

# A New View of Language Development: The Acquisition of Lexical Tone

Leher Singh and Charlene S. L. Fu  
*National University of Singapore*

Research in first language development draws disproportionately from nontone languages. Such research is often presumed to reveal developmental universals in spite of the fact that most languages are tone languages. Recent research in the acquisition of tone languages points to a distinct course of development as compared to nontone languages. Our purpose is to provide an integrated review of research on lexical tone acquisition. First, the linguistic properties and origins of tone languages are described. Following this, research on the acquisition of tones in perception and production is reviewed and integrated. Possible reasons for the uniqueness of tone in language acquisition are discussed. Finally, theoretical advances promised by further research on tone acquisition and specific research directions are proposed.

Several decades of experimental research in language acquisition have led to momentous discoveries about the young language learner. From these findings, we can chart normative pathways to language proficiency from infancy through the formative years and, as a consequence, advance basic theories of language development. To date, research in the field of developmental psycholinguistics has focused primarily on children's acquisition of Romance and Germanic languages (e.g., English, Spanish, and French), thus drawing from a limited range of language systems that use vowels and consonants to contrast word meanings (nontone languages). However, most human languages are tone languages and use lexical tone in addition to vowels and consonants to draw meaningful distinctions between words by systematically manipulating vocal pitch (Yip, 2002). For example, in Mandarin Chinese, the mostly widely spoken tone language, the word "ma" assumes different meanings based on the tone in which it is produced. "Ma" means "mother" when spoken in Tone 55, "hemp" when spoken in Tone 35, "horse" when spoken in Tone 214, and "to scold" when spoken in Tone 51. The use of tone is widespread among languages of the world with an estimated 60%–70% of languages presumed to be tone languages (Yip, 2002). As

such, the majority of research on normative language acquisition draws disproportionately from a minority of languages. A natural consequence is that foundational theories of language development may not necessarily generalize to the population.

Our objective is to provide a review and integration of literature that investigates the acquisition of lexical tone over the first 3 years of life. We begin by providing an introduction to lexical tone, defining properties of lexical tones and describing cues relevant to the psycholinguistic processing of tone. Following this, we review and integrate literature on the perception of tones within three central domains of first language acquisition: (a) the formation of phonetic categories for tone, (b) the maturation and stabilization of tone categories, and (c) the integration of lexical tone into a child's developing lexicon. We then progress to a review of the development of tone production abilities in early childhood, focusing on constraints on early tone production and on the pathway to mature tone productions within the tone language learning child. Further to this, we discuss possible reasons why tone may develop in ways that are unique in comparison to vowels and consonants. We then progress to situate our review of empirical research within broader theories and models of normative language development. Finally, we propose future empirical directions for research in tone language development that could potentially inform and advance the field of first language acquisition.

---

We are grateful to the Ministry of Education Tier 1 Academic Research Fund grant (FY2013-FRC2-009) and a start-up grant from the National University of Singapore to Leher Singh to support the research reviewed in this manuscript. We thank Thilanga D. Wewalaarachchi and Liang Hui Wong for valuable comments on this manuscript.

Correspondence concerning this article should be addressed to Leher Singh, Department of Psychology, National University of Singapore, 9 Arts Link, Singapore 117570. Electronic mail may be sent to psylys@nus.edu.sg.

© 2016 The Authors

Child Development © 2016 Society for Research in Child Development, Inc.  
All rights reserved. 0009-3920/2016/xxxx-xxxx

DOI: 10.1111/cdev.12512

### An Introduction to Lexical Tone

Lexical tone refers to syllable-level shifts in the fundamental frequency contour that draw distinctions between the meanings of words. Lexical tone is primarily identified by fundamental frequency (or pitch), drawing on specific cues such as mean pitch and pitch contour (Chandrasekaran, Sampath, & Wong, 2010; Gandour, 1983, 1984; Whalen & Xu, 1992). However, lexical tone is not determined by pitch movements alone. Rather, there are other secondary cues to tone identity, most commonly reported to be duration, amplitude, vocal range, and the register of a speaker (Blicher, Diehl, & Cohen, 1990; Edmondson & Esling, 2006; Leather, 1983; Liu & Samuel, 2004; Whalen & Xu, 1992; Wong & Diehl, 2003). Secondary cues support native listeners' identification of tones when pitch cues are removed or obscured (Liu & Samuel, 2004; Whalen & Xu, 1992). Therefore, although tone is most commonly defined by pitch changes in psycholinguistic research, in actual fact, a broader range of acoustic cues work in concert to drive tone identification. However, the research reviewed herein focuses predominantly on the perception and production of tones as defined by pitch cues.

