

Final Syllable Lengthening (FSL) in infant vocalizations*

SUNEETI NATHANI

The University of Georgia

D. KIMBROUGH OLLER

The University of Memphis

ALAN B. COBO-LEWIS

The University of Maine

(Received 1 November 2000. Revised 1 March 2002)

ABSTRACT

Final Syllable Lengthening (FSL) has been extensively examined in infant vocalizations in order to determine whether its basis is biological or learned. Findings suggest there may be a U-shaped developmental trajectory for FSL. The present study sought to verify this pattern and to determine whether vocal maturity and deafness influence FSL. Eight normally hearing infants, aged 0;3 to 1;0, and eight deaf infants, aged 0;8 to 4;0, were examined at three levels of prelinguistic vocal development: precanonical, canonical, and postcanonical. FSL was found at all three levels suggesting a biological basis for this phenomenon. Individual variability was, however, considerable. Reduction in the magnitude of FSL across the three sessions provided some support for a downward trend for FSL in infancy. Findings further indicated that auditory deprivation can significantly affect temporal aspects of infant speech production.

[*] This work was supported in part by a dissertation grant from the Purdue Research Foundation. We gratefully acknowledge the invaluable assistance of Michael Lynch in designing the study and of Pat Blackwell and Ayasakanta Rout in determining reliability. We would also like to thank the reviewers and Joe Dinapoli for their comments. Portions of this work were presented at the 1999 Child Phonology Conference in Wales. Address for correspondence: Suneeti Nathani, Assistant Professor, Communication Sciences and Disorders, 528 Alderhold Hall, The University of Georgia, Athens, GA 30602, USA. tel: (706) 542 4602. fax: (706) 542 5348. e-mail: snathani@coe.uga.edu

INTRODUCTION

Prosody, which includes the components of rhythm and intonation, plays a pivotal role in adult speech. For instance, rhythm and intonation help divide up the continuous speech stream into units that are optimally sized for coding by the human processing capacities thereby increasing the efficiency of perception and production (Martin, 1972). There are numerous theoretical treatments of prosody or rhythm in speech (e.g. Selkirk, 1984; Oller & Lynch, 1992), all of which posit the syllable (or the related concept of ‘mora’) as the universally minimal rhythmic unit. Higher order rhythmic units are universally built from syllables into feet, phrases, and so on, according to all these standard treatments. Implementations of rhythmic units in each natural language in the world are, however, unique. For example, one temporal aspect of speech rhythm, namely Final Syllable Lengthening (FSL), is widespread but not universal.

FSL, as the name implies, refers to lengthening of the final syllables of units, such as phrases and sentences (Oller, 1973). FSL is strong in English, French, and German whereas in Japanese or Finnish, FSL is weak or may be absent (Delattre, 1966; Klatt, 1976; Oller, 1979; Hoequist, 1983). Given these cross-linguistic differences in FSL, it would be tempting to conclude that FSL is a learned product of ambient language influences such that over time, speakers learn to employ the amount of FSL appropriate for their language to mark the ends of units. Moreover, because every language employs some device to mark units, e.g. initial stress in Hungarian (Konopczynski, 1993) or final stress with FSL in French, it has been hypothesized that FSL and other final-syllable effects (F_0 declination, for example) represent a perceptual phenomenon in that they serve to cue the listener to unit boundaries (Oller, 1973; Klatt, 1976).

Rhythm clearly, however, has biological roots that are not limited to speech and apply to a variety of perceptual and motor systems. For example, finger movements and speech perception show evidence of rhythmic organization similar to that of speech (Allen, 1975). With a few exceptions (e.g. FSL is absent in some nonfluent aphasics; see Baum & Boyczuk, 1999), FSL is also fairly robust across a variety of conditions, e.g. different speaking rates (Weismer & Ingrisano, 1979; Cummins, 1999) and different vowel types (Erickson, 2000). This widespread and robust occurrence of FSL has led some researchers to suggest that FSL may simply be the result of a generalized propensity to slow down at the end of motoric sequences. Thus, FSL may be more indicative of a general motor phenomenon rather than a learned perceptual phenomenon (e.g. Oller, 1973; Klatt, 1976).

It may eventually turn out to be that some aspects of FSL in speech initially emerge due to a general (perhaps innately specified) tendency of organisms to arrange serial information in a hierarchical grouping of units. These aspects

might then eventually be modified and/or other aspects of FSL might emerge due to ambient language influences (e.g. Robb & Saxman, 1990).

One way by which we might be able to differentiate aspects of FSL that are biological from those that are learned is by examining prespeech infant vocalizations. Because prespeech vocalizations do not contain any adult linguistic structure, aspects of FSL observed in prespeech vocalizations would indicate that those aspects arise due to a general tendency that is either innate or acquired early in life with minimal or nil linguistic influence. Modification of final-syllable effects or occurrence of new aspects later in infant vocalizations would presumably be the result of language-specific, or learned, influences. Examining FSL in deafness offers another way to test divergent explanations regarding the origins of FSL. Features of FSL that are exhibited in the vocalizations of prelinguistic deaf infants would be consistent with the possibility that these features are innate, or acquired with only minimal auditory information, or possibly acquired via non-auditory modes (Lynch, 1996). On the other hand, features of FSL that are observed only in normally hearing infants would suggest that perceptual factors, such as auditory experience, play a major role in the adequate development of these features.

A notable, though difficult-to-interpret literature exists on FSL in infant and young children's vocalizations. Results from these studies, summarized in Table 1, yield a confusing picture regarding the emergence of FSL. Some investigations reveal little or no FSL in infants' vocalizations (Oller & Smith, 1977; Halle, de Boysson-Bardies & Vihman, 1991; Levitt & Wang, 1991; Lynch, Oller, Steffens & Buder, 1995) whereas others note consistent FSL (Zlatin Laufer, 1980; Stark, 1989; Robb & Saxman, 1990; Halle *et al.*, 1991; Levitt & Utman, 1992; Konopczynski, 1993; Snow, 1994; Vihman & DePaolis, 1998). Closer examination of the data in Table 1, however, suggests that FSL may not be as unpredictable as it appears given the disparate outcomes that have been obtained. The occurrence of FSL is influenced by a variety of factors, such as sample size, sample characteristics, ages studied, language background of participants, and criteria used to determine FSL. Once these factors are taken into account, discrepant results across investigations become easier to reconcile with a U-shaped developmental trajectory of FSL. FSL appears to be present in the vocalizations of typically developing two- to six-month old infants (e.g. Zlatin Laufer, 1980) and appears to be absent somewhere in the middle of the first year of life (e.g. Oller & Smith, 1977). FSL then appears to re-emerge sometime at the end of or after the first year of life (e.g. Robb & Saxman, 1990), and may be especially spurred along toward early re-emergence by the influence of very consistent FSL in the ambient language, as is the case with French (e.g. Levitt & Wang, 1991). However, the exact age of re-emergence of FSL is not clear, as ages of re-emergence of FSL have varied across studies.

