Registers in Infant Phonation

Eugene H. Buder, Valerie F. McDaniel, Edina R. Bene, Jennifer Ladmirault, and D. Kimbrough Oller, *Memphis, Tennessee*

Summary: The primary vocal registers of modal, falsetto, and fry have been studied in adults but not *per se* in infancy. The vocal ligament is thought to play a critical role in the modal-falsetto contrast but is still developing during infancy (Tateya and Tateya, 2015).[41] Cover tissues are also implicated in the modal-fry contrast, but the low fundamental frequency (f_o) cutoff of 70 Hz, shared between genders, suggests a psychoacoustic basis for the contrast. Buder, Chorna, Oller, and Robinson (2008)[6] used the labels of "loft," "modal," and "pulse" for distinct vibratory regimes that appear to be identifiable based on spectrographic inspection of harmonic structure and auditory judgments in infants, but this work did not supply acoustic measurements to verify which of these nominally labeled regimes resembled adult registers. In this report, we identify clear transitions between registers within infant vocalizations and measure these registers and their transitions for f_o and relative harmonic amplitudes (H1-H2). By selectively sampling first-year vocalizations, this manuscript quantifies acoustic patterns that correspond to vocal fold vibration types not previously cataloged in infancy. Results support a developmental basis for vocal registers, revealing that a well-developed ligament is not needed for loft-modal quality shifts as seen in harmonic amplitude measures. Results also reveal that a distinctively pulsatile register can occur in infants at a much higher f_o than expected on psychoacoustic grounds. Overall results are consistent with cover tissues in infancy that are, for vibratory purposes, highly compliant and readily detached.

Key Words: Infant vocal development–Registers–Falsetto–Modal–Vocal fry.

INTRODUCTION

At least three primary registers are agreed to occur in adult phonation, known most widely as modal, falsetto, and fry.[1] Modal, which is most typical, is the default register for speaking and singing in habitual fundamental frequency (f_o) range, corresponding to "chest" voice in singers. Falsetto occurs in higher f_o ranges, corresponding to "head" voice in singers. Vocal fry occurs in the lowest f_o ranges, also being called "creaky," "laryngealized," or "glottalized" voice by linguists who have observed this phonatory contrast to be phonemic in some languages.[2,3] In these contexts, registers are controlled adjustments of phonation and may be assumed as such to be well-controlled behaviors. However, harmonic and waveform patterns auditorily resembling falsetto and fry, along with other nonmodal patterns such as subharmonics, have been observed in infancy[4,5] and, in this context, can be regarded as the naturally and spontaneously occurring regimes of an inherently nonlinear dynamic system.[6–8]

Švec et al[9] contrast the spontaneity of natural "bifurcations" that are to be expected in highly nonlinear dynamic systems such as phonation with what they characterized as the "prevailing opinion" (of that time) that a muscular adjustment is required for register shifts. In a nonlinear dynamic bifurcation, sudden changes in vibratory regime may occur abruptly as a control parameter (such as longitudinal vocal fold tension or subglottal pressure) is smoothly varied. The richness of vibratory regimes studied by Buder et al[6] is consistent with the view that the bifurcation principle is operative in infant phonation (although underlying mechanisms and parameters remain unresolved).

Explication of levels of analysis and associated terminology can help maintain clear conceptual distinctions between vibratory mechanisms, acoustic outputs, and auditory qualities. Phonation can be characterized as resulting from interactions between airflow and vocal fold vibrations. These interactions typically correspond to clear acoustic effects. Moreover, as voice scientists also classify vocal outputs based on auditory qualities, the adjective "phonatory" is often applied to that level as well. Scientific study of adult registers has been conducted at all three levels but seems to have been rooted primarily in auditory impressions in the context of acoustic and mechanistic investigations. In the present article, we will also use the term register as an auditory-acoustic construct intended to inform understanding of the underlying phonatory mechanisms, and thus the term "phonation" here is intended to encompass both vibratory and acoustic phenomena. Along similar lines in our usage, the term "vocal quality" originates from the auditory domain with concomitant acoustic markers, pointing toward potentially distinct phonatory mechanisms. "Regimes" refers to the broadest set of vocal fold vibration possibilities as distinguished primarily via inspection of the harmonic structure of phonatory output.

In the present article, we are most interested in register transitions as presumably spontaneous phenomena in the phonatory exploration that characterizes much of vocal development in the first year of life.[10] Although falsetto and fry are the most widely used terms for nonmodal registers in contemporary literature, in the context of infant phonatory types, it seems preferable to return to the terms originally coined by Hollien[11]—"loft" and "pulse." These terms also connote acoustic bases more directly and, as such, are relatively agnostic with respect to underlying mechanisms. These terms also leave open

Accepted for publication December 20, 2017.

Eugene H. Buder and D. Kimbrough Oller received support for this research from Grants R01 DC006099 and DC011027 from the National Institute on Deafness and Other Communication Disorders. Oller's role in this research was also founded by an endowment from the Plough Foundation, which supports his Chair of Excellence at the University of Memphis.

From the School of Communication Sciences and Disorders, The University of Memphis, Memphis, Tennessee.

Address correspondence and reprint requests to Eugene H. Buder, School of Communication Sciences and Disorders, The University of Memphis, 4055 N. Park Loop, Memphis, TN 38111. E-mail: ehbuder@memphis.edu

Journal of Voice, Vol. 33, No. 3, pp. 382.e21–382.e32
0892-1997

https://doi.org/10.1016/j.jvoice.2017.12.013

the possibility that occurrence of similar phonation types in infancy are distinct from those used intentionally for performance and linguistic purposes.

Although not entirely conclusive, a well-developed literature on registers in adult phonation has established some consensus on the underlying mechanisms that distinguish the primary nonmodal registers. Loft is generally thought to involve lateralization of the muscular body of the vocal folds and a decoupling of the cover tissues from the body, yielding a relatively small vertical contact area to participate in vibration.[1] The resulting voice quality is therefore not only higher in f_o but also relatively thin sounding, with a conspicuously higher amplitude of the first harmonic relative to the second and overall greater spectral slope. Pulse, which is lower in f_o than modal but also with a distinctive "rasping" quality, is generally thought to involve greater activation of the body with a lax cover,[1] and is associated with reduced airflow and a longer closed time.[12]

The study of registers in infancy is of particular interest because of differences between adult and infant vocal fold composition. Developmental studies of the vocal mechanisms have reinforced as well as expanded the many ways in which we know that the infant is not just a smaller version of an adult vocalizer.[13] The vocal fold tissues themselves differ substantially in the first year of life from those of the mature adult, the lamina propria being relatively undifferentiated and the macula flavae mostly undeveloped.[14] Although this topic has been investigated fairly systematically in the anatomic and medical literature, revealing a protracted schedule of tissue differentiation that begins fetally but extends into adolescence,[15–18] there are very few studies that have analyzed noncry phonation with these tissue mechanics in mind; see, however, Fuamenya et al[19] for recent work in the context of crying.

The cover-tissue difference in infants is particularly relevant to questions regarding registers; in adults, both loft and pulse purportedly involve cover-body configurations that are quite different from those used to produce modal. Specifically, cover-body decoupling in loft phonation is thought to involve the ligamental nature of the intermediate and deep layers of the lamina propria. However, the distinctions between these connective layers are essentially lacking in the cover tissues of infant vocal folds. Pulse is thought to be associated with a lax cover in adults, yet the cover is particularly compliant in infancy. A basic theoretical question therefore arises in the search for registers in infant phonation. Given the presumed roles of cover tissues in loft and pulse register production, and the known differences in infant vocal fold cover tissues, do registers occur in infancy, and if they do, how do they differ from adult registers? This question requires acoustic criteria for the nonmodal registers.

F_o is a direct acoustic correlate of the two nonmodal registers when produced by mature speakers but does not by itself distinguish those registers from modal phonation. Infants do not produce registers on demand, so criteria other than f_o are needed. Buder et al[6] identified harmonic spacing patterns consistent with pulse and loft f_o ranges but did not provide independent acoustic markers. Physiological register shifts between modal and falsetto may occur at different f_o values depending on the individual, often with rather large pitch jumps.[9] As pitch may shift without a concomitant register shift, distinctive vocal quality measurements are needed.

