

Discourse and Intonation Development in the First-Word Period

Heather L. BALOG*, Felicia D. ROBERTS* et David SNOW*

ABSTRACT

The purpose of the current study was to examine the vocal productions of young children acquiring Standard American English and to determine whether intonation production is influenced by discourse context and developmental level (*i.e.*, age). Participants were 24 children with typically developing speech and language, ages 12 to 23 months. A cross-sectional design was used. Data were collected from language samples during 30- to 40-minute play sessions. Child interactions were analyzed according to three discourse categories: *co-participatory*, *initiation*, and *narrowed focus*. Intonation was analyzed using measures of accent range, contour inventory size and contour maturity. The results supported a previous finding that the contour inventory measure of intonation is consistently more useful in demonstrating differences in production related to contour direction and discourse context. An analysis of the contour maturity of intonation revealed that falling contours were more stable than rises, and that rises were produced with more stability in communicative contexts.

KEY-WORDS: INTONATION DEVELOPMENT, DISCOURSE DEVELOPMENT, PRODUCTION, TODDLERS

* Heather L. Balog, Ph.D., CCC-SLP, Department of Communication Sciences and Disorders, Wayne State University, 207 Rackham Building, 60 Farnsworth, Detroit, MI 48202, E-mail: hbalog@wayne.edu

RÉSUMÉ

Développement discursif et intonatif pendant la période des premiers mots

Le but de la présente étude était d'examiner la production vocale de jeunes enfants en cours d'acquisition de l'anglais américain standard afin de déterminer si les contours intonatifs réalisés étaient influencés par le contexte discursif et le niveau de développement (i.e. l'âge et le stade lexical). 24 enfants, âgés de 12 à 23 mois et présentant un développement typique de la parole et du langage constituent la population étudiée. Il s'agit d'une approche transversale. Les données ont été recueillies à partir d'échantillons de parole, durant 30 à 40 minutes, produite lors de sessions de jeux. Les interactions des enfants ont été analysées en fonction de trois catégories de discours: la co-participation, l'initiative et les focalisations. L'intonation a été analysée en mesurant les accents et la taille de l'inventaire des contours.

Les résultats ont fourni des premiers constats: la mesure de l'éventail des contours intonatifs semble plus adaptée pour faire apparaître les différences en production que celle de l'orientation des contours et indique que cette mesure a été pertinente pour déterminer les différences de production dans les différents contextes de discours. Une analyse complémentaire révèle également que les contours descendants sont plus stables que les montants, et que les contours montants sont produits avec plus de constance selon les contextes de communication.

MOTS-CLÉS : DÉVELOPPEMENT DE L'INTONATION, ÉMERGENCE DU DISCOURS, PRODUCTION, JEUNES ENFANTS

INTRODUCTION

The acquisition of language in the second year of life (*i.e.*, 12 to 23 months) is characterized by emerging skills in all areas of language. These areas include: form (*i.e.*, phonology, morphology, syntax), content (*i.e.*, semantics), and use (*i.e.*, pragmatics). While much is known about the more obvious aspects of language development (*e.g.*, sound and word acquisition, grammar, etc.), less is known about some of the more subtle aspects of early language development. Two such areas are prosody and pragmatics.

Prosody encompasses the suprasegmental phonological aspects of speech, which include: pitch, length and loudness. Each of these characteristics is carried over multiple speech segments. Intonation (*i.e.*, changes in pitch and length) is used to mark pragmatic functions of utterances. Pragmatic language encompasses a wide variety of language use features, some of which are related to intonation (*e.g.*, communicative intention or discourse context).

The primary focus of this study was the exploration of intonation production in very young language learners and its relationship to emerging discourse skills. The intent of this study was to utilize adult judgments of discourse context as a means for exploring variation in intonation production during the second year of life. Discourse context and intonation are related in that they both reveal important pragmatic information during communication. Discourse context can help define the relationship between a child's utterance and the preceding adult utterance. Intonation is also related to pragmatics in that contour direction (*i.e.*, falling or rising) may help indicate intentions such as statements, commands, requests, or questions (Cruttenden, 1997). Given that discourse and intonation share a pragmatic role, it is plausible that exploring the relationship between the two will provide new insight into the developing language system.