TABLE 1. *Findings from studies of FSL in infants' and young children's productions*

Study	Language	Ages studied	Vocal/linguistic abilities	Sample analysed	FSL findings
Oller & Smith (1977)	English	0;8-1;0	Infants produced reduplicated sequences	Reduplicated vocalizations	Minimal
Zlatin Laufer (1980)	English	0;0-0;6	All infants were in the expansion stage at end of study	Disyllables	Present, especially in the second 12 weeks of life
Stark (1989)	English	0;0-0;8	Infants were in the expansion stage at end of study	All vocalizations	Present after the second month of life
Robb & Saxman (1990)	English	0;8-2;2	Matched infants at preword, single-word, and multiword stages	Disyllables	Present at all stages, more marked in closed than in open syllables
Halle, de Boysson-Bardies & Vihman (1991)	French, Japanese	1;2.5-1, 11.2	25-word stage	Disyllables	Present in French; only one Japanese infant showed FSL
Levitt & Wang (1991)	English, French	English : 0;7-1;0 French : 0;5-1;1	Babbling	Reduplicated vocalizations	Minimal in English; present in French
Levitt & Utman (1992)	English, French	0;5, 0;8, 0;11, and 1;2	Vocal play to first words	Reduplicated vocalizations	Present at 0;11 and 1;2; more pronounced for French
Konopczynski (1993)	French	0;9-2;0	Vocal play to word combinations	All vocalizations	Emerges in interactions with adults between 1;1-1;4
Snow (1994)	English	1;4-2;1	Single words to word combinations	All vocalizations matched for phonetic category and affect level	Variable at single word stage; consistent after word combinations
Lynch <i>et al.</i> (1995)	English	0;2-0;4, 0;6-0;8, and 0;10-1;0	Prelinguistic	All utterances and phrases	None in utterances; present in phrases at 0;2-0;4 and 0;10-1;0 but not at 0;6-0;8
Vihman & DePaolis (1998)	English, French	English : 0;10-1;5; French : 0;10-1;6	Matched at 4-word and 25-word point	Disyllables	Adult like FSL at 50-word stage for French; English showed FSL only in iambs; 4-word point, FSL present but less than adult like ratios for French

Apparent exceptions to this trend can be partly explained by methodological considerations. For instance, Levitt & Wang (1991) reported that French infants exhibited FSL in their reduplicative babbles even in the middle of the first year of life. The set of reduplicative babbles used for analysis for each infant, however, were collapsed across ages, ranging from 0;5 to 1;1. Consequently, it is difficult to determine whether FSL occurred across the entire age range or whether FSL was restricted to the later ages. In addition, a very small set of reduplicative babbles, less than 10, was used to analyse FSL for two of the five French infants.

Another apparent exception to the absence of FSL in the middle of the first year of life is seen in the study by Robb & Saxman (1990). These investigators compared disyllabic productions of children, aged 0;8–2;4, with similar linguistic abilities, i.e. children were matched at preword, single word, and multiword stages, and reported reliable FSL at all levels of linguistic development, including infants, aged 0;8. Of the seven children examined, however, only three were observed at eight months of life. Furthermore, data for these three infants were collapsed with the other infants in the group, aged 0;8–1;1. It is, therefore, not possible to determine the exact ages when FSL was present; it may have been present anytime within this age span.

In the Lynch *et al.* (1995) study, no FSL was detected in utterances produced by infants, aged 0;2 to 0;4, 0;6 to 0;8, and 0;10 to 1;0. FSL was, however, observed in phrases in infants aged 0;2 to 0;4 and 0;10 to 1;0, but not in infants aged 0;6 to 0;8. Phrases represent the third level of organization in a presumed rhythmic hierarchy of infant vocalizations (Oller & Lynch, 1992). Syllables constitute the first level of organization in infant productions and utterances are the second level. Work in training observers in our laboratory has indicated that sequences of vocalizations classified as phrases in the Lynch *et al.* investigation may have been classified as utterances in previous investigations (e.g. Stark, 1989). Thus, phrasal results from the Lynch *et al.* study may be partly comparable to those obtained from prior research on FSL in utterances and are therefore, compatible with the hypothesized U-shaped trend. It should be noted that the definition of an utterance has varied considerably across different investigations. We have, therefore, restricted the use of the term 'utterance' in this article to refer only to the technical utterance level posited by Oller & Lynch (1992) and used 'vocalization' elsewhere.

To reiterate, once methodological differences are taken into account, the data, as reported in Table 1, suggest that FSL is present early in life, tends then to disappear in the middle of the first year of life, and makes a variable appearance toward the end of the first year of life or later, perhaps dependent upon the ambient language. Systematic investigation of FSL across the entire first year of life and beyond is, however, necessary in order to substantiate this claim. The present work is an attempt to begin sorting this issue out, in a study of eight normally hearing infants in the first year of life.

To our knowledge, only one prior investigation has examined English-learning infants over the entire first year of life (Lynch *et al.*, 1995). Unlike Lynch *et al.*, however, we matched infants according to level of vocal development rather than age because it is well known that infants of the same chronological age can vary considerably in their level of vocal maturity. Consequently, contradictory findings of FSL from infants of the same age across studies might be partly explained on the basis of differences in vocal maturity among these infants. Prior investigations with older infants have revealed significant differences in FSL with the acquisition of new linguistic milestones, such as production of two-word vocalizations (Snow, 1994). Given these findings, changes in FSL in younger infants with changes in their level of vocal development, e.g. onset of canonical syllable production, might be expected. To examine this possibility, we matched infants according to level of vocal development. Possible FSL was observed at three levels of infant vocal development: prior to canonical syllable production (precanonical), during canonical syllable production (canonical), and following canonical syllable production (postcanonical).

Previous studies have varied as to the types of vocalizations examined for FSL. When only reduplicated vocalizations are used for analysis (e.g. Oller & Smith, 1977), there is a necessary substantial reduction in sample sizes. Reduplicated vocalizations are preferred by many investigators because FSL in reduplicated sequences is not confounded with possible inherent segment duration effects. In contrast, other investigators, such as Zlatin Laufer (1980), used DISYLLABIC vocalizations (allowing final and nonfinal types to differ phonetically) because this approach results in a much larger sample size. Still other investigators (e.g. Stark, 1989) used all vocalizations, irrespective of type, for the analysis of FSL. As Lynch and colleagues (1995, 1996) opined, it may be that FSL in infant utterances occurs primarily at a global level and therefore, the true nature of FSL may be revealed only when looking at all types of utterances. The present work included all utterances (disyllabic, multisyllabic, etc.) in the analysis of FSL. To further avoid possible confounding effects of sample size, syllable duration values were weighted according to the number of utterances available for analysis. Sub-analyses were also conducted on disyllabic utterances for comparisons with previous literature.

Although some prior investigations have included monosyllabic vocalizations (assuming each monosyllable to be a final syllable) in investigating FSL (e.g. Snow, 1994), we excluded these vocalizations from our analyses to avoid confounding FSL and syllable compression (or syllable isochrony effects). Syllable compression, or reduction in duration of syllables, occurs when the number of syllables within a unit increases, and is a phenomenon that is well documented to occur in a variety of adult languages (e.g. Oller, 1973). For example, the syllable ‘cup’ may be compared with ‘puh’ in ‘puppy’ (Snow,

1994). In this circumstance, however, it is possible that 'cup' would be longer than the 'puh' in 'puppy' not only because of FSL effects, but also because 'cup' is monosyllabic, whereas 'puh' occurs within a disyllabic word.

