these form levels and function categories. Given the cross-sectional research design of the current study, however, it is possible that differences in the frequencies of form levels and function categories between younger and older infants were obtained because different infants were included in the various age groups and not because of age effects. This concern is somewhat mitigated by the fact that the strength of the association between form Levels 1–3 and function Categories 1–3 for the younger infants was consistent for the entire data set when both younger and older infants were included, even though different infants comprised the younger and older age groups. Of course, these conclusions would need to be substantiated by larger, longitudinal studies that can directly examine the effects of age on form–function relationships.

It may also be that caregiver response, rather than age, is the link between form and function. When an infant produces advanced forms, the caregiver may interpret them as inherently communicative in function because they resemble the sounds of mature speech. Indeed, Masataka and Bloom (1994) and Gros-Louis, West, Goldstein, and King (2006) have shown that caregivers differentiated
their responses based on the form of infants’ vocalizations. Masataka and Bloom noted that mothers of infants age 0;3 selectively responded to syllabic infant vocalizations (speech-like vocalizations that were resonant and produced toward the front of the mouth) and ignored vocalic vocalizations (sounds produced toward the back of the mouth). Gros-Luis et al. (2006) observed that six of 10 mothers differentiated their responses based on the infant’s productions. They responded with play vocalizations to vowel-like sounds and with imitations to consonant–vowel sounds. In the present study, because function category assignments were partially based on caregiver responses, perhaps they were biased by the form of infants’ vocalizations, such that vocalizations containing later-emerging forms were assigned later-emerging functions and early-emerging forms were assigned early-emerging functions due to the caregiver’s differential responses to various forms. Furthermore, although Stark et al. (1993) emphasized that their function category judgments were based on facial expression, gesture, or other movements, it is possible that judges’ access to the auditory characteristics (form) of the vocalizations may have inadvertently biased their judgments. It would be useful for future studies to conduct function judgments with infant vocalizations rendered inaudible.

It should also be noted that there were some overlaps in the operational definitions within the form and function coding schemes, especially for the Reflexive form level and Reflexive function category. Not only were the names of this category/level identical, but also the elements included within them were defined almost identically. Thus, the observed strong association between the Reflexive form level and the Reflexive function category may have simply been an artifact of the coding schemes rather than a true association. However, given the limited diversity of child vocalization forms and functions, especially at early ages, such commonalities appear to be natural phenomena rather than solely a by-product of analysis.

All form levels and function categories did not appear, however, to be tightly coupled. For example, although not statistically verified, Expansion level vocalizations appeared to be associated with three different function categories (Reactive, Activity, and Communicative). This multifunctional nature of Expansion level vocalizations is in alignment with the proposal made by Oller (2000) that similar speech-like sounds can be produced by infants across a variety of social contexts (i.e., contextual freedom). Advances in function similarly did not also always coincide with advances in form. Although infants ages 0;9–1;1 and 1;4–1;8 primarily used Expansion, Basic Canonical Syllables, and Advanced Forms to express the most advanced Communicative functions, they also used vocalizations from earlier-emerging form levels of Reflexive and Control of Phonation to express Communicative functions.

McCune, Vihman, Roug-Hellichius, Delery, and Gogate (1996) observed a similar phenomenon with grunts. Grunts were found to index physiological function early in the first year of life but then also served communicative functions at the transition to speech (McCune et al., 1996).

Grunts, as defined by McCune et al., were characterized by sudden glottal onset, absence of supraglottal closure (except for lip closure), and short duration. Although it is difficult to be absolutely certain, such grunts would probably be characterized as quasi-resonant nuclei included under the Reflexive level in our scheme. Oller (2000) similarly concluded that grunts would be classified as quasi-resonant nuclei under his scheme. Taken together, McCune et al.’s results and our preliminary results suggest that infants may be versatile in their use of some forms and functions.