The expression of lexical tone is somewhat dependent on other phonetic segments, such as vowels and consonants. Tone is typically instantiated on the vowel, but the realization of tone interacts with surrounding consonants (Hombert, Ohala, & Ewan, 1979; Lee, 2000). However, lexical tone is crucially different to vowels and consonants in two primary ways. First, lexical tone entails syllable-level changes, while vowels and consonants entail segmental changes. As a result, tone is sometimes classified as a suprasegmental cue, whereas vowels and consonants are incontrovertibly classified as segmental cues. Second, there are differences in the acoustic composition of vowels, consonants, and tones, most notably in the distribution of acoustic energy, which is clearly observable in spectral analysis. During speech production, the vocal tract permits the concentrated passage of acoustic energy at particular frequencies (while limiting the passage of sound at other frequencies). Formants refer to the high concentrations of acoustic energy that are permitted to pass through the vocal tract. Vowels and consonants are related to formant properties (specifically, frequency values and transitions) and are not typically defined in terms of fundamental frequency ( $F_0$ ; Ladefoged & Disner, 2012). In contrast, lexical tone is not conventionally defined by formant struc-

ture, but primarily identified by  $F_0$  variation (Gandour, 1978). As such, the primary acoustic-phonetic cues associated with tone production differ from that associated with the production of consonants and vowels.

The origin of tone in human languages (known as "tonogenesis") has been extensively documented in South East Asian languages, such as Mandarin Chinese and Vietnamese (see Haudricourt, 1954; Matisoff, 1973; Thurgood, 2002). Tone is thought to have evolved due to the monosyllabic nature of particular languages, which resulted in syllables that contained a wide variety of consonant sounds that were used to contrast meaning. Gradually, differences between certain consonants were neutralized. Tones then reportedly emerged as a way to maintain distinctiveness between similar-sounding words (Matisoff, 1973). In particular, it is proposed that in early phases of the evolution of tones, pitch height of tones originated from the attrition of syllable-initial consonant contrasts and pitch contours from the attrition of syllable-final consonants (Haudricourt, 1954).

Relevant to their linguistic origin, tone languages are not randomly distributed throughout the world. Rather, they aggregate within certain geographical regions where they serve as the linguistic norm. This patterning has led to conclusions that contact with neighboring languages has contributed to the emergence of particular tone languages (e.g., tone in Vietnamese is thought to have emerged from intensive contact with Chinese languages; Matisoff, 2001). Regions with a high concentration of tone languages include sub-Saharan Africa and continental and mainland South East Asia (Indochinese Peninsular), Central America, the Caribbean, and the Amazon basin. Tone languages are rare in Europe, North Africa, and America. The geographical aggregation of tones may be due to the fact that tone languages may have spread via contact with other tone languages in the region. A second proposed reason (not mutually exclusive from the first) is that there may be a significant concordance between the population frequency of specific genetic expressions and the areas in which tone languages are commonly spoken, suggesting that some populations may also have a genetic predisposition to tone. Two particular genes, ASPM-D and MCPH-D, have been thought to affect neural organization in subtle ways that favor perception and acquisition of linguistic tone (Dediu & Ladd, 2007).

Tone languages are not all alike and are broadly defined in two ways: contour tone systems and register tone systems. Contour tones involve complex

changes in the pitch contour to impart changes in word meaning (e.g., Mandarin Chinese). Register tone languages draw from a set of relatively level (flat) tones to distinguish meaning via tone, such as many Bantu languages. An example of the inventory of a contour tone language and a register tone language can be seen in Figure 1a and b. It should be noted that some languages combine contour and register tones, such as the Kru languages spoken in the Southwest Ivory Coast and Southern Liberia. In addition to the form assumed by tones, the linguistic functions of tones also vary by geographical location. In East Asia, tone is typically lexical, whereas in many African languages, tone often serves a more significant grammatical function rather than lexical function.

A family of languages often likened to tone languages are pitch-accent languages. Although speakers of tone and pitch-accent languages process vocal pitch in similar ways (Burnham et al., 2014), the extent to which tone is used in tone languages differs from that of pitch-accent languages. In tone languages, each syllable within a word has a distinct tone assignment that determines its meaning. However, in pitch-accent languages, such as Swedish and Japanese, variation in pitch is imposed on a subset of syllables, often within a word, and not all words receive a pitch accent. Pitch-accent languages are therefore sometimes described as “reduced” tone languages, although their linguistic status is not uncontroversial (see Hyman, 2009).

For the most part, experimental research on the acquisition of lexical tone has focused on East Asian tone languages and more specifically on Mandarin

Chinese, Cantonese, and Thai. The tone inventories of commonly studied tone languages (Mandarin Chinese, Cantonese, and Thai) are listed in Table 1. Tones are conventionally denoted using a numerical system devised by Chao (1930). Within the Chao numerals system, each tone is denoted by a numeric sequence that describes the onset and offset pitch of a syllable using a scale of 1–5. There is a possible additional medial number in the case of a complex pitch contour such as those with an inflection point. In this sequence, the lowest pitch is represented by 1 and the highest by 5. For instance,

Table 1  
*The Tonal Inventories of Mandarin Chinese, Cantonese, and Thai*

Language	Chao tone numerals	Tone contour
Mandarin (Chao, 1951)	55	High level (level)
	35	Midrising (rising)
	214	Low dipping (dipping)
	51	High falling (falling)
Cantonese (Francis et al., 2008)	55	High level
	25	High rising
	33	Midlevel
	21	Low falling
	23	Low rising
Thai (Diller, 2004)	22	Low level
	33	Mid
	11	Low
	52	Falling
	45/55/454	High
	24	Rising