Compared with loft, pulse in adult phonation is more clearly distinguished from modal by f_o alone, transitioning remarkably at the same value of c. 70 Hz for both men and women.[1] Although investigators tend to agree that pulse is produced by a distinct vibratory mechanism,[1] psychoacoustics alone could explain why 70 Hz is the value at which individual glottal pulses become audibly distinct.[1] For infancy, in which f_o values at transitions to pulse have not previously been determined, and because salient pitch shifts to pulse have not been reported as they have for modal-loft transitions, it is again important to consider additional acoustic measures in conjunction with f_o. A defining feature of pulse phonation is the critical damping of each glottal pulse excitation before the next pulse occurs, causing a distinctive waveform shape marked by the appearance of temporal gaps between glottal pulses.[1] This waveform-shape criterion operates independently of harmonic amplitudes and thus can serve as an independent feature that may signal pulse register in infancy.

Relative amplitudes of the first two harmonics have been considered as indicating glottal status in registers, most definitively for loft.[20,21] Relative harmonic amplitude measures, including the difference between the first and the second harmonics, were introduced in the empirical literature by Hanson[22] as a reflection of the relatively open glottal configuration found in female voices: first harmonics were higher amplitude relative to the second in women.[22] More recent studies indicate that the amplitudes of the first two harmonics are sensitive not only to open quotient and breathiness but also to a variety of phonation types, including register.[23] Examination of the literature, in general, supports the use of both f_o and H1-H2 to distinguish both between modal and loft and between modal and pulse registers.

Table 1 summarizes previous studies that reported either transitional f_o or representative H1-H2 values for registers in nondisordered speakers. The literature employing these measures for studies of register is surprisingly spotty, with *no* adult studies having characterized both registers together in terms of both measures. Although scientists developing theories of voice production *per se* are clearly interested in all possible manners of voice production,[1,28] the phenomena of loft and pulse registers are treated in applied studies as distinct domains: loft (ie, falsetto) has been of primary interest in singing, and normative pulse (ie, "creak," "laryngealization," or "glotttalization") has been investigated as signaling phonemic contrasts, or marking phrasal boundaries,[29] in speech. Speaking and singing are distinct manners of voice production, so it also makes sense to study them distinctively.

Although acoustic measures may corroborate distinct registers in opportunistically observed vocalizations, identification of registers in such materials may begin with auditory judgments. Abrupt pitch shifts have been widely observed in infant phonation,[4,5] but not all of them are necessarily associated with register transitions, so auditory and acoustic classifications of distinctive qualities across such shifts remain necessary. Similarly, "growl" vocalizations are often marked by low-frequency pulsing,[6,30] but precise comparisons in acoustic terms across transitions are needed to demonstrate whether and in what ways these

TABLE 1.
References Reporting f_o and H1-H2 Values for Modal vs Nonmodal Registers

Citation	Population	f_o (Hz)	H1-H2 (dB)
		Modal vs loft	
Neiman et al[20]	Adult males	Loft: ≥250, modal: lower	Loft: H1 > H2 (in 23 of 25), modal: H1 < H2 (in 18 of 25)
Salomão and Sundberg[21]	Adult males singing	Various, with regions of overlap	H1 > H2, but with 14.2-dB greater difference in loft than in modal*
Švec et al[9]	Excised male larynx (aerodynamically excited)	Modal to loft†: 168 ↗ 332 Loft to modal†: 234 ↘ 146	
		Modal vs pulse	
Gordon and Ladefoged[2]	Male speaker of Zapotec		Pulse: H1<H2, modal: H1>H2
McGlone[24]	Adult males	Pulse: 34–51	
	Adult females	Pulse: 28–49	
McGlone and Shipp[25]	Adult males	Pulse: 18–65, modal: 87–117	
Blomgren et al[26]	Adult males	Mean pulse = 49, mean modal = 117	
	Adult females	Mean pulse = 48, mean modal = 211	
Avelino[27]	Three adult males		Modal: H1-H2 > pulse H1-H2 (in two of three)‡
	Three adult females		Modal: H1-H2 > pulse (in three of three)

* Measured from the spectrum of the inverse filtered flow glottogram.
† Note that the loft value after the upward jump is higher than the loft value before the downward jump, and vice versa for modal values, indicating a "bi-stable" region.
‡ Female values all positive, male values all but one (pulse) negative.

qualities and apparent registers are categorically distinct. Our approach emphasizes that within-vocalization register transitions specifically satisfy two scientific concerns raised by difficulties in observing infant phonatory patterns: (1) listening across a transition optimizes auditory identification, and (2) some of the possible physiological variables affecting register shifts may be held constant across the transition. Under these conditions, significant changes in acoustic quantities optimally support the premise that registers are driven by distinctive phonatory mechanisms.

In summary, the main goal of the present study was to document and quantify distinctive registers in infancy by investigating the nature of these registers as produced by immature vocal folds. Again, to our knowledge the investigation of both loft and pulse using basic acoustic metrics in one study has not previously been published even with adult subjects, much less with children or with infants. The theory of adult phonation suggests that because of the absence of a vocal ligament in infants, clear breaks between modal and loft are unlikely to occur at this stage of development, and there is also little basis on which to predict that infants will exhibit a clearly distinct pulse register at any specific frequencies. Secondarily, acoustic quantification of the observed registers will enhance metrics used to classify infant phonation into distinct types, helping to associate them with regimes and distinct vocal qualities. This latter goal will support yet unfulfilled objectives in studies of infant vocal development, specifically identifying the physiological and acoustic bases for the protophone categories purported to form a key basis (that of systematic contrastive sound production) for subsequent language development.[31–33]

MATERIALS AND METHODS

Participants and recordings

Two 20-minute recording sessions for each of three typically developing female infants were selected for coding at each of three ages: "early" (3–4 months), "mid" (5–7 months), and "late" (9–11 months) for a total of 18 sessions (360 minutes). Infants were both video and audio recorded in a laboratory equipped with four cameras and set up as a child's playroom. For representativeness, recordings were paired, one from a session in which the mother and infant were freely interacting and the other during a mother-experimenter interview session in which infants were otherwise often vocally active "separately."

The infants were fitted with custom-built vests that housed a wireless microphone system (Samson Airline UHF AL1 transmitter (Samson Technologies, Hicksville NY), equipped with a Countryman Associates low-profile low-friction flat frequency response MEMWF0WNC capsule (Countryman Associates, Menlo Park, CA), sending to a Samson UHF AM1 receiver). The vest configuration followed an original design developed by Buder and Stoel-Gammon,[34] with the microphone capsule housed within a Velcro patch and oriented to maintain the mouth-to-microphone distance at approximately 10–15 cm. *TF32* software[35] operating a DT321 acquisition card (Data Translation, Inc., Marlborough, MA) was used to digitize the infant signals at 48 kHz after low-pass filtering at 20 kHz via an AAF-3 anti-aliasing filter board.

Materials

Recording sessions were analyzed in the *Action Analysis, Coding, and Training* (*AACT*) environment,[36] which presents synchronized video and audio and allows users to demarcate and label intervals on spectrographic displays during coding. All noncry, nonlaugh, and nonvegetative protophone vocalizations[30,33] meeting minimal audibility and duration criteria (>50 milliseconds) were coded in breath group units.[37] These vocalizations were subsequently coded into intervals (referred to below as "segments") representing the following mutually exclusive and exhaustive phonatory regimes[6]: modal (clear, parallel, and moderately spaced harmonics), high modal (more widely spaced harmonics or an audible falsetto quality), pulse (very closely spaced harmonics, widely spaced glottal pulses, and a "zipper-like" sound), subharmonics (lower amplitude harmonics appearing in between main harmonics), biphonation (two different sets of harmonics moving in nonparallel directions), closed stops (within-vocalization adduction-caused gaps in phonation), open stops (within-vocalization abduction-caused gaps in phonation), and chaos (very unclear harmonic structure with aperiodicity in glottal pulses); see Buder et al[6] for more detailed definitions and examples.