Discourse context

Bloom, Rocissano and Hood (1976) reported on the early role of discourse context. Their study focused on the relationship between a child's conversational turn-taking and the preceding utterance of an adult conversational partner in children 19 to 36 months. They found that children produced utterances with varying grammatical complexity depending on the surrounding discourse context. Balog and Roberts (2004) adapted Bloom et al.'s discourse descriptions for use with younger children. Their goal was to broaden the discourse categories so that both non-lexical and lexical productions could be categorized. Balog and Roberts categorized utterances into 3 discourse contexts: *co-participatory* (which included child utterances that were part of the general discourse interaction and which were reliably judged by adult raters to relate to the previous adult utterance), *initiation* (which included utterances that were truly spontaneous or those in which a child did not maintain the topic of conversation), and *narrowed focus* (which included utterances in which the child seemed to be vocalizing privately). The narrowed focus context was considered non-communicative, in

the sense of self-directed, whereas the co-participatory and initiation contexts were considered interpersonally directed and communicative.

Intonation development

Intonation is one of the earliest aspects of phonology to develop. For example, children were found to use different pitch directions between 4 and 8 months to distinguish early cries (D'Odorico, 1984). In that study, cries labeled as indicating request or call functions were found to have rising pitch, compared to those categorized as discomfort cries, which were found to have falling pitch. Studies of older toddlers determined that, in general, falls are used for intentions like labeling or commenting and rises are used for intentions in which there is less certainty (*e.g.*, requests) (Snow & Balog, 2002).

One approach to measuring intonation in young children is that of nuclear tone theory (Cruttenden, 1997), which stems from the British tradition of autosegmental phonology. This approach has been most widely used with developmental research, similar to that of the current study and is easily applied to toddler productions (*i.e.*, babbling and monosyllabic utterances). Finally, nuclear tone theory is distinct from other autosegmental theories (*e.g.*, ToBI, Beckman and Elam, 1997) in that it distinguishes pre-nuclear and nuclear tones, a feature which is important for the current data. Measurement of intonation in the nuclear tone is so important because it is at that point in an utterance where pragmatic information is communicated.

One of the primary tenets of the nuclear tone theory is that intonation is analyzed on the *nuclear tone* of each utterance. The nuclear tone begins on the *nuclear accent*, which is the primary stressed syllable of an *intonation group*. Together, the nuclear accent and the syllables that follow to the end of the tone group make up the nuclear tone. Nuclear tones are described in terms of three tonal features (Cruttenden, 1997): 1. direction (*i.e.*, falling or rising), 2. accent range (*i.e.*, amount of pitch change), and 3. complexity (*i.e.*, small changes in direction that create contours such as *rise-fall-rise*).

Recent work has utilized careful measures of intonation to demonstrate developmental patterns and the relationships between intonation and emerging pragmatic skills (Balog & Snow, 2007; Balog & Brentari, 2008). Balog & Snow used both measures of accent range and contour inventory to define intonation development in toddlers. Balog & Brentari used similar measures to determine how intonation contours matched behavioral patterns in early communication. Following the principles of the nuclear tone theory, those studies measured intonation using accent range (*i.e.*, the change in pitch), contour inventory (*i.e.*, the number of contours produced), and contour maturity (*i.e.*, whether the child produced adultlike contour shapes and pitch change).

Summary of study rationale

In the current study, intonation was analyzed in varying discourse contexts using measures of accent range, contour inventory and contour maturity. Following Balog and Snow (2007), it was hypothesized that the contour inventory and

maturity measures would demonstrate differential use of intonation in various discourse contexts, whereas measures of accent range would fail to highlight these differences.

A second hypothesis of this study was that there would be variability across age and, more importantly, an interaction between the effects of discourse context and age. It was expected that the influence of discourse context on intonation would change as children developed more sophisticated language skills. Support for this hypothesis came from previous reports of intonation changes occurring across two distinct age groups (*i.e.*, before and after 18 months) (Snow & Balog, 2002). This study sought to determine whether or not the effects of discourse context on intonation production changed during the developmental period from 12 to 23 months. Finally, it was hypothesized that measuring intonation in terms of contour maturity would demonstrate more stable (*i.e.*, more mature contours) intonation production for falling compared to rising contours and that children would use more high and medium contours in their productions associated with communicative discourse contexts compared to less communicative contexts.