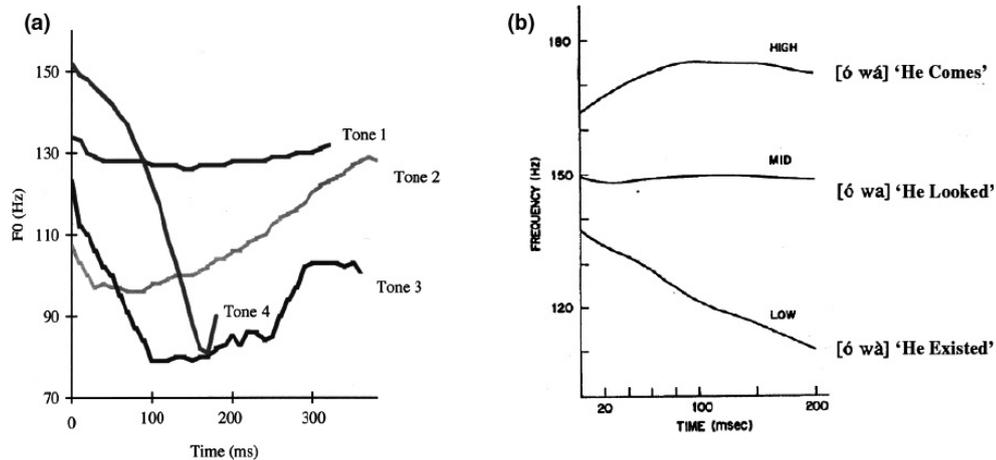


Figure 1. (a) The Mandarin tone inventory (contoured tones; from Wang et al. 1999). (b) The Yoruba tone inventory (register tones; from Hombert et al., 1979).

Tone 55 describes a high-level tone (high-onset-high-offset pitch) and Tone 35 describes a rising tone (medium-onset-high-offset pitch). It should also be noted that these numerals are relative to a speaker's vocal range, and do not reflect absolute pitch values. In this review, all tones will be represented by their Chao tone numerals for the purposes of simplicity.

### The Formation of Lexical Tonal Categories

One of the first steps to mastering a tone language is to determine that tones serve a lexical function within the learner's native language and as such, they must be accurately discriminated by the tone language learner. In all languages, the ability to perceive native phonological contrasts is a crucial component of first language acquisition and is thought to foreshadow the development of formal language skills (Kuhl, 2004; Werker, 2003).

A vast body of research has demonstrated that infants begin life with the ability to distinguish a wide variety of vowels and consonants whether or not these sounds are contrasted within the infant's native language (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Polka & Werker, 1994; Werker & Tees, 1984; but see Best & Tyler, 2007). This has led to the characterization of newborn infants as "universal listeners" (Kuhl, 2011). Infants then undergo major developmental changes in phonetic perception over the 1st year of life, retaining the ability to distinguish vowels and consonants that feature in the native language, and demonstrating a decline in their ability to distinguish many foreign vowels and consonants (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Kuhl et al., 2006; Polka & Werker, 1994; Werker & Tees, 1984). This shift in the perception of vowels and consonants toward the properties of the native language is known as phonological attunement and is recognized as a critical phase in first language development (Curtin, Byers-Heinlein, & Werker, 2011; Kuhl, 2004; Werker, 2003).

More recently, there have been several initiatives to investigate phonological attunement of lexical tone. Studies have investigated early sensitivity to lexical tone in two domains: the auditory discrimination of lexical tones in infancy and the association of lexical tones with meaning. Collectively, these studies have revealed important differences in the emergence of a phoneme inventory for tone as compared to vowels and consonants.

### *The Development Lexical Tone Categories: Evidence From Infant Phoneme Discrimination*

#### *Discrimination of Lexical Tone in Tone Language Learners*

Experimental approaches to tone discrimination originated through a study by Harrison (2000) who compared tone-language-exposed (Yorùbá) infants with nontone-language-exposed (English) infants on responses to tone contrasts. Using a visual reinforcement paradigm where infants were exposed to an auditory stimulus and then conditioned to make a headturn in response to a change in stimulus, 6- to 8-month old Yorùbá-learning and English-learning infants were tested on their sensitivity to tone contrasts in Yorùbá. The results of this study demonstrated that while Yorùbá-exposed infants were able to distinguish some Yorùbá tones, English-exposed infants were unable to discriminate any of the tone contrasts. This study provided the first experimental evidence that early tone discrimination is modified by tone language experience. Harrison's findings complemented a long-standing tradition of research on vowel and consonant perception in revealing effects of native language experience on phonetic perception in infancy.

Since Harrison's (2000) initial investigation, several other studies pursued the question of when infants attune to tone, revealing a more complex developmental timetable. In a study by Tsao (2008), Mandarin-learning infants from 10 to 12 months of age were tested on their abilities to distinguish Mandarin tone contrasts. The tone pairs to be discriminated varied in acoustic salience. For instance, Mandarin Tones 55 and 214 are a highly distinct tone pair, Tones 35 and 51 are moderately distinct from each other, and Tones 35 and 214 are minimally distinct from one another. While Mandarin Chinese infants were able to discriminate Mandarin tones, they showed differences in performance based on the tone pair tested: Tones 55 and 214 were discriminated with higher accuracy than Tones 35 and 214 and Tones 35 and 51 (which yielded comparable performance). Moreover, the highly discriminable pair, Tones 55 and 214, were better discriminated when familiarized with Tone 55 and tested on Tone 214 than vice versa. Effects of stimulus-specific discriminability reported by Tsao mirror those reported in studies of the attunement of vowels and consonants (Anderson, Morgan, & White, 2003; Narayan, Werker, & Beddor, 2010). Effects of direction of testing are consistent with prior studies reporting directional asymmetries

in vowel discrimination (Kuhl et al., 2006; Polka & Bohn, 2003).