Of special interest in the current investigation, the high modal code was a stand-in for a *possible* loft register: coders were trained only to mark very high-pitched intervals *or* intervals in which the thin and weak quality of a nonmodal "falsetto" voice was salient. As the existence of a true loft register in infancy had not been verified in prior literature, no training was provided to certify that regime coders could reliably distinguish loft from modal (but we note that the adult literature generally seems to lack that certification as well). Hence, the HiModal codes in our dataset included cases that might have been judged to involve loft if such a code had been included in the coding protocol, but the codes also included many segments that were high-pitched but still perceived to have been produced by the same mechanism as lower pitched modal segments. Pulse was deemed to be a more salient quality in infancy, reliably distinguished on the basis of close harmonics, critically damped glottal pulses, and a specifically raspy "zipper" or "frog" quality.[6] Pulse regimes mark growl protophones, but so can other low-f_o or even mid-f_o "rough" regimes, such as subharmonics, biphonation, and chaos.[30]

In all, 2445 noncry, nonvegetative, and nonlaugh protophone vocalizations were identified and coded for regimes. From this corpus, an exhaustive search was made, via inspection of spectrograms and listening, for vocalizations in which quality transitions between modal and "high modal" codes or between modal and pulse codes were audible within the utterance. Transitions could be in either order but with no intervening regimes. The following paragraphs detail how we inspected for markers of loft and pulse, independently of the f_o and H1-H2 measurements that would then be used to characterize the registers, to corroborate whether the previous regime coding boundaries represented a register transition and where that transition occurred.

For modal↔ loft transitions, only those in which there was a perceptually very salient difference in vocal quality and spectral tilt across the transitions were retained for analysis. Transitions were marked accordingly. For modal↔ pulse transitions, clear temporal gaps between glottal pulses should be observed characterizing the pulse segments.[1] Although waveform shape was one consideration in the previous pulse regime coding, a screening criterion for such temporal gaps had not been utilized.

For this purpose, Root Mean Square (RMS) amplitude contours were extracted with a very narrow 0.5-millisecond window to inspect for "temporal gaps" between glottal pulses (see Figure 1). Inspection of these contours was a primary consideration but was not always sufficient: variations in overall vocal intensity and noise background had to be taken into account when judging where the RMS amplitude had reached the ambient floor, and in such cases, waveform morphology could also be a consideration. Furthermore, as is especially conspicuous in Figure 1B, an interim "transitional" segment was sometimes observed in which the temporal gaps were lower in amplitude than in the adjacent modal segment but did not actually reach the amplitude. Most often, such gaps between modal and pulse segments were well under 50 milliseconds.

Figure 2 provides a schematic to illustrate where acoustic measures were extracted relative to the identified transitions. Measurement locations differed somewhat between the two register types because of the transitional gaps between modal and pulse that were not observed in modal↔ loft transitions. A primary question regarding f_o is the size of the shift between registers; for this question, it was important to make f_o measurements without variable gaps between registers. On the other hand, as H1-H2 measures were applied here to characterize the registers as such, it was important to accommodate transitional gaps between modal and pulse (especially in cases such as that illustrated in Figure 1B). In summary, all measures were taken as closely "adjacent" to the transitions as possible except for H1-H2 in pulse, which was taken as "near" to transitions as possible while still ensuring that the measures clearly represented pulse phonation.

Employing the criteria identified in the preceding three paragraphs for modal↔ nonmodal transitions, analysts endeavored to find five clear examples of each transition type from each of the three infants at each of the three ages. However, sometimes fewer than five, or even just one good example, could be found to represent an infant or age. The resulting sample is detailed in Table 2 of the Results section.

The *AACT* software was used for extracting harmonic amplitude values. *AACT* implements the *TF32* acoustic analysis library for extracting parameters such as harmonic frequencies and amplitudes, given an initial f_o cursor placement in a (optionally frequency-zoomed) narrowband spectrum. The *AACT-TF32* software automatically generates the following seven values for harmonic amplitude extraction: the time associated with the cursor placement; the fundamental frequency (regional f_o as determined by frequency cursor placement); the regional-peak

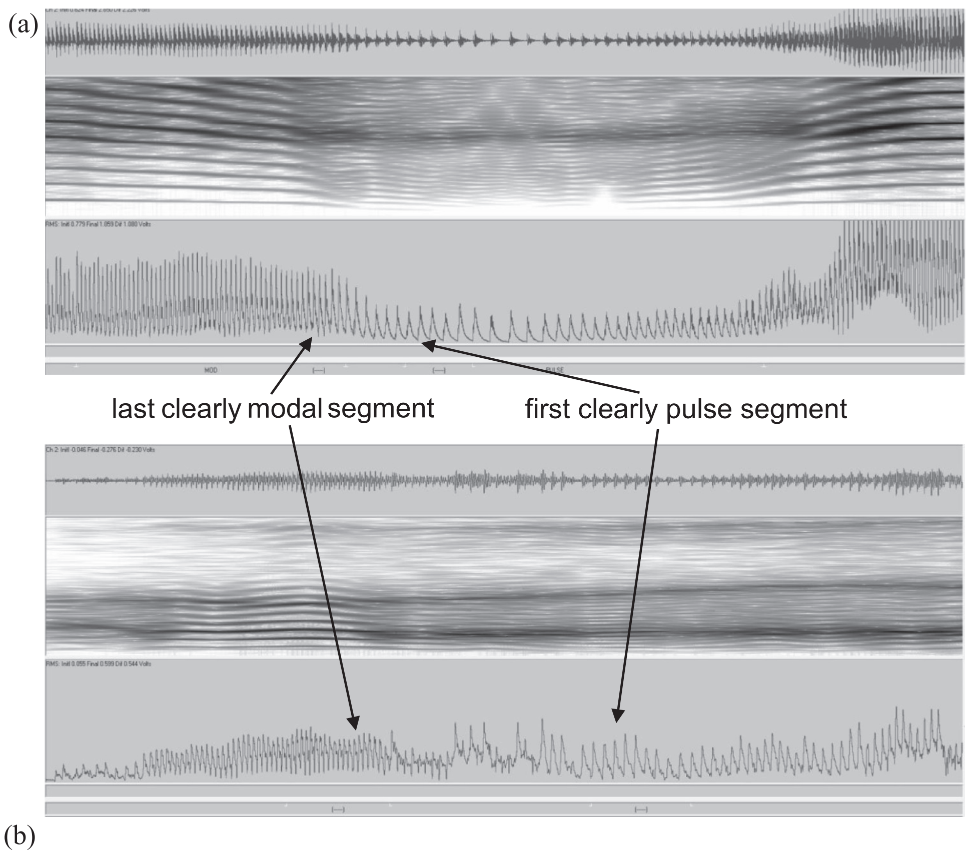

FIGURE 1. Illustrations of RMS procedure for modal-pulse discrimination. Each panel, extracted from *TF32*, displays the waveform on top (c. 0.5 second), a narrowband spectrogram (c. 0–4 kHz), and RMS amplitude with a 0.5-millisecond window. Label displays beneath the amplitude trace include "(————)" codes indicating the segments of interest, also indicated by arrows. Panel **A** presents a particularly clear example with only a brief transition gap, whereas panel **B** presents the most difficult case encountered in this corpus, with an extended transition due to brief, intervening regimes and an unclear amplitude floor due to noise overlay. The samples are from different infants but both at the youngest age represented in the data.

frequency and amplitude for the first harmonic, H1 (taken as f_o for this study), the regional-peak frequency and amplitude close to double H1, which was taken to be H2; and finally, H1-H2 as the amplitude of H2 subtracted from the amplitude of H1. The harmonic spectrum was visually inspected to confirm the accuracy of each extraction. Materials with ambiguous harmonic structure due to overlapping vocalizations, noise, or unclear harmonic structure were measured in consultation with coauthors until consensus was reached. In many cases, problems with identification of harmonics occurred because of very low intensity or very low frequency, causing either or both of the first harmonics to fall below the spectral amplitude floor. Such cases were excluded from the final data. These problems account for somewhat reduced numbers of observations in some categories, especially for pulse H1-H2 measures. H1-H2 measures were not transformed according to vowel identity as in the H1*-H2* measures proposed by Hanson[22] primarily because formant measurements are notoriously difficult in infant materials because of sparse harmonic sampling, nasality, and a host of other issues.[38] We note, however, that prior literature on registers in adult phonation has also presented raw H1-H2 measures (see Table 1).