### Overall Study Limitations

As previously reviewed, there are differences across researchers in classifying form and function. Therefore, although difficult to ascertain, different or more nuanced results might have been obtained if other classification schemes for form and function had been used (e.g., Buder et al., 2013, Hsu & Fogel, 2001). For example, some vocalizations that were classified as Communicative in the present study could have been subclassified into symmetrical (both mother and infant mutually engaged), asymmetrical (mother is actively engaged but infant is passively engaged), or unilateral (mother is actively engaged but infant is disengaged and observing something else) by Hsu and Fogel (2001).

In the present study, we used classification schemes that operationally defined vocalization types and that had previously been shown to be both reliable and valid (Nathani et al., 2006; Stark et al., 1993). In addition, the schemes used in the present study allowed us to classify a range of vocalization forms and function across the entire 20 months of life. But future investigations could use different schemes to ascertain whether and how the choice of classification schemes influences the detection of form–function relationships. Similarly, with regard to analytical frameworks, a multinomial regression model was used to analyze the data in the present study because of the ordered nature of the form levels and function categories. Other analytic frameworks, such as log linear analysis or odds ratios with modifications for ordered data, might yield new and different insights into the nature of form–function relationships (Bakeman & Quera, 2011).

Furthermore, there were some issues with using an archival database. Even though most functions categorized by Stark et al. (1993) are consistent with most prelinguistic functions that are currently discussed in the literature, there are some discrepancies. For example, Stark et al. assigned separate function categories to sounds made while gazing at objects versus actively interacting with objects. However, recent reports in the literature assigning different communicative functions to these two contexts could not be found. It may be that future investigations do not need to separate out these two contexts. Similarly, a breath-group criterion (Oller & Lynch, 1992) is typically used to separate utterances in contemporary investigations rather than the 2-s rule used by Stark et al. A larger sample of utterances would have been available if the breath-group rule, rather than the 2-s rule, which could consolidate several vocalizations into...
the same utterance, had been used to separate out utterances. Post hoc analysis, however, indicated that even with this more encompassing 2-s rule, it was possible to identify short and unitary utterance types; 13% of all utterances were isolated quasi-resonant nuclei. Although utterance durations were not available for the present study, single quasi-resonant nuclei are presumably the shortest unitary utterance type based on our operational definitions.

Finally, the cross-sectional research design of the current study made it difficult to be certain whether observed differences between younger and older infants were a result of age effects or were because different infants were included in the various age groups. Associations of specific forms and functions included within each form level and function category also could not be obtained because of small sample sizes. For example, squeals and marginal syllables were included under Expansion level vocalizations because they signaled maturity of the vocal tract. However, these two vocalizations are very disparate in form, and it is possible that they serve entirely different functions. A larger, longitudinal study would improve detection of more subtle age-related changes in form–function relationships, especially given known variability in infant vocalization behaviors at these ages.

**Theoretical and Clinical Implications**

The findings from this preliminary study suggest that complex relationships appear to exist between forms and functions across the first 3–20 months of life. Whereas some forms appear to be tightly coupled to functions, other forms appear to be free to vary across contexts, and vice versa. Thus, different forms and functions were associated in different ways. Early-emerging forms (e.g., Control of Phonation forms) were more often associated with early-emerging functions (e.g., Reactive functions), and later-emerging forms (e.g., Advanced Forms) were more often associated with later-emerging functions (e.g., Communicative). Expansion level forms appeared to be, however, associated with multiple functions. Thus, our findings provide new perspectives on theoretical proposals made by Shatz (1983) and Oller (2000) regarding form–function relationships in infancy.

Our results also suggest several preliminary implications for future research on form–function relationships and clinical assessment and intervention practices. Clinical implications from the present study should, however, be considered very cautiously given the small sample size and because only typically developing children were observed. In particular, the results suggest that forms should be examined together with their associated functional domains. Although causal relationships cannot be inferred from mere associations, the results show that certain vocalization forms tend to occur more often in particular contexts. If coupling relationships between forms and functions are verified in future investigations, the likelihood of eliciting a particular form may be increased (or decreased) depending on the context. Unusual coupling of forms with functions might also turn out to be an early diagnostic indicator of language learning difficulties. Furthermore, assuming future verification for forms and functions that are not as tightly coupled, it might be prudent to examine these forms in more than one context to ensure that the child is not restricted in his or her usage or form repertoire. Consideration of form–function relationships is also likely to have an important role in intervention planning. Facilitating the emergence of a form without regard for function may not be an efficient approach to early intervention. Examining the validity of these preliminary results and their implications will, however, require several additional studies of both typically developing children and children with communicative delays and deficits.