Tsao's (2008) study investigated tone discrimination in infants between 10 and 12 months, typically defined as the terminal phase of native phonological attunement for vowels and consonants (Werker & Tees, 1984). Developmental changes in tone perception at earlier phases of infancy were systematically explored by Yeung, Chen, and Werker (2013). Using a habituation-based paradigm, Yeung et al. investigated age-based changes in tone perception by comparing tone- (Cantonese or Mandarin) and nontone-language-exposed (English) infants at 4 and 9 months on discrimination of Cantonese tones. The primary dependent measure was fixation time to a visual display during repeated presentations of particular tone strings (e.g., Tone 25 and Tone 33) in comparison to fixation time to the same visual display when the same tones alternated (e.g., tokens of Tone 25 interspersed with tokens of Tone 33). The authors discovered that English-exposed infants discriminated Cantonese tone contrasts at 4 months but demonstrated attenuation in this ability at 9 months. By contrast, Cantonese- and Mandarin-exposed infants could discriminate Cantonese tones under particular conditions (i.e., when trained on Cantonese Tone 25, discrimination abilities were relatively strong, but when trained on Tone 33, their ability to discriminate tones were much weaker). Moreover, Cantonese- and Mandarin-exposed infants showed different responses to one tone (Cantonese Tone 25), which the authors attributed to differences in how this tone relates to the native tone inventories of Cantonese and Mandarin. Yeung et al., therefore, demonstrated a decline in sensitivity to lexical tones for nontone learning infants between 4 and 9 months or age, but no decline for tone language learning infants over this period.

In contrast to previous results obtained with vowels and consonants, Yeung et al. (2013) demonstrated very specific effects of language experience on infant phonetic perception as early as 4 months of age: English-, Cantonese-, and Mandarin-exposed infants each demonstrated different discrimination abilities that accorded with the properties of their native language at this stage. This study suggests that phonetic categories for tone evolve earlier than those for vowels and consonants. Furthermore, this study called into question existing notions of infants under the age of 6 months as being universal listeners, suggesting that this description may not apply to the large proportion of infants learning a tone language.

Although tone categories seem to emerge early in development, this leaves open the question of whether early tone categories are phonologically well specified and by extension, whether they are truly "native." Ideally, emergent tone representations should be very precisely tuned to the native tone inventory, as appears to be the case for nascent vowel and consonant categories (see Best, 1995). In a study that bears on the phonological specificity of early tone categories, Mattock and Burnham (2006) sampled English- and Chinese- (Mandarin or Cantonese) learning infants at 6 and 9 months. Infants were tested on their ability to discriminate Thai tone contrasts. Mattock and Burnham reported that English-exposed infants demonstrated attenuation in tone sensitivity between 6 and 9 months, while Chinese-learning infants exhibited a sustained sensitivity to Thai tones over this period. Both groups demonstrated similar sensitivities to musical tone, suggesting that tone attunement may generalize across tone inventories (i.e., from Cantonese or Mandarin to Thai), but not from voiced sourced to instrumental pitch contrasts.

Mattock and Burnham's (2006) report of attunement to Thai tones in Chinese-learning infants is difficult to reconcile with those obtained by Yeung et al. (2013), whose findings suggest a high degree of phonological specificity of tone attunement as early as 4 months. There are several methodological differences that may account for the discrepancy in findings reported by Mattock and Burnham and Yeung et al. First, it is highly possible that performance is governed by the precise relation between the tones used as experimental stimuli and the learners' native tone inventory (as previously documented for vowels and consonants by Best, 1995, and for tones by Reid et al., 2014). For example, in contrast to stimuli used by Yeung et al., the Thai stimuli employed by Mattock and Burnham incorporated lengthier tone segments comprising acoustically distinct tone pairs. The inclusion of a highly salient contrast may have facilitated Thai tone discrimination in Chinese-learning infants. Second, the dependent measures extracted from each study (percent correct from a conditioned headturn paradigm employed by Mattock & Burnham, 2006, vs. visual fixation to a neutral stimulus upon presentation of tones in a stimulus alternating paradigm by Yeung et al., 2013) are quite different in the demands they place on the infant, in their relative sensitivity, and, therefore, in their interpretive scope. In particular, conditioned headturn paradigms require infants to execute a learned behavior

in response to a distinction taught in the laboratory (and presumably also accrued outside of the laboratory), whereas visual habituation tracks infants' response to stimulus changes based on experience primarily acquired outside of the laboratory. These differences may modify the pattern of results and limit the degree to which these studies can be compared. Moreover, Mattock and Burnham's sample infants were on average bilingual, who may maintain more open-ended phonological systems (see Kuhl, 2011), whereas Yeung et al. tested monolingual infants. Finally, it is possible that the Thai tone contrasts employed by Mattock and Burnham, in addition to being highly salient in general, may have assimilated to native tone contrasts in Chinese, in which case discrimination of non-native Thai tones by Chinese infants would be predicted. The mapping between non-native stimuli and the native phonological inventory is a well-attested determinant of early phonological attunement of vowels and consonants (Best, 1995). Such relation may also determine phonological attunement to tone.