During training of two analysts for H1-H2 measures from the entire corpus of regime segments identified as loft or modal, a large sample for assessing reliability was obtained, demonstrating good intercoder correlation for H1-H2 (Pearson $r(353) = 0.90$). A third analyst placed location markers for the pulse measures; reanalysis of 20% of the pulse-associated measurement locations by a reliability coder revealed a strong yet somewhat lower correlation (Pearson $r(84) = 0.84$). Intercoder reliability was also obtained on 20% of the data for the location of pulse↔modal transition locations using the RMS-based technique; coder differences, on average, differed by less than 3 milliseconds (standard deviation: 27 milliseconds).

As a first approach to the topic with no specific hypothesis-driven framework in terms of such measures in infant register samples, the data are summarized descriptively by univariate statistics (*t* test, analyses of variance [ANOVAs]) on the measures and examined within groups as appropriate (eg, across transitions). Nonetheless, even with a modestly sized dataset and large overall variability, basic results of interest were significant in the third-order alpha level, $P < 0.001$, so we have reported all effects with alpha level 0.05 or less, and no "number of tests" adjustments have been applied.

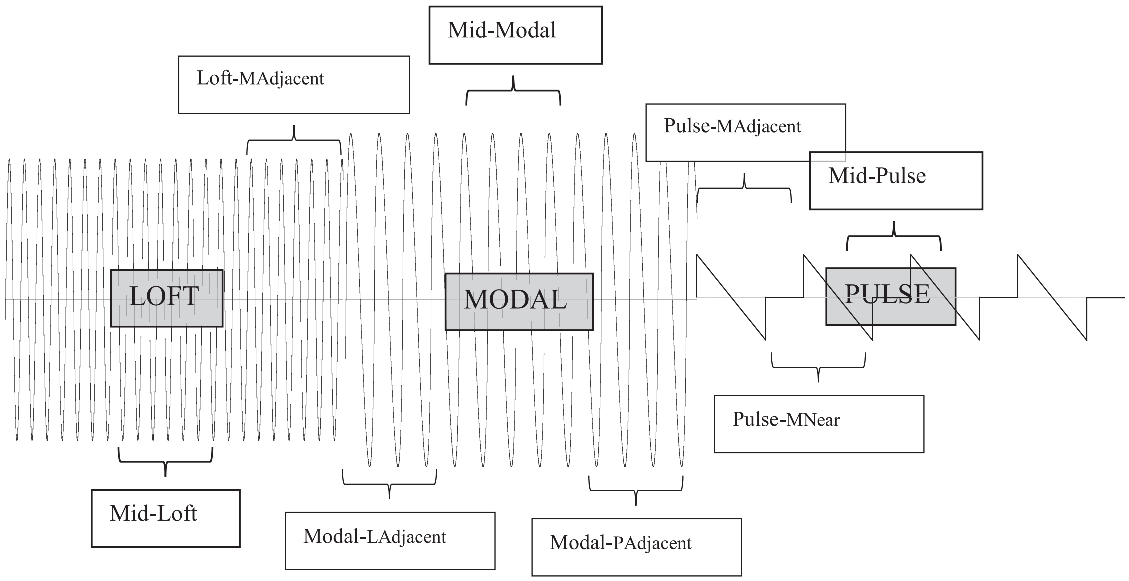

FIGURE 2. Schematic of f_o and H1-H2 measurement locations ("waveform" shapes in this schematic do not depict actual glottal waveforms and are merely intended to be evocative of register distinctions). Measurements were taken at "mid" locations for all three registers, and adjacent to loft-modal transitions. In pulse-modal transitions, transitional segments such as those illustrated in Figure 1 were typically observed: for this reason, f_o measures were still taken adjacent to transitions, but *H1-H2 measures were taken for pulse only "near" the transition*. This latter criterion meant that for pulse, a 0.5-millisecond windowed RMS contour had to reveal a continuous sequence of glottal pulses that were critically damped exhibiting temporal gaps for H1-H2 measurement.

TABLE 2.
Observations of Within-utterance Register Transitions

Infant	Age (mo)	Vocalizations With Loft	Transitions	Vocalizations With Pulse	Transitions	Modal Segments
AD	3	6	6 M-L, 3 L-M	10	7 M-P, 3 P-M	18
	6	3	1 M-L, 3 L-M	8	5 M-P, 3 P-M	12
	9	4	5 M-L, 3 L-M	6	3 M-P, 3 PM	16
EA	3	7	5 M-L, 5 L-M	5	2 M-P, 1 P-M	13
	5	8	6 M-L, 5 L-M	9	2 M-P, 7 P-M	19
	10	5	3 M-L, 2 L-M	5	1 M-P, 4 P-M	10
SM	4	12	8 M-L, 9 L-M	8	5 M-P, 3 P-M	25
	6	10	8 M-L, 5 L-M	7	1 M-P, 6 P-M	20
	11	10	9 M-L, 4 L-M	5	1 M-P, 4 P-M	16
	Totals:	65	51 M-L, 39 L-M	63	27 M-P, 34 P-M	149

Notes: The tallies of transitions for each infant and age represent the number of clear examples that could be readily found among the 40 minutes of recorded examined but not necessarily the maximum number to be found. Smaller numbers (five or fewer) reflect a paucity of examples (eg, infant EA at the youngest age for pulse), but in other cases, more than the listed number were readily identified (eg, infant SM for loft or infant AD at younger ages for pulse), but only the clearest exemplars were retained to provide representation across infants and ages. The total number of transitions = 151.
Abbreviations: L-M, loft to modal; M-L, modal to loft; M-P, modal to pulse; P-M, pulse to modal.

RESULTS

Occurrence of registers

Table 2 lists vocalizations selected from the initial pool of 2445 according to the criteria listed in Materials and Methods. This subset does not represent an exhaustive inventory of all transition occurrences but does include the utterances submitted for analysis and demonstrates that all contiguous and clearly distinct registers occurred within vocalization at least once at each age for each infant. Sample sound files with descriptions are provided as Supplementary Data.

Acoustic characteristics of registers and register transitions

Register differences

An ANOVA confirmed that the three registers were clearly distinct by f_o in hertz ($F(2, 292) = 238$; $P < 0.001$) and also in each pairwise comparison (Tukey honestly significant difference, $P < 0.001$). These effects were stronger yet in semitones ($F(2, 292) = 681$), where the distributions were clearly more well normalized as seen in Figure 3B below by the symmetrical whiskers, the distribution of outliers both above and below the median,