METHOD

Participants

Children ($n = 24$; 12 to 23 months old) and their mothers participated in the current study. Children were included if they were developing normally, passed a hearing screening, and were being raised in a monolingual American English home. Once the above criteria were met, the children were categorized into two age groups.

Procedures

The children and their mothers participated in a naturalistic play setting as described in earlier work (*e.g.*, Balog & Snow, 2007). Audio recordings were made using TELEX FMR-70 wireless microphones and a Marantz PMD430 stereo cassette recorder. A video recording of each session was made using a Sony Digital Handycam.

Phonetic transcription

Speech-like utterances (defined as having at least one voiced syllabic element and an outward breath stream) from all 24 children were transcribed phonetically. Inter-rater reliability for transcription was completed on 10% of each child sample for utterance boundary (reliability of 92%), transcription of consonants (reliability of 72%), stress placement (reliability of 79%), and word meaning (reliability of 86%). Overall reliability was 82%. While reliability for the current study was low, it was consistent with reports of low reliability for transcriptions of toddler productions compared to older children (Stockman, Wood, & Tishman, 1981).

Behavioral analysis

Child utterances (a total of 4,434) were analyzed for discourse context using the following categories: co-participatory (CP), initiation (IN), narrowed focus (NF), and undetermined (U). CP and IN utterances were considered *communicative* and NF utterances were considered *non-communicative*. CP utterances (totaling 2,436) had joint attention. Following definitions of *interactive* utterances in Vihman & Miller (1988), CP utterances occurred within a reasonable conversational timeframe compared to the previous adult utterance. Other indicators of co-participation included gestures that signaled communication, eye gaze, and checking behaviors (all of which are aspects of joint attention).

IN utterances (totaling 1,260) had many of the same behavioral characteristics as CP interactions, but indicated a change in topic. For both categories, lexical cues were important indicators, but behaviors such as pointing and other gestures, eye gaze, checking, and object focus were also used to determine a change in topic or attention.

NF utterances (totaling 354) were those in which the child vocalized within an already established interaction but narrowed his or her attention into a more personal space. This category included utterances produced when the child was playing by him/herself, or there was little, if any, eye gaze or checking behavior, similar to Vihman & Miller's (1988) non-interactional category. Utterances were categorized as *undetermined* (totaling 384 utterances) if their conversational context could not be determined. The undetermined category was typically utilized in cases where the child was not facing the camera or was not visible on the video tape. Inter-rater agreement was completed on 10% of each sample. Statistical analysis using Cohen's kappa to correct for chance agreement yielded an agreement level of 0.49, which is considered fair (Fleiss, 1981; Bakeman & Gottman, 1997).

Intonation analysis

Intonation was measured acoustically using TF32 (Milenkovic, 2001), a signal analysis program using a 16-bit resolution and a sampling rate of 22kHz. Measures were completed on the nuclear tone of single syllable utterances. The nuclear tone was measured from the beginning of the vocalic nucleus to the final boundary of the tone (Allen & Hawkins, 1980). Boundaries were determined by the first clear and periodic cycle to the last in the vocalic nucleus. For each tone, a fundamental frequency (f_0) contour (*i.e.*, a pitch contour) was created using TF32.

Intonation was measured in three ways. First, the basic directionality (*i.e.*, falling or rising) of the intonation contour was determined. Secondly, the accent range (*i.e.*, pitch change) was measured by calculating the difference between the minimum and maximum f_0 pitch points. Accent range measures were calculated in semitones to equalize pitch height differences (see Burns & Ward, 1982). Finally, each contour was classified by contour shape (to yield a contour inventory count and contour maturity measures; see Figure 1, adapted from Figure 1 in Balog & Snow, 2007).