The robustness of FSL was evaluated in the present study by including eight deaf infants, i.e. infants with severe-to-profound hearing losses due to various causes, matched with normally hearing infants according to level of vocal development. If, in fact, FSL was observed in the vocal samples of these deaf infants, it would indicate that FSL is fairly robust with respect to handicapping conditions, such as deafness. If, on the other hand, FSL was absent or aberrant in deaf infants, it would indicate that perceptual modelling is necessary for the typical development of FSL.

Scant evidence is available regarding general prosodic features of vocalizations in deaf infants (Stark, 1972; Kent, Osberger, Netsell & Hustedde, 1987). The available evidence suggests that intonation and intensity are perhaps the parameters most likely to be affected in the absence of audition. Kent *et al.* (1987) reported that although one deaf infant had a similar range and peak value of F_0 as his normally hearing twin brother, he produced more variable F_0 contours than his normally hearing twin. Furthermore, instances of vocal fry or noise were common in the vocalizations of the deaf infant. Stark (1972) observed that deaf infants frequently produced vocalizations with very little variation in pitch and intensity.

Several investigations have been conducted on prosody in the productions of older deaf children and deaf adults. These investigations reveal significant difficulties with some rhythmic parameters, e.g. intonation, for deaf children (e.g. Calvert, 1961; Monsen, 1979). The key finding regarding duration is that deaf speakers often produce longer syllables and sentences (often much longer) than their hearing counterparts (Calvert, 1961). This finding does not, however, apply to all individuals with hearing impairment (Monsen, 1979). It is difficult to ascertain whether the observed temporal aberrations are due to a general slowing of production such that all syllables are lengthened, or whether the slowing is an artifact of poor control over specific temporal parameters, such as FSL, because the studies cited did not tease apart absolute and relative temporal duration values. The present study investigated duration of both nonfinal and final syllables in order to determine whether longer durations observed in deaf speech are merely due to generally overlong productions or due to difficulty with relative temporal phenomena, such as those seen in FSL.

Only one known prior investigation of a single acochlear (completely deaf) infant (Lynch, 1996), aged 2;3 to 3;6, has examined FSL specifically in relation to deafness. As in the case of the Lynch *et al.* (1995) study on hearing infants, FSL was revealed only in phrases but not in utterances in the acochlear infant. FSL in phrases, however, did not stabilize until around 2;9. Lynch further noted that the deaf infant produced overlong syllables and

utterances within phrases. When results from Lynch *et al.* (1995) and Lynch (1996) are combined, it appears that both nonfinal and final syllables of the deaf infant were longer in comparison to those produced by normally hearing infants. The FSL ratio of the deaf infant was, however, similar to that of normally hearing infants thereby providing some support for the hypothesis that deafness results in a generalized slowing of speech production. Results from the Lynch (1996) study should, however, be cautiously interpreted because they are based upon a single infant whose deafness was caused by a rare condition. Although complete loss of audition can provide valuable insights on the role of perceptual modelling in phrasing, it does not represent the condition of the deaf population in general because small amounts of residual hearing are usually present even in the general deaf population. In addition, the role of audition in FSL would be better understood if a larger sample size could be studied.

To summarize, the present study sought to determine the occurrence of FSL in utterances produced by normally hearing and deaf infants at different levels of vocal development. The primary research questions were as follows:

- (1) What is the pattern of FSL in prelinguistic utterances? Eight normally hearing infants in the first year of life were evaluated.
- (2) Does FSL vary with vocal maturity? This question was examined by measuring FSL at three levels of vocal development.
- (3) Is FSL influenced by auditory status? Eight deaf infants were matched with eight normally hearing infants.

METHOD

Participants

Eight full-term normally hearing infants and eight full-term deaf infants were longitudinally studied. Infants in each group were matched for socioeconomic status and language background, Spanish or English, bilingual or monolingual.

Gender was approximately balanced in the deaf group. In the normally hearing group, there were seven male infants and one female infant. The gender mismatch was not expected to influence results because gender differences in the qualitative characteristics of prelinguistic vocal development have not been demonstrated in any previous work (e.g. Lynch *et al.*, 1995). Normally hearing infants were recruited via mail solicitation. Infants' auditory status was screened using computer-controlled visually reinforced audiometry and standard tympanometry. Infants' hearing levels and tympanograms were found to be within normal parameters. In addition, they had unremarkable medical and developmental histories.

Unaided better-ear pure tone averages for deaf infants using behavioural audiometry revealed severe-to-profound and profound hearing losses. Severe

TABLE 2. *Demographics for deaf infants*

ID	Gender	Language	SES	Aetiology	Age at identification	Degree of loss	Age at amplification
9	Male	English	High	Genetic	0;1	Severe	0;1
10	Male	English	High	Genetic	0;3	Severe-to-profound	0;3
11	Male	English	Mid	Unknown	1;2	Profound	1;3
12	Female	Spanish	Low	Unknown	1;1	Profound	1;2
13	Male	Bilingual	Mid	Mondini's dysplasia	1;9	Profound	1;11
14	Female	English	Low	Unknown	1;1	Profound	1;3
15	Female	Bilingual	Low	Genetic	1;7	Profound	1;11
16	Male	Bilingual	Mid	Unknown	1;7	Profound	1;8

hearing loss was defined by hearing thresholds in the range of 71–90 dB HL and profound hearing loss was defined by hearing thresholds at 91 dB HL or greater. Two of the eight deaf infants had some complications at birth including hyperbilirubinemia and cold stress syndrome, but hearing loss in these infants was not attributed to their birth complications. A summary of the demographic characteristics of the deaf infants is provided in Table 2.

Six deaf infants attended a total communication classroom; two infants were enrolled in an oral communication programme. All eight infants were fitted with hearing aids appropriate for their degree of hearing loss. In addition, six infants were fitted with tactile aids and received amplification through an FM system.

Recording environment and procedures

Vocalizations of normally hearing infants were recorded every two weeks in a sound-treated room that was equipped with quiet toys, a bed, and high-fidelity recording equipment (Marantz PMD-221 audiocassette recorder and Bose PM-10 external microphone). The infant, one parent, and a recording assistant were typically present. Infants could move freely within the room and play with the toys. Parents were encouraged to elicit vocalizations from infants by eye contact and reciprocal vocalizations. Parents sometimes played with the infant, engaged in face-to-face interaction with the infants, or conversed with the recording assistant. The recording assistant also sometimes played or interacted with the infants. Parents and the recording assistant attempted to be silent when the infant vocalized. A session usually lasted about 20 min. Participants were reimbursed at the end of each session.

Monthly recordings for six deaf infants were made using a Sony CFS-720 recorder in a quiet room that was designated for speech training. A speech-language pathologist and the infant were typically present in the recording

session. The speech-language pathologist used a variety of age-appropriate activities and materials (e.g. bubble blowing), to elicit vocalizations from the infants. The remaining two deaf infants were recorded under the same conditions and procedures as normally hearing infants.