The findings from this preliminary study suggest that both coupling and decoupling relationships exist between forms and functions in infancy. This area of investigation is clearly worthy of increased attention because of its potential for advancing both theoretical and clinical knowledge.

**Acknowledgments**

 Portions of this work were presented at the 2004 International Conference on Infant Studies, Chicago, IL. We gratefully acknowledge the contributions of L. E. Bernstein and M. E. Demorest in investigating the concept of vocalization function in an earlier article (Stark et al., 1993). We are also indebted to Mallika Raman for her assistance with data coding and Nicole Lazar for her assistance with statistical analysis. Finally, these studies would not have been possible without the visionary work of our late mentor, Rachel E. Stark.

**References**


### Appendix A

#### Levels of Vocalizations Forms and Associated Vocalization Types

<table>
<thead>
<tr>
<th>Levels</th>
<th>Examples of vocalization forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reflexive</td>
<td>Fixed signals (e.g., crying); vegetative sounds (e.g., burps); and a protophone, quasi-resonant nuclei. Quasi-resonant nuclei are primitive vowel-like brief sounds with muffled resonance and contain energy mostly restricted to low frequencies.</td>
</tr>
<tr>
<td>2. Control of Phonation</td>
<td>Fully resonant nuclei protophones, which are vowel-like sounds with greater resonance than quasi-resonant nuclei but may still contain features (e.g., harshness) that make them difficult to be recognized as specific adult vowels. Other protophone types include primitive consonant-like sounds (e.g., clicks, raspberries, goos) and primitive consonant- and vowel-like combinations. A fixed signal, namely chuckle and laughter, is also included at this level.</td>
</tr>
<tr>
<td>3. Expansion</td>
<td>Vowel-like sounds that are fully recognizable as mature adult-like vowels (e.g., [a]), squeals, and marginal babbling. Marginal babbling refers to a series of primitive consonant- and vowel-like combinations with slow formant transitions from the consonant-like to the vowel-like element.</td>
</tr>
<tr>
<td>4. Basic Canonical Syllables</td>
<td>Canonical syllables, which are consonant–vowel combinations with rapid formant transitions between the consonant and the vowel, making them sound adult-like (e.g., [ba], [ninini]; Oller, 1980, 2000). May be produced as a singleton, as disyllables, or as a series. This level also includes whispers.</td>
</tr>
<tr>
<td>5. Advanced Forms</td>
<td>Closed syllables (e.g., VC, CVC), diphthongs, and jargon (i.e., strings of canonical syllables with varying intonation and stress patterns). Protophones at this level are considerably more complex in their articulatory and phonatory characteristics than protophones at any other level.</td>
</tr>
</tbody>
</table>


### Appendix B

#### Categories of Vocalization Function and Associated Vocalization Types

<table>
<thead>
<tr>
<th>Categories</th>
<th>Examples of vocalization types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reflexive Sound Making</td>
<td>Vegetative sounds, fussing and crying sounds, grunts made during movements, and sighs.</td>
</tr>
<tr>
<td>2. Reactive Sound Making</td>
<td>Vocalizations produced during face-to-face caregiver–infant interaction and vocalizations produced during prolonged visual regard of objects (e.g., lights in a room), with a positive or neutral facial expression.</td>
</tr>
<tr>
<td>3. Activity or Object Sound Making</td>
<td>Vocalizations produced as infants actively explore the environment and objects in it but without apparent communicative intent (e.g., sounds made when crawling, banging, or shaking objects).</td>
</tr>
<tr>
<td>4. Communicative Intent Sound Making</td>
<td>Vocalizations made when the infant appears to be intentionally communicating with caregivers, for example, using vocalizations to draw attention to objects or activities, greetings, or naming an object.</td>
</tr>
</tbody>
</table>