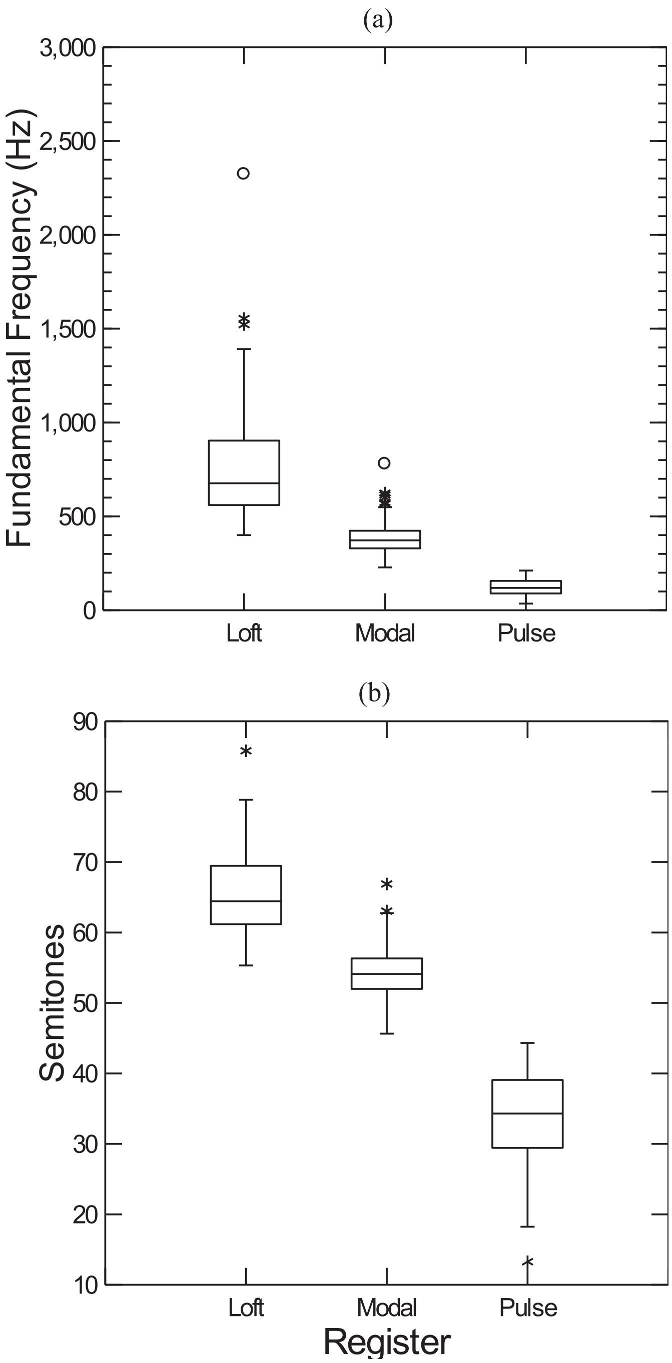

FIGURE 3. Box and whisker plots depicting f_o in (**A**) hertz and (**B**) semitones, measured midregister.

and the absence of extreme outliers above the median. See Table 3 for values and numbers of f_o observations. Note that modal and loft showed an overlap, whereas modal and pulse registers were nonoverlapping. It is notable that pulse in the infant data was, on average, at a frequency that would be in a modal range for adults—122 Hz—and was observed to be as high as 211 Hz even at regime midpoint.

ANOVA also indicated an overall difference between registers in these data by the H1-H2 measure ($F(2, 276) = 93.4$, $P < 0.001$), but pairwise comparisons by Tukey honestly significant difference were significant at $P < 0.001$ only for loft vs

TABLE 3.
f_o Values of Registers in Hertz at Midpoints of the Regime Segment

	Minimum	Mean	Maximum	n
$Loft_{mid}$	399	791	2322	90
$Modal_{mid}$	228	384	779	148
$Pulse_{mid}$	35	122	211	58

TABLE 4.
H1-H2 Values in Decibel of Registers at Midpoints of the Regime Segment

	Minimum	Mean	Maximum	n
$Loft_{mid}$	−12	9.3	33	89
$Modal_{mid}$	−22	−6.0	17.2	148
$Pulse_{mid}$	−17	−6.0	6.4	42

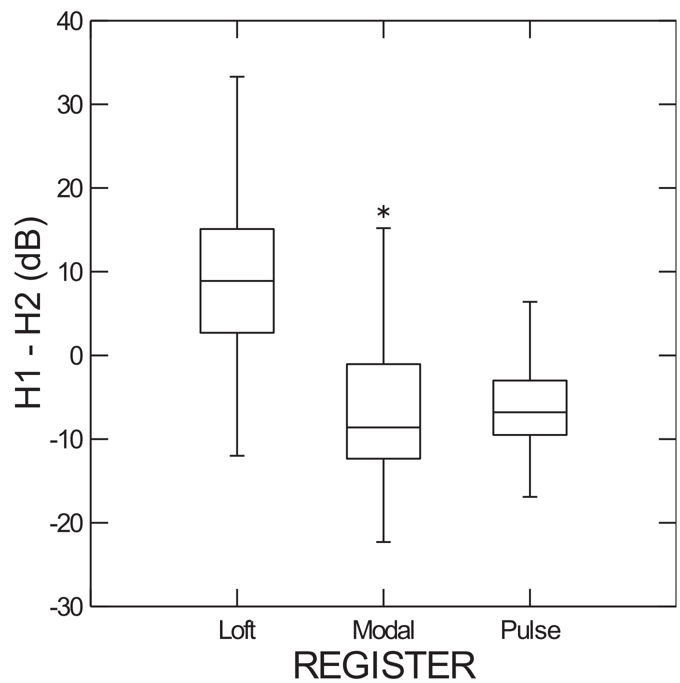

FIGURE 4. Bar charts depicting H1-H2 transitions by type and direction.

the other two registers. See Table 4 and Figure 4: both modal and pulse phonations showed a high second harmonic relative to the first in infant phonation, whereas loft phonation in the infants yielded the expected higher first harmonic relative to the second. There was still much variation within registers and overlap across registers, but loft was distinct in both measures.

Register transitions

We now turn to inspection of the measures taken adjacent to transitions. Recall that for these comparisons, as schematized in Figure 2, the f_o data were taken as close to the transition as possible, adjacent to it, in both types of register shifts. Loft transitional

TABLE 5.
f_o Values in Hertz of Nonmodal and Modal Registers Adjacent to Transition Points by Transition Types

Transition	Register	Minimum	Mean	Maximum	Mean Transition	Shifts
Modal→loft	Modal	**289**	*428*	**603**	313 Hz (9.1 ST)	428 ↗ 741
n = 51	Loft	440	741	1219		
Loft→modal	Loft	**396**	*687*	**1254**	−272 Hz (−8.4 ST)	687 ↘ 415
n = 39	Modal	290	415	647		
Modal→pulse	Modal	**136**	*267*	**377**	−66 Hz (−5.3 ST)	267 ↘ 201
n = 29	Pulse	70	201	293.7		
Pulse→modal	Pulse	**124**	*206*	**275**	96 Hz (6.5 ST)	206 ↗ 302
n = 32	Modal	193	302	560		

Notes: Bold entries indicate the last adjacent values in the direction of transitions. Italics indicate the average last adjacencies.
Abbreviation: ST, semitone.

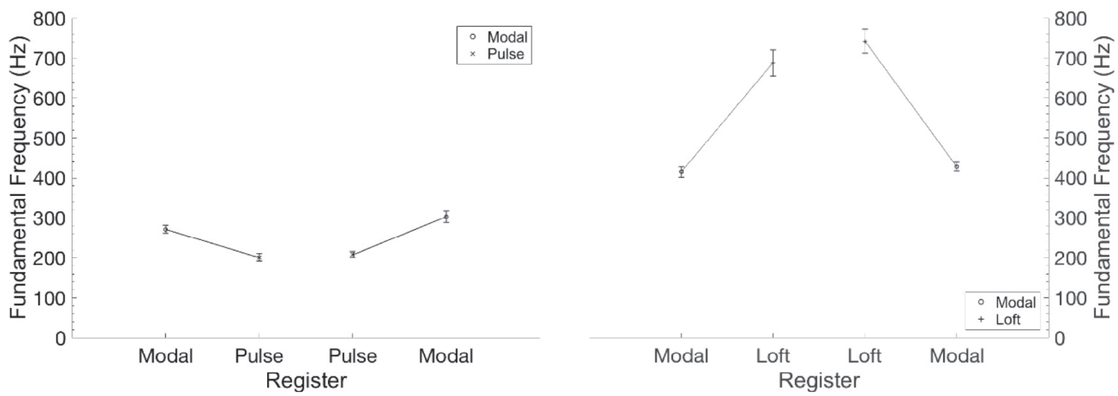

FIGURE 5. f_o transitions by type and direction. Error bars are standard errors.

H1-H2 measures were also taken adjacent to shifts, but pulse transitional H1-H2 measures were taken as near to the transitions as possible, so long as pulse-by-pulse temporal gaps had been clearly observed (again, typically no more than 10–20 milliseconds separated from clearly modal regimes). Table 5 lists f_o measure statistics by transition type for each register separately, and also includes the average changes for each transition type in both hertz and semitones, and regions of overlap between the upward and downward going transitions. Figure 5 depicts the transition differences in f_o.