Criteria for selection of sessions

The onset of canonical syllable production was used as the criterion to select three recording sessions for analysis for normally hearing infants. Canonical syllables (e.g. [ba]) are speech-like, have a consonant-like element (other than a glottal stop or glottal fricative) and adult-like vowels (or syllabic nuclei) and include rapid formant transitions which are generally less than 120 ms. Onset of canonical syllable production was estimated to occur when the canonical babbling ratio (or CBR, the total number of canonical syllables per total number of syllables) in a session was first equal to or greater than 0.2 (Steffens, Oller, Lynch & Urbano, 1992). This ratio was chosen because it has been confirmed that parents and naïve observers reliably identify the production of canonical syllables when the ratio of canonical syllable production reaches 0.2. The canonical session was always the first session to be selected for analysis because it served as the basis for selection of the other two sessions for an infant. One of the other two sessions was a 'precanonical' session that had fewer than 20% canonical syllables. The other session was a 'postcanonical' session that had greater than 20% canonical syllables. In general, selection of precanonical sessions was aimed at those recorded approximately three months prior to the canonical sessions, and selection of postcanonical sessions was aimed at those recorded approximately three months subsequent to the canonical sessions. Variations were based upon availability of recorded sessions. Assignment of sessions to the three vocal development categories was based upon preliminary determinations of CBRs obtained in codings of tape recordings at the University of Miami. Subsequently, data were more extensively recoded to determine CBRs by the first author.

For deaf infants, either the session in which 20% canonical syllable production was first observed or the session corresponding to the teacher-designated or parent-reported age of onset of canonical syllable production, whichever occurred earlier, was designated the canonical session. Teacher or parent report was included because Oller & Eilers (1988) have observed that designating the age of onset of canonical babbling in the laboratory is especially difficult for deaf infants as these infants often demonstrate inconsistent production of canonical syllables across sessions.

As with the hearing infants, recorded data were not always available at three months before and after the designated canonical session. In such cases, the selected precanonical session was the earliest session recorded prior to the onset of canonical syllable production, irrespective of whether or not it

TABLE 3. *Ages corresponding to the selected precanonical, canonical, and postcanonical sessions*

Normally hearing infants				Deaf infants			
ID	Precan	Can	Postcan	ID	Precan	Can	Postcan
1	0;3	0;5.15	0;8.15	9	0;8	0;11	1;1
2	0;4	0;7	0;8	10	1;0	1;1	1;4
3	0;3.15	0;7	0;10	11	1;10	2;2	2;3.15
4	0;3.15	0;7	0;10	12	2;0.15	2;2	2;4.15
5	0;6	0;7	0;11.15	13	2;1	2;4	2;8
6	0;5	0;7.15	0;9	14	2;4	2;7	2;10
7	0;5.15	0;8	0;11	15	2;8.15	2;11.15	3;1
8	0;3.15	0;8	0;10	16	3;8	3;9	4;0
<i>M</i>	0;4.8	0;7.4	0;9.23	<i>M</i>	2;0.11	2;2.28	2;5.15
S.D.	0;1.3	0;0.24	0;1.6	S.D.	0;11.1	0;11.4	0;11.7

Precan, Precanonical; Can, Canonical; Postcan, Postcanonical.

occurred three months prior to the canonical session, and postcanonical sessions were likewise assigned based on the available sample that best met the three-month criterion. Ages corresponding to the selected precanonical, canonical, and postcanonical sessions for each infant are provided in Table 3.

Analysis of infant vocalizations

The audiotape recordings for each infant were analysed in three phases:

Phase 1. Protophones, i.e. vocalizations without links to obvious biological functions, as occur with vegetative sounds, (e.g. cough, sneeze) or obvious social functions, as occur with fixed signals (e.g. cry, laugh), were identified by a trained judge. The term ‘protophones’ has been adopted recently and includes vocalization types such as cooing and babbling (Oller, 2000). Protophones have been referred to earlier as ‘nonvegetative’ sounds and have been selected for analysis in a number of previous studies (e.g. Lynch *et al.*, 1995) because they are assumed to be precursors to meaningful speech (Oller & Lynch, 1992).

Utterances were generally defined as vocalizations or groups of vocalizations separated from all others by audible ingressive breaths or separable in accord with adult judges’ intuitions about utterance boundaries (Oller & Lynch, 1992). Utterances were coded by a primary judge who designated counter numbers for their locations on recordings. A second judge coded a randomly selected subset of vocalizations (10% from each infant group). Cohen’s kappa for interjudge agreement of utterance determinations for all infants was 0.66.

Phase 2. A trained adult judge used the Logical International Phonetics Programs (LIPP; Oller, 1991) software package to conduct infraphonological

TABLE 4. *Number of utterances analysed for each infant across the three sessions*

Group	ID	Multisyllabic utterances			Disyllabic utterances		
		Precan	Can	Postcan	Precan	Can	Postcan
Hearing	1	7	28	29	4	14	26
Hearing	2	22	19	27	17	11	10
Hearing	3	5	32	26	3	21	17
Hearing	4	11	13	14	9	8	12
Hearing	5	21	15	20	14	6	15
Hearing	6	28	18	18	23	12	14
Hearing	7	25	14	10	17	8	8
Hearing	8	15	20	13	9	8	8
Deaf	9	24	11	14	12	8	14
Deaf	10	11	5	2	5	4	2
Deaf	11	49	28	19	16	16	13
Deaf	12	5	13	7	4	10	6
Deaf	13	17	33	25	8	9	10
Deaf	14	1	24	17	1	21	7
Deaf	15	17	16	16	13	8	18
Deaf	16	33	17	24	21	7	15

coding of syllables in the utterances in order to determine canonical babbling ratios and thereby confirm level of vocal development. An additional trained judge coded a randomly selected subset of vocal samples (10% from each infant group) in order to establish reliability of the primary judge's codes. Cohen's kappa for interjudge agreement for infraphonological codes was 0.72.

Phase 3. Each selected session(s) contained at least 50 utterances. Although an attempt was made to include all utterances in the acoustic analysis, approximately one-fourth of all utterances that had low signal-to-noise ratios and had simultaneous vocalizations from another speaker had to be discarded because the acoustic features of such vocalizations were difficult to determine. The remaining utterances were low-pass filtered at 10 kHz and digitized at a sampling rate of 22.1 kHz using the Sound Blaster 16 card and CSpeech speech analysis software package (Milenkovic & Read, 1992) running on an IBM-compatible microcomputer. The number of utterances used for analysis for each infant across the three sessions is provided in Table 4. Amplitude displays (amplitude across time), wide-band spectrograms (300 or 600 Hz analysis filter), and, if necessary, narrow-band spectrograms (60 Hz analysis filter) were generated.

Identification of syllables and utterances was conducted using procedures outlined by Lynch *et al.* (1995) and Lynch (1996). Syllables were identified on the spectrograms and amplitude displays as occurrences of marked acoustic energy that were generally less than 500 ms in duration for normally hearing infants. Syllables typically contained either a vowel-like nucleus in isolation or a consonant- and vowel-like element in sequence, with either element

occurring first in the syllable. Utterances were identified by their syllabic content and were either monosyllabic or multisyllabic.