#### *Discrimination of Lexical Tone in Nontone Language Learners*

While tone perception has been investigated in tone language learners, a useful corollary is to investigate how nontone language learners perceive lexical tones. Every language incorporates pitch contrasts in the form of intonational variation. In nontone language learners, it is possible that the specific type of experience infants acquire with pitch variation in their native language may determine responses to lexical tones. For example, one might expect tone sensitivity to vary for stress- and syllable-timed language learners given that stress-timed languages contain greater pitch variation. This question was posed by Mattock, Molnar, Polka, and Burnham (2008) in a study involving 4-, 6-, and 9-month-old English- and French-learning infants, learning stress- and syllable-timed languages, respectively. Infants were tested on their sensitivity to Thai tone contrasts. Both French- and English-learning infants demonstrated a similar pattern of performance: All infants could discriminate tones at 4 and 6 months, but not at 9 months of age. This suggests that sensitivity to tone in nontone language learners is not contingent on the extent to which pitch variation is incorporated into the native language.

A series of studies have reported perceptual attunement to lexical tone in the 1st year of life (Harrison, 2000; Mattock & Burnham, 2006; Mattock

et al., 2008; Yeung et al., 2013) as previously reported for vowels and consonants. However, in striking contrast to findings with vowels and consonants, the decline in tone discrimination appears to be transient and subject to subsequent reversal, demonstrated in a recent study by Liu and Kager (2014). Liu and Kager sampled nontone-language-learning infants and charted tone discrimination at several age groups between 5 and 18 months. Their investigation of infants' sensitivity to a highly discriminable Mandarin tone contrast (Tones 55 and 51) revealed a U-shaped discrimination pattern emerged across age groups. Specifically, infants between 8 and 12 months showed the weakest evidence of discrimination, whereas discrimination was relatively strong prior to 8 months and, perhaps more intriguingly, after 12 months. This finding stands in sharp contrast to prior studies with vowels and consonants, many of which have consistently demonstrated a decline in discrimination of non-native contrasts after 10–12 months (Kuhl et al., 1992; Polka & Werker, 1994; Werker & Tees, 1984). Liu and Kager then manipulated the tone contrasts to systematically reduce differences in fundamental frequency, making their stimuli harder to contrast. In response to this, only the youngest (5–6 months) and oldest (17–18 months) infants succeeded in discriminating similar sounding tone contrasts. Therefore, it appears that tone perception in learners of nontone languages assumes a non-normative developmental pattern in relation to that documented for vowels and consonants.

Liu and Kager (2014) further investigated the degree of flexibility associated with tone discrimination within the period of decline (11–12 months) by assessing distributional learning of tone categories during this phase. Liu and Kager familiarized 11- to 12-month old nontone-language-learning infants with two distributional profiles for tone. One profile was a bimodal distribution, previously shown to elicit the formation of two phonetic categories, and the other was a unimodal distribution, previously shown to elicit a single phonetic category (Maye, Werker, & Gerken, 2002). In their sample of 11-month old infants learning a nontone language (Dutch), the bimodal distribution resulted in two tone categories and the unimodal distribution resulted in a single category. This provides empirical evidence that the reduction in tone discrimination observed between 11 and 12 months is associated with residual flexibility and openness to tone as a source of phonemic contrast such that tone categories could be created based on varying statistical properties of the input.

### *A Summary of the Development of Phonetic Categories for Tone*

On account of recent research on the acquisition of the phonetic category structure for tone, a developmental pathway to tone language attunement is gradually coming into focus. As compared to consonant and vowel systems, the trajectory of phonological attunement appears quite distinct for lexical tone. In particular, what was traditionally conceived of as a language-general stage of phonological development (birth to 6 months) appears not to be so for tone-exposed infants, who already show primitive lexical tone categories based on ambient language exposure as early as 4 months. A natural question to emerge from these findings is why tone categories are evident prior to vowel and consonant categories for native learners. While there are no definitive answers to this question as of yet, we proposed three candidate reasons for the early establishment of tone categories.

First, tone is largely driven by variation in vocal pitch. Infants show an early inclination toward vocal pitch variation in speech, which is thought to underlie a range of listening preferences, such as a preference for infant-directed prosody, for native language prosody, and for emotional prosody (Fernald, 1992; Gleitman, Newport, & Gleitman, 1984; Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009; Singh, Morgan, & Best, 2002). A heavy attentional focus on pitch variation at an early age may lead to early analysis of elements of phonology that are convolved with pitch, such as lexical tone. As such, sources of variation driven by pitch shifts may be more precociously analyzed than those driven by segmental variation, such as vowels and consonants.