Loft-modal f_o *transitions*

Paired t tests confirmed that both upward and downward shifts traversed significantly distinct end points: modal→loft: $t(50) = 13.9$, $P < 0.001$, and loft→modal: $t(38) = -11.2$, $P < 0.001$. These effects were even larger when assessed on the semitone scale: modal→loft: $t(50) = 22.0$, $P < 0.001$, and loft→modal: $t(38) = -15.9$, $P < 0.001$. The average last hertz value of modal before a transition to loft (428) was about the same as the average first hertz value of modal after a transition from loft (415), and a t test affirmed that these values were statistically indistinguishable ($t(88) = -0.77$, $P = 0.44$). The first hertz value of loft after a transition from modal (741) was also comparable with the last value of loft before a transition to modal (687), and these were also statistically indistinct ($t(88) = -1.21$, $P = 0.23$). The upward and downward shifts appeared comparable in Hz and their magnitudes were statistically indistinct ($t(88) = 1.2$, $P = 0.22$).

Pulse-modal f_o *transitions*

Paired t tests confirmed that both upward and downward shifts traversed significantly distinct end points: modal→pulse: $t(31)=-5.46$, $P < 0.001$, and pulse→modal: $t(28) = -5.94$, $P < 0.001$. These effects are comparable when assessed on the semitone scale: modal→pulse: $t(31) = -4.76$, $P < 0.001$, and pulse→modal: $t(28) = 6.41$, $P < 0.001$. The average last hertz value of modal before a transition to pulse (267) was smaller than the average first hertz value of modal after a transition from pulse (302), and this difference did reach statistical significance ($t(59) = -2.05$, $P < 0.05$), but this level of significance might be viewed with caution, given the number of tests applied to these data (and when one high outlier value was removed from the post–pulse modal data the P value was 0.08). The first hertz value of pulse after transition from modal (201) was comparable with the last hertz value of pulse before transition to modal (206), and these were not statistically distinct ($t(59) = -0.4$, $P = 0.68$). The upward shift of 96 was larger than the downward shift of −66, but, apparently because of variability, these magnitudes were not statistically distinct ($t = -1.51$, $P = 0.14$).

Table 6 lays out the H1-H2 measures in the same format as f_o measures in Table 5, and Figure 6 depicts transition line charts

TABLE 6.
H1-H2 Values in Decibel of Nonmodal and Modal Registers Close to Transition Points by Transition Types

Transition	Register	Minimum	Mean	Maximum	Mean Transition
Modal→loft	Modal	**–26.3**	*–7.4*	**9.9**	+14.5
n = 51	Loft	–16.4	7.1	37.8	
Loft→modal	Loft	**–17.9**	*8.0*	**31.5**	–15.2
n = 39	Modal	–21.5	–7.2	8.9	
Modal→pulse	Modal	**–19**	*–9.2*	**6.9**	+0.8
n = 29	Pulse	–20.3	–8.4	4.2	
Pulse→modal	Pulse	**–21.3**	*–8.0*	**6.2**	–4.3
n = 32	Modal	**–39.1**	*–12.3*	**9.8**	

Notes: Bold entries indicate the last adjacent values in the direction of transitions. Italics indicate the average "last" adjacencies.

for this measure. Recall that for midregister, depicted in Figure 4 and quantified in Table 4, large H1-H2 differences were observed in relative harmonic amplitudes for loft vs modal and loft vs pulse, but not for modal vs pulse.

Loft-modal H1-H2 transitions

Large shifts in average H1-H2 values were observed in both transition directions, reflecting an abruptly occurring change between a relatively strong second harmonic in modal to a relatively strong first harmonic in loft. Paired *t* tests revealed that these harmonic amplitude changes were statistically significant across the modal-loft transition ($t(50) = -7.61$, $P < 0.001$) and across the loft-modal transition ($t(38) = 7.13$, $P < 0.001$). Note in Table 4, however, that a very large range of relative harmonic amplitudes was observed in both modal and loft—this range is examined further in more detail. The shifts appear to be comparable in both directions, and their magnitudes were statistically indistinct ($t(88) = 0.02$, $P = 0.99$).

Pulse-modal H1-H2 transitions

Because H1-H2 differences were not found between pulse and modal midregister, it is not surprising that they were also not observed to be very different across the transitions. More surprising is the observation that there were somewhat less negative H1-H2 differences in pulse compared with nearby modal values, indicating comparatively somewhat less amplitude in the second harmonic, as in loft. After modal to pulse transitions, the second harmonic was higher very slightly relative to the first, but not with statistical significance according to matched pair *t* test ($t(31) = -0.66$, $P = 0.51$). Across pulse-modal transitions, a small overall increase in the second harmonic relative to the first was significantly distinct at the $P < 0.01$ level ($t(29) = 2.61$, $P < 0.05$). Again, this small effect must be viewed with caution, given the small numbers of observations and larger number of tests conducted on this dataset overall, but it remains of interest as tending oppositely to expectation based on adult models.

Variation within registers

As seen in Figure 4 for midregister variation and the transitional value ranges listed in Table 6, there was considerable variation in relative harmonic amplitudes in these data both within and across registers. Some of these variations may be related to fundamental frequency: The overall Pearson correlation taken between f_o and H1-H2 for all the available measures in this corpus was significantly positive ($r = 0.50$, $P < 0.001$), possibly reflecting a general covariation in these acoustic variables with pressure, amplitude, glottal resistance, and other parameters likely to cause H1-H2 changes. However, the scatter plot depicted in Figure 7 suggests that this overall correlation was just as much if not entirely due to variation across the register clusters in this space. Furthermore, if the registers are produced by different vibratory mechanisms,[1] patterns of covariation might also be expected to differ within register.

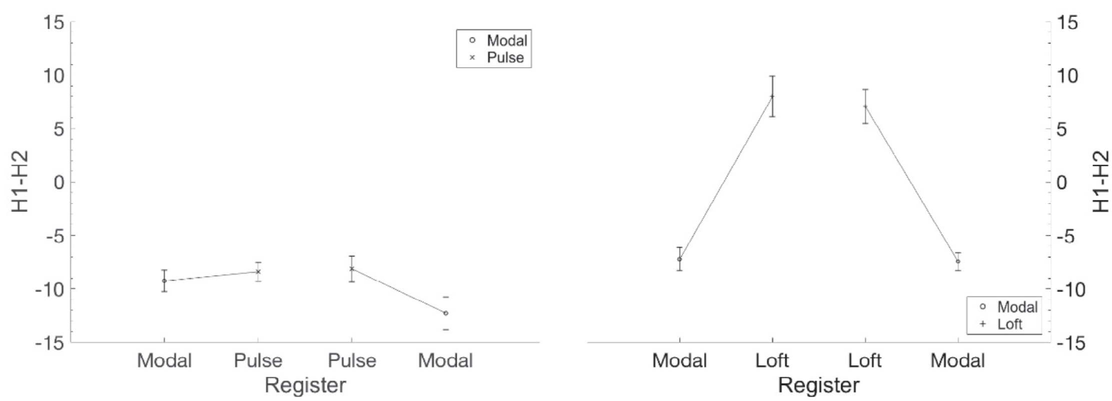

FIGURE 6. Box and whisker plots depicting H1-H2, measured midregister.

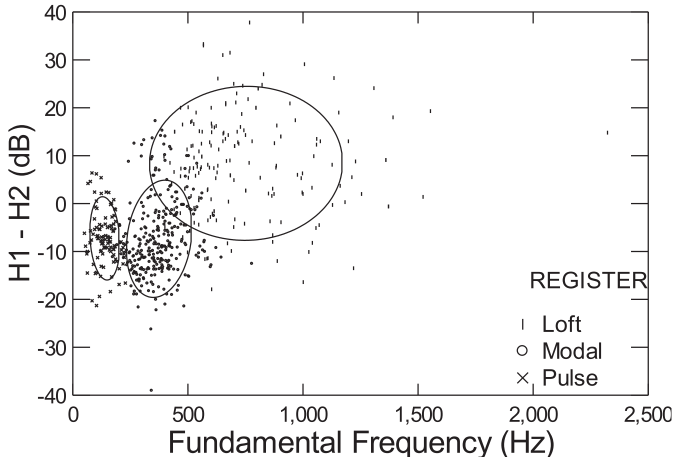

FIGURE 7. Scatterplot of all H1-H2 measures over f_o (inclusive of midregister, loft and modal adjacencies, pulse-adjacent modal, and modal-near pulse).