The duration of all syllables and utterances was then measured. A second judge measured the duration of a randomly selected set of vocal samples (10% from each infant group). The intraclass correlation coefficient for interjudge reliability was 0.99. The intraclass correlation coefficient was used because it is specifically designed to assess interjudge agreement on continuous variables, such as duration. Syllable duration measurements of the second judge were, on the average, within 6% of the original judge's measurements, indicating that the procedure was highly reliable.

RESULTS

Duration of all nonfinal syllables in multisyllabic utterances was compared to duration of final syllables of utterances. It should be noted that because multisyllabic utterances of all syllable lengths were combined in the analyses, syllable compression effects were confounded in the evaluations with FSL effects. Syllable compression effects refer to the reduction in duration of syllables that occurs when the number of syllables within a unit increases (e.g. Oller, 1973). In order to enable comparisons with previous investigations, additional sub-analyses were conducted on disyllabic utterances, where syllable compression effects are controlled. Table 4 provides the number of multisyllabic and disyllabic utterances analysed for each infant across the three sessions.

Experimental design

The data were analysed via hierarchical linear modelling. Fixed effects were position (nonfinal, final), session (precanonical, canonical, postcanonical), and hearing status (normally hearing, deaf), and their interactions. Random effects were infant (nested within hearing status) and session (nested within infant). The dependent measure was syllable duration. Syllable durations were weighted by number of utterances per cell because unequal numbers of utterances were obtained for each infant. The drawback of small sample sizes was, therefore, reduced. In addition, because residuals were not normally distributed (the data were skewed to the right) and cell standard deviations were approximately proportional to cell means, durations were log-transformed. In the transformed data, skew was greatly reduced and variance heterogeneity was mitigated as required by ANOVA. Table 5 presents mean weighted syllable durations of the log-transformed data, back-transformed to the original millisecond units (equivalent to weighted geometric means of the original data) and FSL ratios across the three sessions. The standard errors were obtained by propagating the errors from the log-transformed data. Table 6 presents unweighted arithmetic means and corresponding FSL ratios

TABLE 5. *Weighted geometric nonfinal and final syllable durations (in ms) and FSL ratios for normally hearing and deaf infants across sessions*

Group	Precanonical		Canonical		Postcanonical	
	Nonfinal	Final	Nonfinal	Final	Nonfinal	Final
Hearing						
<i>M</i>	368.021	472.891	343.001	394.61	311.685	338.638
S.E.	(50.469)	(85.479)	(20.863)	(38.607)	(39.461)	(51.632)
FSL ratio	1.28		1.15		1.09	
Deaf						
<i>M</i>	311.041	469.041	381.013	569.885	374.355	520.689
S.E.	(89.833)	(81.779)	(86.583)	(99.040)	(41.236)	(84.323)
FSL ratio	1.51		1.5		1.39	

TABLE 6. *Unweighted arithmetic nonfinal and final syllable durations (in ms) and FSL ratios for normally hearing and deaf infants across sessions*

Group	Precanonical		Canonical		Postcanonical	
	Nonfinal	Final	Nonfinal	Final	Nonfinal	Final
Hearing						
<i>M</i>	378.05	527.275	346.555	408	316.295	355.398
S.E.	(35.894)	(81.411)	(17.791)	(37.636)	(31.62)	(47.876)
FSL ratio	1.35		1.17		1.11	
Deaf						
<i>M</i>	448.957	497.089	457.63	598.433	360.089	771.853
S.E.	(78.665)	(46.173)	(41.23)	(77.241)	(21.297)	(250.221)
FSL ratio	1.26		1.3		2.12	

of the untransformed data (not used in the ANOVA) as a standard of comparison for prior studies wherein weighting and transformation have not been utilized.

Results for multisyllabic utterances revealed a significant interaction effect for hearing status \times position, $F(1, 42) = 10.27, p < 0.01$. *Post-hoc* pairwise comparisons revealed that durations of normally hearing and deaf nonfinal syllables were similar, $F(1, 14) = 2.95, p > 0.05$; deaf infants, however, produced longer final syllables than normally hearing infants, $F(1, 14) = 6.17, p < 0.05$. Nonfinal and final syllable duration values for normally hearing and deaf infants are provided in Figure 1.

There was a significant main effect for position, $F(1, 42) = 73.87, p < 0.01$, such that final syllables were longer than nonfinal syllables. FSL ratio for normally hearing infants was 1.17 and the FSL ratio for deaf infants was 1.46. FSL ratios for normally hearing infants decreased across the three

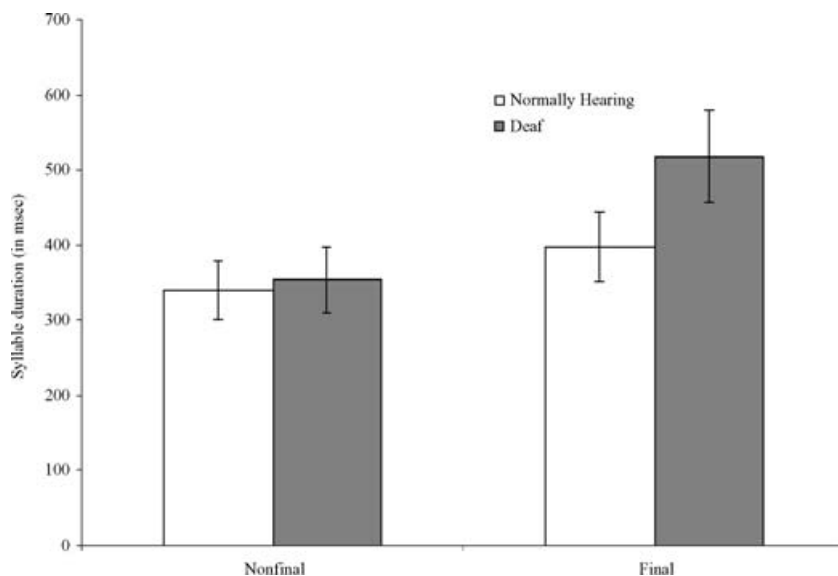


Fig. 1. Weighted geometric mean (and s.e.) nonfinal and final syllable durations (in ms) for normally hearing and deaf infants.

sessions whereas it was relatively consistent across sessions for deaf infants (see Tables 5 and 6). The main effect for hearing status, $F(1, 14) = 4.49$, $p = 0.05$, was marginally significant. Deaf infants ($M = 428.32$ ms, $S.D. = 146.35$ ms) produced longer syllables than normally hearing infants ($M = 368.07$ ms, $S.D. = 118.05$ ms). No interaction effect for hearing status \times session \times position or main effect for session was found. Within-group *post-hoc* analyses revealed a significant main effect for position for normally hearing infants, $F(1, 21) = 26$, $p < 0.001$, such that final syllables of normally hearing infants were longer than their nonfinal syllables. Deaf infants also showed significant FSL, $F(1, 21) = 53$, $p < 0.001$.