A second possibility for the early establishment of tone categories relates to processing load. Of the tone languages sampled in empirical research thus far, there are fewer tones than consonants or vowels. For example, there are 21 onset consonants, 5 monophthong vowels (with several additional diphthongs and triphthongs), and only 4 lexical tones in Mandarin Chinese. Likewise, Thai has 21 consonants, 24 vowels, and 5 tones. This may greatly reduce the learning burden for deriving and consolidating tone inventories relative to consonant and vowel inventories and may contribute to the early emergence of tone sensitivities.

Third, the status of tone as a suprasegmental unit of speech may favor its early analysis and acquisition. Suprasegmental phonology is accessible to infants much earlier than segmental phonology

and is arguably the only source of phonological information available prenatally (e.g., Mehler et al., 1988). There is evidence for fetal learning of suprasegmental information prenatally (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998). Furthermore, suprasegmental cues are prioritized over segmental cues in linguistic analyses when they are placed in competition with one another (e.g., Matys, Jusczyk, Luce, & Morgan, 1999). It is therefore possible that a natural scalar ordering exists in the development of phonetic categories from global units to local units in terms of ease of computation, which may favor the precocious analysis of tones.

Each of the candidate explanations presented above await further empirical evidence in order to convincingly explain why infants may get a “head start” with lexical tone and furthermore, why tone attunement may be associated with an extended and more complex timetable (Liu & Kager, 2014). In the next section, the link between tone perception and vocabulary development will be explored.

### *Integration of Tone Into Word-Level Representations: Evidence From the Lexical Level*

The integration of tone into the native vocabulary introduces a different set of demands for tone and nontone language learners: Nontone language learners must disregard pitch contrasts in searching for the lexical identity of a word, and tone language learners must selectively incorporate phonemic pitch variation in determining the meanings of words. Both of these processes have been investigated in recent research.

### **Lexical Integration of Tone (Nontone Language Learners)**

With regard to the first issue of whether nontone language learners disregard tone variation when hearing words, there have been several studies to investigate this issue using words varying in vocal pitch. At the earliest age group, Singh, White, and Morgan (2008) familiarized English-learning infants at 7.5 and 9 months with monosyllabic words. Infants were then exposed to pitch-matched and pitch-mismatched forms of the same words in the context of test passages. At 7.5 months, infants only equated pitch-matched words and it was not until 9 months that infants equated pitch-mismatched forms. At these ages, infants were not expected to have meaning associated with the word forms they heard and their task was primarily

to find repetitions of words in speech and not to demonstrate an association between these words and their meanings. Nevertheless, infants underwent a developmental progression in their integration of pitch into native word forms between 7.5 and 9 months.

To address the possibility that pitch may be more closely integrated in a word-learning situation, Quam and Swingley (2010) tested recognition of newly learned word forms when the words were pitch matched and pitch mismatched in 30-month-old English-learning toddlers. In their design, participants were taught a new nonword, a “*deebo*,” as a label for a new toy. During a subsequent test phase, the target word was presented identically to the trained pronunciation, with a change in pitch, or with a change in vowel (i.e., “*dahbo*”). Toddlers showed linguistically mature responses to vowel and tones, selecting the target as a referent when it was labeled with the correct pronunciation and when it was labeled with a tone variant of the correct pronunciation. They did not accept a label for the target that had undergone a vowel change. This suggests that by 2.5 years of age, learners of a nontone language have already categorized vocal pitch changes as lexically irrelevant.

It should be noted that the pitch variants used by Singh et al. (2008) and Quam and Swingley (2010) were not actual tone contrasts. They were pitch contrasts that were not drawn from any lexical tone inventory. It is possible that nontone language learners would have responded differently to true lexical tone contrasts in a word-learning context. In a study employing lexical tone contrasts drawn from Mandarin Chinese, Singh, Tam, Chan, and Golinkoff (2014) familiarized nontone language learners with novel objects labeled with novel words. Infants were then tested on recognition of learned words via a preferential looking paradigm when labels were correctly pronounced, altered due to a change in tone (Mandarin Tones 35 vs. 51) or altered due to a change in vowel. English-learning infants were tested at 18 and 24 months. As early as 18 months, English-learning infants linked words to their tones, such that tone variants were treated as mispronunciations. Later at 24 months, English-learning infants accepted tone variants as differing realizations of the same word. A complementary set of findings was obtained by Hay, Graf Estes, Wang, and Saffran (2015) using a habituation-based paradigm: Hay et al. (2015) demonstrated recognition on the part of English-learning infants that lexical tones did not determine meaning at 19 months.

### Lexical Integration of Tone (Tone Language Learners)

Several studies have investigated the companion question of when tone language learners integrate tone into words. In a study designed to investigate whether tone language learners encode lexical tone in infancy, Singh and Foong (2012) tested English-Mandarin bilingual infants on their abilities to recognize familiarized English words varying in pitch and Mandarin words varying in tone. Each infant went through two testing sessions, one in English and one in Mandarin. During each session, infants were familiarized with two words and then tested on recognition of each word in sentences. In English test sentences, one word was changed in pitch from the familiarization set and one word was matched in pitch. In Chinese test sentences, one word was changed in lexical tone from the familiarization set and one word was matched in tone. Infants were tested at three ages: 7.5, 9, and 11 months. At 7.5 months, infants differentiated words based on tone contrasts in Mandarin and pitch contrasts in English. At 9 months, infants equated tone contrastive forms in Mandarin and equated pitch contrastive forms in English. However, at 11 months, infants distinguished tone variants in Mandarin, but equated pitch variants in English, demonstrating language-specific responses to pitch movements across each of their languages. These findings suggest bilingual infants undergo an important transition between 7.5 and 11 months in appreciating the lexical role played by pitch in each of their native languages.