Examining solely within registers, a small but significant positive correlation was obtained for the modal values ($r = 0.18$, $P < 0.01$), whereas the others were nonsignificant at zero or negative values (loft $r = 0.01$, $P = 0.89$, and pulse $r = -0.15$, $P = 0.13$). Furthermore, when only midregister values were examined, all correlations became nonsignificant, and the negative correlation for pulse register became more so at $r = -0.23$. In conclusion, f_o and H1-H2 measures were not observed to covary except when all three registers were combined, and so this variation may be epiphenomenal to the distribution across clusters and not characteristic of the mechanisms except possibly within modal. For the purpose of using f_o and H1-H2 as acoustic classifiers distinguishing registers, we therefore proceed to analyze modal↔ nonmodal distinctions for the two nonmodal registers separately.

Acoustic classification of registers

Table 7 presents logistic regression modeling results using f_o and H1-H2 as predictors, and each of the two pairwise modal vs nonmodal register contrasts as outcomes. Note that pulse was fully predicted by f_o, with no further predictive value when H1-H2 was added. Loft was best predicted by f_o, but H1-H2 added a good bit of final variance explained, increasing prediction from 0.86 to 0.90. Overall, at 0.90 prediction of loft with only two measures, and 0.96 for pulse with only one, the measurement set contributes a great deal of explanation for the categories.

The f_o threshold value for modal vs pulse in this dataset was 208.5 Hz, conforming nicely with the general rule of thumb of 200 Hz that "regime" coders had used when identifying pulse.[6] Note again, however, that this would be a remarkably high f_o at which to hear individual pulses, yet pulse was the percept for these infants: a generally rasping "washboard," "zipper," or "frog" sound, occurred despite the high f_o. The result is surprising because adult pulse is reportedly perceived at the much lower f_o where individual pulses become detectable.

With both variables predicting modal vs loft, the combination threshold values were $f_o = 516$ Hz and H1-H2 = 1.6 dB. This f_o cutoff was slightly lower than the threshold for f_o alone, and this H1-H2 threshold was slightly more negative than that in the harmonics alone model. In combination, the measures increased prediction but at a negligibly small improvement in comparison with either measure alone. However, given the large gaps across loft↔ modal transitions and the large variation in both measures within registers, it is almost certainly of value to utilize both measures when discriminating these registers. Examining correctly classified instances, it could be seen that sometimes one measure "rescued" the other. For example, the segment with the lowest f_o that was still identified as loft—$f_o = 396$ Hz—had an H1-H2 measure of 9.8 dB, well above the cutoffs. The highest f_o still identified as modal—$f_o = 779$ Hz—had an H1-H2 measure of –12.6 dB. These observations validate perceptual impressions that perceived registers in these infants were not simply a matter of pitch.

DISCUSSION

Each of the three infants in the present study demonstrated within-vocalization register transitions at least once in relatively brief recorded samples at each of the ages examined during their first year of vocal development. The loft-modal transitions were particularly salient, corresponding in the acoustic data to large jumps in f_o and large changes in H1-H2, typically positive for perceived loft and negative for perceived modal. Pulse-modal transitions were equally evident, however, and were also signaled by clear acoustic markers. Although these transitions were distinguished better by f_o than by H1-H2, pulse surprisingly trended toward more positive H1-H2 values: given the theoretical bases for this measure, such values would presumably result from a *less* pressed configuration of the glottis in comparison with the immediately adjacent modal phonation.

The harmonic amplitude measures alone did not clearly support for the infants the distinct mechanism for pulse phonation posited

TABLE 7.
Logistic Regression Results for Discriminating Registers Across Transitions

	H1-H2			f_o			H1-H2 and f_o	
	Prediction	$Area_{ROC}$	Cutoff	Prediction	$Area_{ROC}$	Cutoff	Prediction	$Area_{ROC}$
Loft-modal	0.74	0.88	2.8	0.86	0.97	523	0.90	0.98
Pulse-modal	0.62*	0.48*	NA	0.96	0.996	209		

* Only f_o is significant.
Abbreviation: NA, not applicable.

in the adult literature. Moreover, perceived register transitions occurred around 200 Hz, casting doubt upon the applicability for infants of the psychoacoustic argument from adult literature, suggesting that pulse is identifiable because individual pulses are only audibly distinct when f_o is lower than 70 Hz. By the same token, the simple fact that infant vocal folds will naturally vibrate at higher frequencies does not explain the perceived register transition at the higher f_o. Furthermore, the lack of clearly distinct H1-H2 values in perceived pulse register seems to indicate that the infants were not producing pulse by "constricting" the glottis as is generally thought to occur in adults.[3]

Most broadly, in comparison with adult phonation, these results indicate that the loft register operates very similarly in infants, although it is signaled by somewhat higher f_o, whereas the pulse register in infancy may operate quite differently from that in adults. The findings should be interpreted in the context of known differences between adult and infant vocal fold composition, specifically regarding the relatively undifferentiated and compliant lamina propria in infants. In formulating contemporary myoelastic-aerodynamic theory, van den Berg himself[39] identified the ligament as most likely needed for loft phonation. The stiffness of adult ligamental layers, in conjunction with macula flavae, contributes measurable differences to natural vocal fold vibration frequencies.[40] As it is established that these two aspects of vocal fold histology—ligament and macula flavae—are underdeveloped during possibly the entire first year of life,[41] it is most interesting to observe that infants produce a distinctive loft register quite readily, apparently contradicting the general assumption that a ligamental layer is critically involved.

A propensity for infant vocal folds to produce sound signaling a pulse register, even at relatively high f_o values but without distinctive H1-H2 values, may be understood in terms of infant tissue characteristics. "Deligamented" lamina propria are likely to be highly compliant and also more readily detachable from the body. Such loose and detachable vocal fold cover tissues could provide the lax cover and marginalized body tissues that fry phonation is thought to involve, potentially explaining the lack of distinctive H1-H2 characteristics even for glottal pulses that exhibit critical damping as evidenced in temporal gaps between pulses. In other words, a reason that this perceived register did not correspond to an anticipated "pressed" marker, signaled by a high second harmonic (relative to nearby modal register), might be attributable to the very compliance of these cover tissues. Furthermore, this loose detachable aspect might also explain the occurrence of falsetto style vibration without the involvement of a ligament. The natural protophone categories that result (often called vowel-like sounds, squeals, and growls) are routinely recognized by caregivers and form anchors for vocal interaction while infants are still prelinguistic.[10]

Turning to the observed threshold values obtained in the logistic regression results, these values will be useful in classification work underway toward understanding the acoustic dimensions of infant protophones.[10,30] Although previously our regime codes included only an agnostic high modal category, the f_o and harmonic amplitude measures reported here suggest that a loft register is indeed present in infants. The register distinction may provide the basis for distinguishing two types of "squeals" in infant vocalizations, one as corresponding to loft, and the other, high-pitched modal. Register distinctions made on an acoustic basis can now be added to other acoustic dimensions, such as f_o variability and other nonmodal regimes, and the combination should help in modeling of human coding of infant protophone types.[42,43] Perceptual coding of phonation by young children with autism suggests that the register distinction may be important for understanding audible signs characterizing this population.[44]

Summary

The results presented here demonstrated that within-vocalization transitions between the modal register and the two primary nonmodal registers of loft and pulse occur at each of three different ages during the first year of life in each of the three infants sampled. These transitions resemble those observed in adults, even though the cover tissues of infant vocal folds are quite different from those found in adults. In addition to describing phonatory phenomena not previously documented in infant phonation, the results present specific challenges to voice science by demonstrating that (1) loft phonation does not require a vocal ligament, and (2) pulse phonation can occur audibly at frequencies above 200 Hz.