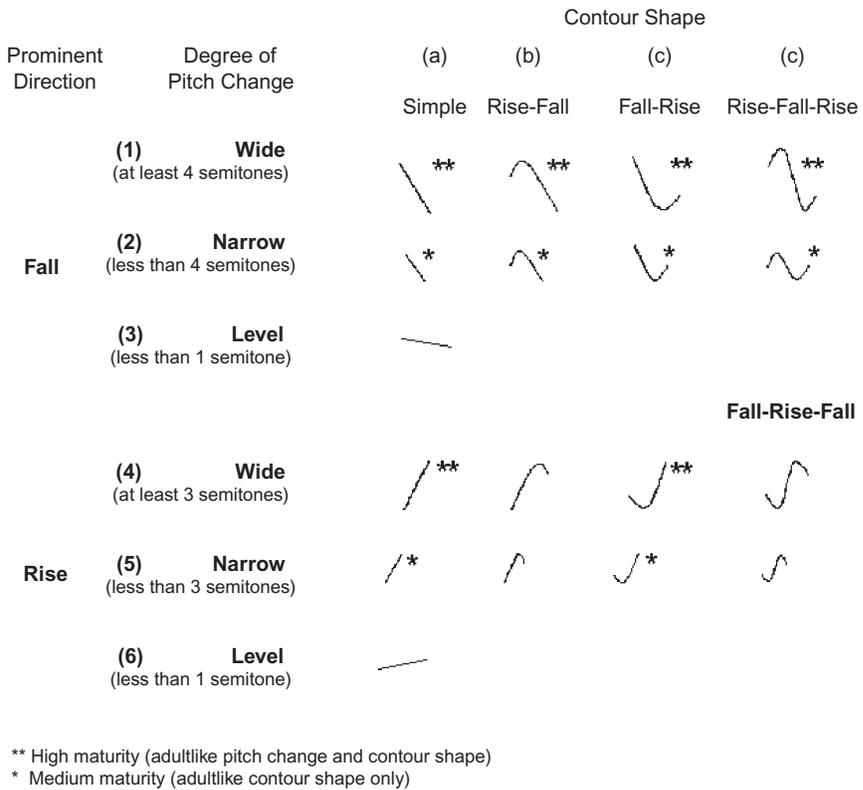


Figure 1.
Contour inventory (adapted from Balog & Snow, 2007).

RESULTS

Effect of utterance meaning

The data in the current study included both meaningful (*i.e.*, lexically rich) and non-meaningful (*i.e.*, babbled) utterances. Previous research has not indicated that the distinction between non-meaningful and meaningful utterances influences intonation in a consistent manner (Snow, 2002). To test this assumption, statistical analysis with meaningfulness as one of the independent variables was conducted.

Of the 24 children, 18 produced both non-meaningful and meaningful utterances with both falling and rising contour productions. The data were analyzed using a three-way ANOVA: age (2) x contour direction (2) x meaningfulness (2). Statistical analysis indicated no main effects or interactions; therefore meaningfulness was not considered a factor in the following analyses.

Accent range measures

A total of 2,358 utterances were analyzed for accent range. The mean accent range measures are reported in Table 1. The data were analyzed using a three-way

ANOVA: age (2) x contour direction (2) x discourse context (3). There were no significant main effects or interactions. When intonation was measured using accent range the analysis did not reveal any differences between falling and rising contours or between the three discourse contexts in toddler productions.

Table 1.

Means (semitones) and Standard Deviations (SD) of Accent Range in Nuclear Tones by Contour Direction (*i.e.*, fall and rise), Discourse Context (*i.e.*, CP, IN, and NF) and Age.

Age (months)	Contour Direction	Discourse Context					
		CP		IN		NF	
		Mean	SD	Mean	SD	Mean	SD
12-17	Falls	4,90	1,81	4,73	2,00	6,35	2,75
	Rises	3,95	1,97	5,04	2,01	4,60	3,03
18-23	Falls	5,37	1,36	5,80	2,01	5,94	3,17
	Rises	5,71	2,07	4,82	2,17	5,07	2,53

Contour inventory

The mean contour inventory size measures are reported in Table 2. These data were also analyzed using a three-way ANOVA: age (2) x contour direction (2) x discourse context (3). When intonation was measured by contour inventory there was a main effect for discourse context ($F = 64.84, p < .001$). Post Hoc analysis of the discourse context data revealed that the children used significantly larger contour inventories in their communicative utterances compared to their non-communicative utterances.

Table 2.

Mean Size and Standard Deviation (SD) of the Contour Inventory in Nuclear Tones by Contour Direction (*i.e.*, falls and rises), Discourse Context (CP, IN, and NF), and Age.