A direct comparison of how tone and nontone language learners respond to tone in word-learning tasks was drawn by Singh et al. (2014). In this study, 18- and 24-month-old infants were trained on novel object–word pairings and tested on recognition of these words when they varied in tones and vowels. Performance was compared across English-learning infants and Mandarin-learning infants with each group being tested in their native language. Results demonstrated that both English and Mandarin learners demonstrated similar patterns of tone integration to one another at 18 months: Both groups defined newly learned words by tone and vowel quality in equal measure. At 24 months, English and Mandarin learners diverged in their interpretation of tone. At this point, Mandarin-learning infants differentiated words based on tone and vowel changes. English-learning infants no longer differentiated words

based on tone contrasts, although they did distinguish words based on vowel contrasts.

While Singh et al. (2014) used a single, highly salient contrast, the effects of less salient contrasts on word recognition were investigated by Gao, Shi, and Li (2011) in a preferential looking paradigm. When tested on their recognition of known words, Mandarin-speaking toddlers (19–26 months) did not appear to detect a subtle tone substitution (Tone 35–214), whereas they did detect a more salient tone contrast (Tone 35–51, which was the same contrast employed by Singh et al., 2014). This suggests that sensitivity to tone contrasts at the lexical level is potentially modified by the salience of a particular tone contrast.

In a direct comparison of the effects of variation in tones, vowels, and consonants on spoken word recognition in Mandarin-learning children, Singh, Goh, and Wewalaarachchi (2015) investigated recognition of familiar words in Mandarin-speaking toddlers between 2.5 and 3.5 years of age. All trials were conducted in Mandarin. On half the trials, the name of the target was correctly pronounced. The other half of the trials entailed a substitution of the correct label with a mispronounced label by changing the vowel, initial consonant, or tone of the original label. Results demonstrated that toddlers recognized all three types of changes to be mispronunciations. However, in terms of their relative sensitivity to each type of mispronunciation, toddlers were highly sensitive to tone mispronunciations and less sensitive to vowel and consonant mispronunciations. It should be noted that preschoolers were also sampled in this study (4- to 5-year-olds). Preschoolers demonstrated an increase in sensitivity to vowel and consonant variation (as one would expect in an older sample), yet their sensitivity to tone variation was reduced as compared to the toddler sample. This reduction in sensitivity with age is reminiscent of conclusions drawn by Liu and Kager (2014) with regard to phonetic discrimination, specifically, that tone categories may emerge early, but that they undergo considerable subsequent refinement along an extended trajectory as compared to vowels and consonants.

#### *A Summary of the Development of the Lexical Integration of Tone*

Findings of studies investigating the integration of tone into words reviewed thus far point to an early and de facto sensitivity to tone as a potential determinant of meaning. This is based on convergent evidence from auditory word recognition,

novel word learning, and familiar word recognition demonstrating a high degree of sensitivity to tone as a determinant of lexical identity. In broad terms, the pattern of results obtained from the lexical level are consistent with those obtained at the phonetic level. In phoneme discrimination, infants appear to show an early sensitivity to lexical tone in comparison to vowels and consonants, but tone sensitivity appears to linger beyond that of vowels and consonants (Liu & Kager, 2014). Correspondingly, in word-learning tasks, infants appear highly sensitive to tone and encode tone information even if their native language does not (Singh et al., 2014) although tone sensitivity appears comparatively low in preschool children as compared to toddlers (Singh et al., 2015). Tone therefore appears to feature prominently in the speech percept in infants and toddlers and to exert potent influences on early perception in tone and nontone language learners alike. However, based on evidence available to date, the trajectory of tone acquisition in perception appears more complex than that associated with phonetic segments (see Figure 2 for an integrated timeline of sensitivity to tone in phoneme discrimination and lexical processing tasks).

In addition to developing a perceptual framework that is aligned with the native tone system, tone language learners must be able to produce lexical tones accurately. While the empirical study of tone perception in early language learners is relatively recent, tone production has been systematically investigated over the past 40 years using analyses of individual cases, corpora, and experimental approaches. As in the case of perception, tone production appears to assume a different maturational course to vowels and consonants.

### **The Development of Tone Production**

As with perception, studies on tone production focus largely on Chinese languages, such as Mandarin and Cantonese with few exceptions. Prior studies in tone productions have proceeded in two directions: establishing a developmental trajectory of when children acquire an awareness of the native tone inventory and on the order of acquisition of specific tones in vocal production. Each of these areas will be treated in sequence.