Results for disyllabic utterances revealed a similar pattern. Significant main effects were found for position, $F(1, 42) = 40.31$, $p < 0.01$, and hearing status, $F(1, 14) = 7.28$, $p = 0.02$. The hearing status \times position effect, $F(1, 42) = 3.84$, $p = 0.06$, was marginally significant, but given the smaller number of utterances in the disyllable analysis than in the full analysis, the less reliable effect is not surprising.

Individual data

Although group data revealed FSL across all three sessions, individual data revealed considerable variability in the magnitude of FSL. The mere occurrence of FSL may not be enough to draw definite conclusions about the

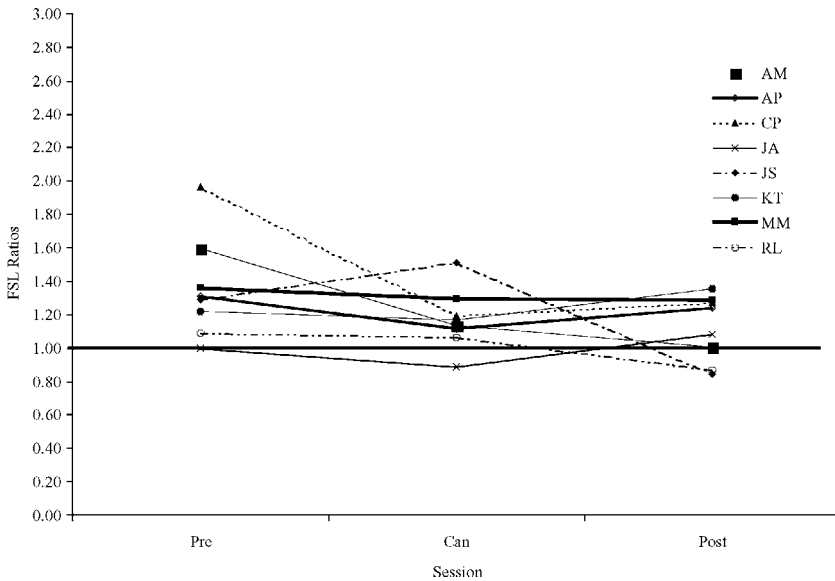


Fig. 2. Individual FSL ratios for normally hearing infants across sessions.

development of FSL. It is clear that FSL must be of some minimal magnitude in order to be perceptible or to achieve practical significance. Klatt (1976) provides data suggesting that in some circumstances a 20% difference from a base syllable duration, or an FSL ratio of 1.2, is necessary for perceptual significance in adult speech. Figures 2 and 3 display FSL ratios for normally hearing and deaf infants across the three sessions. Our figures show that the infant data varied widely around the 1.2 ratio. Even the 1.2 ratio, however, cannot provide a strict criterion of perceptual importance for a variety of reasons. First, Klatt (1976) detailed a series of qualifications about the circumstances of appropriate application of the 1.2 ratio, which may vary with rate of speech, stress patterning and so on. Second, infant syllables may not be treated auditorily as adult syllables in speech are. Furthermore, many other factors (such as F_0 and intensity) influence perception of durations and of boundaries (Klatt, 1976). Consequently, the 1.2 criterion needs to be viewed with caution.

As evident from Figure 2, not all infants showed FSL in every session. Two normally hearing infants, JA and RL, showed especially minimal (or nil) magnitudes of FSL across the three sessions. Of the remaining six infants, only three infants displayed the hypothesized U-shaped development of FSL. However, all six infants showed considerable FSL in precanonical sessions. There appeared to be a reduction in the magnitude and variability of FSL ratios in canonical sessions for normally hearing infants.

FSL IN INFANTS

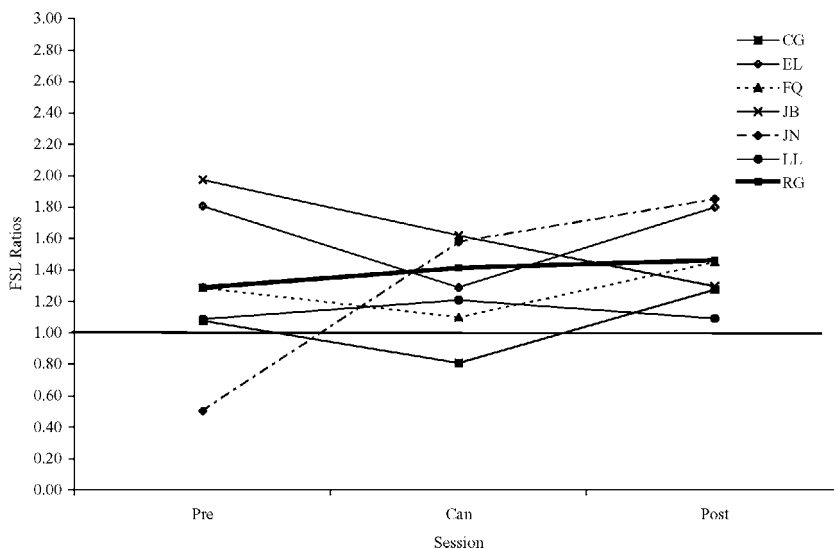


Fig. 3. Individual FSL ratios for deaf infants across sessions.

On the other hand, all deaf infants showed some FSL in at least one of the three sessions. It should be noted that JN, who appears to show no FSL in precanonical sessions and then a marked FSL in canonical sessions, produced only one multisyllabic utterance in the precanonical session thereby reflecting undersampling. It should also be noted that data from only seven deaf infants are displayed in Figure 3 because DH produced only 2 utterances in the postcanonical session with a corresponding FSL ratio > 6.0 , obviously a sampling artifact.

DISCUSSION

The purpose of our study was to assess possible origins of FSL and to examine whether a U-shaped developmental pattern of FSL occurred in the vocalizations of normally hearing infants in the first year of life. In addition, we sought to investigate whether the occurrence of FSL was influenced by auditory status and level of vocal development.

Statistical results indicated that reliable FSL was present throughout the first year of life for normally hearing infants when either all multisyllabic utterances or disyllabic utterances alone were used in the analysis. These results suggest that FSL may have biological roots and may, therefore, be motivated by a basic tendency to slow down at the ends of motoric sequences (Oller, 1973; Klatt, 1976). Further evidence in support of this claim is provided by the fact that reliable and auditorily perceptible FSL was present in

the utterances of deaf infants in the present study. Babbling is furthermore clearly related to other motoric phenomena, and its rhythmic nature very early in life, therefore, resembles that of other physical acts (MacNeilage & Davis, 1990; Thelen, 1991).

Even though FSL may be present in the first year of life, it may not achieve practical significance until it reaches a certain magnitude. Examination of the magnitude of FSL in the normally hearing group reveals that FSL was strongest in the precanonical sessions (FSL ratio = 1.28). The occurrence of strong FSL early in life supports the presumed initial biological basis for this phenomenon. Conclusions about the strength or a threshold criterion for FSL are, however, speculative as FSL is a highly complex phenomenon. Its values vary considerably even in adult speech, ranging from 60% lengthening to near absence in adult English speakers (Smith, 2000). With such large variations, it is clear that infants must face considerable difficulty in determining what the appropriate target level might be for final-syllable lengthening, assuming they notice it at all. In addition, as Klatt (1976) noted, the exact ratio at which FSL becomes perceptible varies under different sentence conditions. Therefore, a cutoff magnitude criterion for the practical significance of FSL, though desirable, is difficult to specify.