Age (months)	Contour Direction	Discourse Context					
		CP		IN		NF	
		Mean	SD	Mean	SD	Mean	SD
12-17	Falls	4,22	1,20	3,44	0,53	1,67	1,66
	Rises	2,89	1,27	2,78	1,30	1,11	0,78
18-23	Falls	4,50	0,53	3,50	0,93	0,88	1,46
	Rises	3,88	0,99	3,25	1,39	0,63	1,06

Contour maturity

Falling contours. The mean percentage of utterances occurring at each maturity level was calculated for each discourse context (refer to Table 3). Statistical analysis using a one-way ANOVA revealed significant differences for all discourse categories for falling contours (for CP, $p < .001$; for IN $p < .001$, and for NF, $p < .001$). Post Hoc analyses using Tukey HSD revealed that CP and IN utterances were characterized by significantly more high and medium maturity contours compared to low maturity contours (high-low comparison, $p < .001$; medium-low comparison, $p < .001$). High and medium contours were not different from one another. Maturity levels were all significantly different for NF utterances (high-medium comparison, $p < .01$; high-low comparison, $p < .001$; medium-low comparison, $p < .001$). The results revealed that for the more communicative discourse contexts (*i.e.*, CP and IN) children used more stable intonation production. In other words, they consistently produced forms that were categorized as having high or medium maturity (therefore, more adultlike). When children produced NF utterances their intonation was characterized by more instability. Interestingly, NF utterances were characterized by the largest percentage of highly mature contours and the lowest percentage of low maturity contours.

Table 3.

Mean Percentage of Occurrence of Utterances in each Discourse Context and Maturity Level Category for Rises and Falls.

Contour Direction	Discourse Context	Maturity Level		
		High	Medium	Low
Falls	CP	50.58%	42.22%	7.78%
	IN	49.06%	46.50%	5.00%
	NF	59.50%	38.94%	0.44%
Rises	CP	38.89%	25.72%	35.39%
	IN	40.28%	25.56%	34.72%
	NF	30.47%	16.79%	52.74%

Rising contours. The mean percentage of utterances occurring at each maturity level was calculated for each discourse context (refer again to Table 3). Statistical analysis using a one-way ANOVA revealed significant differences for the NF discourse context only ($p < .01$). Post Hoc analyses using Tukey HSD revealed that the utterances produced in CP and IN contexts used similar percentages of high, medium and low maturity contours. In other words, there were no significant differences in maturity level for these communicative contexts. NF utterances were characterized by significantly more low maturity contours

compared to medium contours ($p < .01$). These results indicated that intonation production is less stable for rising contours.

DISCUSSION

This study sought to determine the relationship between suprasegmental phonological skills (*i.e.*, intonation) and discourse (a linguistic skill marked by intonation) in early language learners. In order to do so, recently developed methodologies for analyzing both intonation (Balog & Snow, 2007) and discourse (Balog & Roberts, 2004) were applied and expanded.

In terms of methodological contributions to the analysis of developmental intonation, this study solidified the usefulness of the contour inventory measure. The results support previous research (Balog & Snow, 2007) in which the contour inventory measure was found to be more sensitive to intonation differences in early productions. The expanded use of the inventory to measure contour maturity further enhanced our knowledge of early language development.

By measuring intonation in terms of maturity level, the data revealed not only a phonological difference between falling and rising contours (as expected based on previous research; see Kent & Bauer, 1985, Kent & Murray, 1982), but also more stable use of intonation in falls produced in communicative compared to non-communicative discourse contexts (*i.e.*, a pragmatic difference in production). This pragmatic differentiation for communicativeness was not observed for rising contours.

The current study revealed that children differentiated their productions phonologically when engaged in more communicative contexts (*i.e.*, they produced more contour types in communicative contexts, defined as CP and IN in the current study). Of even greater interest, the children differentiated their productions within CP and IN contexts, by producing more contour types in CP contexts. The CP context was characterized by topic maintenance in ongoing adult-child interactions in which toddlers have topical support and are able to maintain that topic. This finding is likely a result of toddler's early awareness of conversational floor and topic maintenance. Their use of reduced contour inventory in their initiated productions could reflect a greater cognitive load to the system and possibly the lack of a model from the adult input (*i.e.*, the toddler's had to expend more energy on topic initiation, thus reducing their ability to increase phonological complexity).