Limitations and future research

A major limitation of the present study is the small number of speakers involved, only three female infants. As with so many other aspects of speech analysis, all the measures of the present study required prior determination of f_o, and early infant phonation is notorious for confounding this determination (eg, due to subharmonics, biphonation, and chaos). With ongoing developments in automated recognition of harmonic structures in spectrograms, however, it can be hoped that a more rapid and ultimately automatic detection of registers in infant phonation might be possible.

With more rapid data collection on a larger number of subjects, applications may extend beyond certification of protophone dimensions and categories in a variety of directions. Age effects, although not found in this limited dataset, ought to be observed beyond the first year of life, especially in a lowering of the pulse f_o threshold. It is possible that infants' use of registers could be imitative or could vary by situation as an indicator of emotion or interactional intent. Quantified discrimination of prelinguistic phonation and distress-call ("crying") phonation has yet to be performed, and knowledge of register-specific modes of infant phonation will provide an essential backdrop for this effort. Early detection of voice-related signs and symptoms of speech disorders is another longer term direction, but meanwhile, clearly much basic voice science is yet to be done, tracing the ontogeny of register range measures with the schedule of ligament and macula flavae developments.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2017.12.013.

REFERENCES

1. Titze IR. *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice Hall; 1994.
2. Gordon M, Ladefoged P. Phonation types: a cross-linguistic overview. *J Phon*. 2001;29:383–406.
3. Keating PA, Garellek M, Kreiman J. Acoustic properties of different kinds of creaky voice, Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, 2015.
4. Kent RD, Murray AD. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *J Acoust Soc Am*. 1982;72:353–365.
5. Robb MP, Saxman JH. Acoustic observations in young children's non-cry vocalizations. *J Acoust Soc Am*. 1988;83:1876–1882.
6. Buder EH, Chorna L, Oller DK, et al. Vibratory regime classification of infant phonation. *J Voice*. 2008;22:553–564.
7. Mende W, Herzel H, Wermke K. Bifurcations and chaos in newborn infant cries. *Phys Lett A*. 1990;145:418–424.
8. Robb MP. Bifurcations and chaos in the cries of full-term and preterm infants. *Folia Phoniatr Logop*. 2003;55:233–240.
9. Švec JG, Schutte HK, Miller DG. On pitch jumps between chest and falsetto registers in voice: data from living and excised human larynges. *J Acoust Soc Am*. 1999;106:1523–1531.
10. Oller DK. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
11. Hollien H. On vocal registers. *J Phon*. 1972;2:124–143.
12. Hollien H, Girard GT, Coleman RF. Vocal fold vibratory patterns of pulse register phonation. *Folia Phoniatr (Basel)*. 1977;29:200–205.
13. Vorperian HK, Wang S, Chung MK, et al. Anatomic development of the oral and pharyngeal portions of the vocal tract: an imaging study a. *J Acoust Soc Am*. 2009;125:1666–1678.
14. Sato K, Hirano M, Nakashima T. Fine structure of the human newborn and infant vocal fold mucosae. *Ann Otol Rhinol Laryngol*. 2001;110:417–424.
15. Boseley ME, Hartnick CJ. Development of the human true vocal fold: depth of cell layers and quantifying cell types within the lamina propria. *Ann Otol Rhinol Laryngol*. 2006;115:784–788.
16. Hartnick CJ, Rehbar R, Prasad V. Development and maturation of the pediatric human vocal fold lamina propria. *Laryngoscope*. 2005;115:4–15.
17. Nita LM, Battlehner CN, Ferreira MA, et al. The presence of a vocal ligament in fetuses: a histochemical and ultrastructural study. *J Anat*. 2009;215:692–697.
18. Rosenberg TL, Schweinfurth JM. Cell density of the lamina propria of neonatal vocal folds. *Ann Otol Rhinol Laryngol*. 2009;118:87–90.
19. Fuamenya NA, Robb MP, Wermke K. Noisy but effective: crying across the first 3 months of life. *J Voice*. 2015;29:281–286.
20. Neiman M, Robb M, Lerman J, et al. Acoustic examination of naturalistic modal and falsetto voice registers. *Logop Phoniatr Vocol*. 1997;22:135–138.
21. Salomão GL, Sundberg J. What do male singers mean by modal and falsetto register? An investigation of the glottal voice source. *Logop Phoniatr Vocol*. 2009;34:73–83.
22. Hanson HM. Glottal characteristics of female speakers: acoustic correlates. *J Acoust Soc Am*. 1997;101:466–481.
23. Kreiman J, Shue YL, Chen G, et al. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *J Acoust Soc Am*. 2012;132:2625–2632.
24. McGlone RE. Air flow during vocal fry phonation. *J Speech Lang Hear Res*. 1967;10:299–304.
25. McGlone RE, Shipp T. Some physiologic correlates of vocal-fry phonation. *J Speech Lang Hear Res*. 1971;14:769–775.
26. Blomgren M, Chen Y, Ng ML, et al. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J Acoust Soc Am*. 1998;103:2649–2658.
27. Avelino H. Acoustic and electroglottographic analyses of nonpathological, nonmodal phonation. *J Voice*. 2010;24:270–280.
28. Roubeau B, Henrich N, Castellengo M. Laryngeal vibratory mechanisms: the notion of vocal register revisited. *J Voice*. 2009;23:425–438.
29. Dilley L, Shattuck-Hufnagel S, Ostendorf M. Glottalization of word-initial vowels as a function of prosodic structure. *J Phon*. 1996;24:423–444.
30. Buder EH, Warlaumont AS, Oller DK. An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. In: Peter B, MacLeod AAN, eds. *Comprehensive Perspectives on Speech Sound Development and Disorders: Pathways from Linguistic Theory to Clinical Practice*. Hauppauge, NY: Nova Science Publishers Inc.; 2013:103–134.
31. Koopmans-van Beinum FJ, van der Stelt JM. Early stages in the development of speech movements. In: Zetterstrom R, ed. *Precursors of Early Speech*. New York: Stockton Press; 1986:37–50.
32. Nathani S, Ertmer D, Stark RE. Assessing vocal development in infants and toddlers. *Clin Linguist Phon*. 2006;20:351–369.
33. Oller DK. The emergence of the sounds of speech in infancy. In: Yeni-Komshian GH, Kavanagh JF, Ferguson CA, eds. *Child Phonology*, Vol. 1. Production). New York: Academic Press; 1980:93–112.
34. Buder EH, Stoel-Gammon C. American and Swedish children's acquisition of vowel duration: effects of vowel identity and final stop voicing. *J Acoust Soc Am*. 2002;111:1854–1864.
35. Milenkovic P. TF32, University of Wisconsin-Madison, Madison, WI, 2001.
36. Delgado RE, Milenkovic P. AACT—Action analysis coding and training software, Intelligent Hearing Systems Corp., Miami, 2017.
37. Lynch MP, Oller DK, Steffens ML, et al. Phrasing in prelinguistic vocalizations. *Dev Psychobiol*. 1995;1:3–25.
38. Robb MP, Cacace AT. Estimation of formant frequencies in infant cry. *Int J Pediatr Otorhinolaryngol*. 1995;33:57–67.
39. van den Berg J. Myoelastic-aerodynamic theory of voice production. In: Kent RD, Atal BS, Miller JL, eds. *Papers in Speech Communication: Speech Production*. Woodbury, NY: Acoustical Society of America; 1991:121–138.
40. Titze IR, Hunter EJ. Normal vibration frequencies of the vocal ligament. *J Acoust Soc Am*. 2004;115:2264–2269.
41. Tateya T, Tateya I. Vocal fold development. In: *Regenerative Medicine in Otolaryngology*. New York: Springer; 2015:161–169.
42. Buder EH. Topography of human vocal development, Symposium on Learning about the Vocal World: Deciphering the Statistics of Communication, Atlanta, GA, 2015.
43. Buder EH, Oller DK. Acoustic structure of common early protophones, Annual Meeting of the American Speech-Language-Hearing Association, Denver, CO, 2015.
44. Wallace MC, Cleary JE, Buder EH, et al., An acoustical analysis of vocalizations of children with autism, Annual Meeting of the American Speech-Language-Hearing Association, Chicago, 2008.