Although statistically significant session effects were not obtained, there was a trend toward reduction in the magnitude of FSL and in the variability of syllable durations with the onset of canonical babbling. This trend is consistent with many prior findings. For instance, Lynch *et al.* (1995) observed FSL in phrases in infants aged 0;2 to 0;4 and 0;10 to 1;0 but not in infants aged 0;6 to 0;8. Because canonical syllable production typically emerges around 0;6, studies that matched infants for age might have observed reduced FSL after 0;6 due to the production of canonical syllables by infants at this time. Canonical syllable production is frequently characterized as a rhythmic, stereotypic behaviour (Thelen, 1991). It is, therefore, plausible that the temporal regularity induced by the production of canonical syllables inhibited the occurrence of FSL such that the final syllable was equated in duration to nonfinal syllables. This might be especially evident in reduplicated sequences, where the same syllable is repeated over and over again. In fact, studies which investigated FSL in reduplicated sequences only, indeed support this claim as they noted nil to minimal FSL (e.g. Oller & Smith, 1977).

This possible influence of canonical syllable production on FSL also has interesting implications for models of prelinguistic vocal development. Previous models of speech and language development have posited unidirectional influences on vocal production such that the establishment of higher-order rhythmic units, e.g. utterances, is believed to promote the acquisition of lower-order rhythmic units, e.g. syllables, because the utterances can serve as frames into which syllables may be inserted (de Boysson-Bardies, Bacri, Sagart & Poizat, 1981). Findings from studies of perception (e.g. Jusczyk,

1998) support this unidirectional hypothesis because they suggest that prosodic cues are available to infants some months before they begin using more segmentally-based cues. Perceptual models assert that the chunks of the speech stream that result from the application of broader prosodic cues provide infants with more opportunities to detect segmental and syllabic cues that could then be used to achieve further segmentation within these chunks. Given present findings, however, it may be possible that influences are bi-directional, rather than unidirectional, such that advances at the syllabic level, namely canonical syllable production, are accompanied by reorganization at the higher-order utterance level, as reflected in the utterance-level reduction of FSL with canonical syllable production. Longitudinal observation of infants at closer intervals will also have to be conducted to document more precisely whether the changes that occur in rhythmic organization correspond with advances in vocal development at the syllabic level.

Based on findings from previous research, we had initially hypothesized a U-shaped development of FSL. Although the reduction of FSL with canonical babbling supported the downward trend of the hypothesized U-shaped curve, the continued attenuation of FSL in postcanonical sessions was contrary to our predictions that FSL would re-emerge later in infancy. It may be that the developmental range sampled in the present study was insufficient to show the predicted pattern strongly and examination of FSL later in life would reveal its re-emergence. The attenuation of FSL in canonical and postcanonical sessions and the absence of FSL in two infants' productions, however, emphasize individual differences, and imply that FSL may be a fleeting phenomenon in infancy. Exposure to the ambient language may be necessary before FSL becomes crystallized in production as evidenced by reports in the literature of cross-linguistic differences in the extent of FSL in the productions of older infants. For instance, FSL appears to be more pronounced in French than in English (Vihman & DePaolis, 1998). It may be that rhythmic organization can proceed in a variety of ways and depending upon the ambient linguistic environment, the initial biological aspects of rhythm may come to be weighted differently such that one language might favour one aspect, e.g. stress in English, whereas another language might emphasize another aspect, e.g. FSL, as in French. Exactly when and how FSL is influenced by the ambient language will have to await more systematic, cross-linguistic investigations of prelinguistic vocalizations to the one-word stage and beyond. Large-scale studies of normally hearing infants are also warranted in order to verify the developmental attenuation of FSL and the extent of individual differences. Values vary considerably even for adult speakers (Smith, 2000) and are influenced by several factors (Klatt, 1976). Consequently, some variability in FSL values is to be expected even in infancy.

Reliable differences in FSL across infants and sessions were, however, not noted for deaf infants. FSL was present in all three sessions for most infants, as

evidenced in Table 5 and Figure 3. In addition, the extent of FSL was greater in the deaf infants than in the hearing infants. Because the deaf infants were older than the normally hearing infants in the study (having been matched on developmental level), the data suggest that the constraints of stereotypic durations induced by canonical syllable production may not apply in a physiologically more mature system. It may also be that because, at the time of the precanonical sessions, the deaf infants had used FSL for a longer time than normally hearing infants, and since FSL was more established, it was more resistant to change with the onset of canonical syllable production than in the hearing infants. Finally, it may be that the temporal regularity associated with canonical syllable production is primarily a byproduct of perception rather than production characteristics. Because deaf infants lacked auditory experience, perhaps they could not (as the hearing infants apparently had done) use auditory information to manipulate and thereby reduce final syllable durations to achieve temporal regularity at the onset of canonical babbling.

Some evidence that audition influences the manifestation of rhythmic stereotypies is provided by recent findings of Ejiri (1998), who noted that rhythmic hand shaking of a rattle continued or increased with an audible rattle but not with an inaudible rattle. The so-called rhythmicity of canonical babbling may be similarly dependent upon perceptual experience. Rhythmic stereotypies that tend to result in equal durations of canonical syllables may be necessary in order to normalize syllable durations so they resemble mature speech. This normalization may take precedence over rhythmic organization of larger units, such as utterances, in normally hearing infants, and thus cause FSL to abate with the onset of canonical syllables. Because deaf infants lack auditory experience to regulate syllable durations, it may be difficult for them to override the inherent tendency toward rhythmic organization of serially ordered productions. Deafness, therefore, not only results in a delay in the onset of canonical syllable production (e.g. Oller & Eilers, 1988) but may also influence the qualitative aspects of canonical babbling when it eventually emerges.

A caveat should be noted about our findings regarding deaf infants in that the deaf infants in the present study were enrolled in extensive intervention programs and this may have influenced the findings. It is possible that the presence of FSL and/or some of its characteristics were influenced by intervention modelling or patterns of reinforcement in therapy. In addition, vocalizations from deaf infants were recorded when they were interacting with the speech-language pathologist during intervention sessions. In more naturalistic interaction situations, the quality of vocalizations might be different. Finally, deaf infants were older than normally hearing infants and may have acquired FSL as a result of their longer listening or lipreading experience or greater physical maturation. Younger deaf infants should be included in future investigations in order to determine the effects of age, length of auditory

experience, and intervention on rhythmic organization. Recent initiatives on universal newborn hearing screening should enable identification of deaf infants at younger ages.