These findings reflect language development characteristics of the first-word lexical stage. The behavioral cues used to identify discourse categories in the current study are not unique to children aged 12 to 23 months or the first-word period. Cues like joint attention, body posture, eye gaze, and gesturing are used by all communicators, regardless of age. These behaviorally based categories might improve the analyses of discourse in older or more linguistically advanced children because of the reliance on more basic skills, as compared to measures such as phonemic inventories or syntactic structure. Behavioral skills may

prove useful in analyzing the discourse interactions of children from varying socioeconomic classes or cultural settings and clinical populations by partially removing the complication of varying language skills.

Issues of reliability for discourse context judgements must be addressed prior to the application of these of the methods to any new population. The reader is referred to the reliability rating of *fair* for categorization of utterances into various discourse contexts. What can be considered statistically reliable does not necessarily translate into useful behavioral reliability. Future research should focus on improving reliability either by defining the discourse contexts for increased clarity or by implementing the use of multiple rater judgements for categorization (e.g., Balog & Brentari, 2008).

Finally, it is hoped that in the future these methodologies can be applied to disordered populations and lead to new methods for the earlier and more accurate diagnoses of communication disorders. It is currently proposed that the study of intonation production in normally developing children will eventually find application in disorders characterized by social, pragmatic, and affective impairment.

REFERENCES

- Allen, G.D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G.H. Yeni-Komshian, J.F. Kavanagh & C.A. Ferguson (Eds.), *Child Phonology. Vol. 1: Production* (pp. 227-256). New York: Academic Press.
- Bakemen, R., & Gottman, J.M. (1997). *Observing interaction: An introduction to sequential analysis (second edition)*. Cambridge, UK: Cambridge University Press.
- Balog, H.L., & Brentari, D. (2008). The relationship between early gestures and intonation. *First Language*, 28, 141-163.
- Balog, H.L., & Roberts, F.D. (2004). Perception of utterance relatedness during the first-word-period. *Journal of Child Language*, 31, 837-854.
- Balog, H.L., & Snow, D. (2007). The adaptation and application of relational and independent analyses for intonation production in young children. *Journal of Phonetics*, 35, 118-133.
- Beckman, M.E., & Elam, G.A. (1997). *Guidelines for ToBI Labelling. Version 3*. Columbus: Ohio State University.
- Bloom, L., Rocissano, L., & Hood, L. (1976). Adult-child discourse: Developmental interaction between information processing linguistic knowledge. *Cognitive Psychology*, 8, 521-552
- Burns, E.M., & Ward, W.D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 241-269). New York: Academic Press.
- Cruttenden, A. (1997). *Intonation (second edition)*. Cambridge, MA: Cambridge University Press.

- D'Odorico, L. (1984). Non-segmental features prelinguistic communication: An analysis of some types of infant cry and non-cry vocalizations. *Journal of Child Language*, *11*, 17-21
- Fleiss, J.L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Kent, R.D., & Bauer, H.R. (1985). Vocalizations of one-year-olds. *Journal of Child Language*, *12*, 491-526.
- Kent, R.D., & Murray, A.D (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, *72*, 353-365.
- Milenkovic, P (2001). *TF32* [computer software]. Madison: Department of Electrical Engineering, University of Wisconsin.
- Snow, D. (2002). Intonation in the monosyllabic utterances of 1-year-olds. *Infant Behavior and Development*, *24*, 393-407.
- Snow, D. (2004). Falling intonation in the one- and two-syllable utterances of infants and preschoolers. *Journal of Phonetics*, *32*, 373-393.
- Snow, D., & Balog, H.L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, *112*, 1025-1058.
- Stockman, I.J., Woods, D.R., & Tishman, A. (1981). Listener agreement on phonetic segments in early infant vocalizations. *Journal of Psycholinguistic Research*, *10*, 593-617.
- Vanderslice, R., & Ladefoged, P (1972). Binary suprasegmental features and transformational word-accentuation rules. *Language*, *48*, 819-838.
- Vihman, M.M., & Miller, R. (1988). Words and babble at the threshold of lexical acquisition. In M.D. Smith & J.L. Locke (Eds.), *The emergent lexicon: The child's development of a linguistic vocabulary* (pp. 151-183). New York: Academic Press.