One significant finding was that nonfinal syllable durations of deaf infants were similar to those of normally hearing infants (refer to Figure 1). This finding may be very important because it suggests that longer durations of rhythmic units that have been reported for deaf individuals (e.g. Calvert, 1961) may not always occur because these individuals have general temporal difficulty with duration but may reflect difficulty with relative temporal patterning. The final syllables of deaf speakers may be much more discrepant from those of hearing speakers than the non-finals. These results should however, be tempered because findings of Lynch (1996) indicate that both absolute and relative temporal patterning were affected in a single acochlear child. Furthermore, nonfinal syllable duration tended to drop in canonical and postcanonical sessions for normally hearing infants as seen in Table 5. Deaf infants, on the other hand, showed a slight increase in nonfinal syllable durations in canonical sessions suggesting some deficits in absolute temporal patterning as well. In any event, present findings suggest that it might be wise to target relative temporal patterning in addition to mere production of canonical syllables in intervention protocols with deaf infants.

In conclusion, our findings show that FSL appears to be biologically motivated at least initially. Results support only the downward trend of the U-shaped developmental function of FSL that has been noted in evaluation of the complex results from many prior studies on FSL in infancy. The FSL seen in precanonical utterances appears to be attenuated at the point of advances in canonical syllable production in the normally hearing infant. In the absence of normal auditory perception, the initial FSL appears to persist and thereby seemingly results in a different pattern of syllable durations than is found in hearing infants. Verification of these conclusions, however, requires large-scale cross-linguistic studies of FSL beyond the one-word stage and examination of younger deaf infants. In addition, closer longitudinal observations of the relationship among perception, FSL, and canonical syllable production are needed. Finally, examination of other aspects of rhythm, e.g. falling final intonation or stress, is necessary in order to validate these preliminary observations regarding the developmental trajectory of FSL and correlated rhythmic phenomena.

REFERENCES

- Allen, G. D. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics* 3, 75–86.
- Baum, S. R. & Boyczuk, J. P. (1999). Speech timing subsequent to brain damage: effects of utterance length and complexity. *Brain and Language* 67, 30–45.
- Boysson-Bardies, B. de, Bacri, N., Sagart, L. & Poizat, M. (1981). Timing in late babbling. *Journal of Child Language* 8, 525–39.

- Calvert, D. R. (1961). Some acoustic characteristics of the speech of profoundly deaf individuals. Ph.D. Dissertation, Stanford University.
- Cummins, F. (1999). Some lengthening factors in English combine additively at most rates. *Journal of the Acoustical Society of America* **105**, 476–80.
- Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* **4**, 183–98.
- Ejiri, K. (1998). Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica* **55**, 226–37.
- Erickson, M. L. (2000). Simultaneous effects on vowel duration in American English: a covariance structure modeling approach. *Journal of the Acoustical Society of America* **108**, 2980–95.
- Halle, P. A., Boysson-Bardies, B. de & Vihman, M. M. (1991). Beginnings of prosodic organization: intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* **34**(4), 299–318.
- Hoequist, C. J. (1983). Syllable duration in stress-, syllable-, and mora-timed languages. *Phonetica* **40**, 203–37.
- Jusczyk, P. W. (1998). Constraining the search for structure in the input. *Lingua* **106**, 197–218.
- Kent, R. D., Osberger, M. J., Netsell, R. & Hustedde, C. G. (1987). Phonetic development in identical twins differing in auditory function. *Journal of Speech and Hearing Disorders* **52**, 64–75.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of Acoustic Society of America* **59**(5), 1208–21.
- Konopczynski, G. (1993). The phonological rhythm of emergent language: a comparison between French and English babbling. *Working Papers in Linguistics* **18**, 1–32.
- Levitt, A. G. & Utman, J. G. A. (1992). From babbling towards the sound systems of English and French: a longitudinal two-case study. *Journal of Child Language* **19**, 19–49.
- Levitt, A. G. & Wang, Q. (1991). Evidence for language-specific rhythmic influences in the reduplicative babbling of French and English learning infants. *Language and Speech* **34**(3), 235–49.
- Lynch, M. P. (1996). The case of total deafness II: phrasing in the prelinguistic vocalizations of a child born with congenital absence of cochleas. *Applied Psycholinguistics*, 1–31.
- Lynch, M. P., Oller, D. K., Steffens, M. L. & Buder, E. H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychology* **28**(1), 3–25.
- MacNeilage, P. F. & Davis, B. (1990). Acquisition of speech production: frames, then content. In M. Jeannerod (ed.), *Attention and performance XIII: motor representation and control*. Hillsdale, NJ: Erlbaum.
- Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review* **79**, 487–509.
- Milenkovic, P. & Read, C. (1992). CSpeech Version 4. Department of Electrical Engineering, University of Wisconsin, Madison.
- Monsen, R. B. (1979). Acoustic qualities of phonation in young hearing-impaired children. *Journal of Speech and Hearing Research* **22**, 270–88.
- Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* **54**, 1235–47.
- Oller, D. K. (1979). Syllable timing in Spanish, English, and Finnish. In H. Hollien & P. Hollien (eds), *Current issues in the phonetic sciences*. Amsterdam: John Benjamins.
- Oller, D. K. (1991). Computational approaches to transcription and analysis in child phonology. *Journal for Computer Users in Speech and Hearing* **7**, 44–59.
- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.
- Oller, D. K. & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development* **59**, 441–9.
- Oller, D. K. & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology. In L. M. C. A. Ferguson and C. Stoel-Gammon (eds), *Phonological development: models, research, implications*. Timonium, MD: York Press.

- Oller, D. K. & Smith, B. L. (1977). Effect of final-syllable position on vowel duration in infant babbling. *Journal of the Acoustical Society of America* **162**(4), 994–7.
- Robb, M. P. & Saxman, J. H. (1990). Syllable durations of preword and early word Vocalizations. *Journal of Speech and Hearing Research* **33**, 583–93.
- Selkirk, E. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press.
- Smith, B. L. (2000). Variations in temporal patterns of speech production among speakers of English. *Journal of the Acoustical Society of America* **108**, 2438–42.
- Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech and Hearing Research* **37**, 831–40.
- Stark, R. E. (1972). Some features of the vocalizations of young deaf children. In J. F. Bosma (ed.), *Third symposium on oral sensation and perception: the mouth of the infant*. Springfield, IL: Charles C. Thomas.
- Stark, R. E. (1989). Temporal patterning of cry and non-cry sounds in the first eight months of life. *First Language* **9**, 107–36.
- Steffens, M. L., Oller, D. K., Lynch, M. & Urbano, R. C. (1992). Vocal development in infants with Down Syndrome and infants who are developing normally. *American Journal of Mental Retardation* **97**(2), 234–46.
- Thelen, E. (1991). Motor aspects of emergent speech: a dynamic approach. In N. Krasnegor (ed.), *Biobehavioral Foundations of Language*. Hillsdale, NJ: Erlbaum.
- Vihman, M. M. & DePaolis, R. (1998). Perception and production in early vocal development: evidence from the acquisition of accent. *Chicago Linguistic Society: The Panels* **34**, 373–86.
- Weismer, G. & Ingrisano, D. (1979). Phrase-level timing patterns in English: effects of emphatic stress location and speaking rate. *Journal of Speech and Hearing Research* **22**, 516–33.
- Zlatin Laufer, M. (1980). Temporal regularity in prespeech. In T. Murry & J. Murry (eds), *Infant communication: cry and early speech*. San Diego, CA: College Hill Press.