#### *The Developmental Trajectory of Tone Production*

Early studies in tone production focused on closely examining tone productions in individual

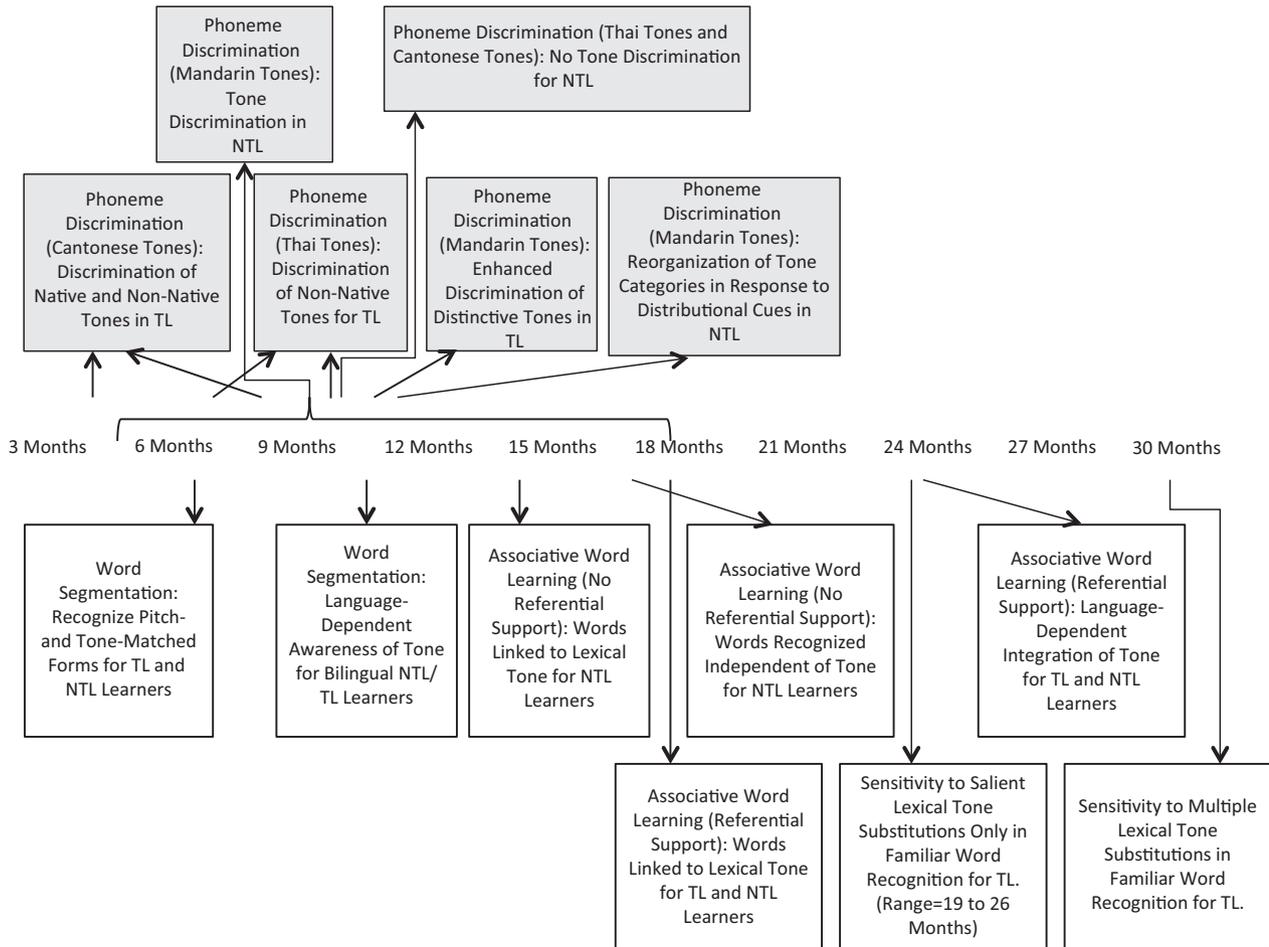


Figure 2. A timeline of the acquisition of tone in phonetic discrimination and lexical integration tasks. TL = tone language learners; NTL = nontone language learners.

children's vocalizations. A case study by Clumeck (1977) documented the production of lexical tones in protowords in one infant exposed to Mandarin Chinese in the 2nd year of life. Prior to the use of tone in association with lexical identity, Clumeck noted the earlier intentional use of pitch as a source of prosodic contrast in the conveyance of emotion. He also observed that the child appeared to have a "hypothesis" that words have a characteristic pitch by the end of the 2nd year of life, demonstrated by the child's adoption of rising pitch contours as a feature of particular Mandarin words. Clumeck observed two additional children from Mandarin Chinese-speaking families (Clumeck, 1980), one in the 2nd year of life and the other in the 3rd year of life. He noted that both children could produce all the tones in the Mandarin inventory throughout the period of observation. However, he also noted that it took a much longer time for the children to distinguish phonemic tone variation from allophonic

tone variation and to therefore demonstrate mature tone knowledge in production.

Case studies by Clumeck (1977) and others (see also Chao, 1951; Tse, 1978; Tuaycharoen, 1977, for additional case studies of tone production) have charted spontaneous tone productions of children over the first 2–3 years of life. In each of these case studies, however, interpretation of findings is arguably limited by several factors. Specifically, the focus on a single child, the absence of comparison of children's productions to mature tone forms, and an absence of more objective measures of early tone productions limit the scope of interpretation. Nevertheless, these initial forays into native tone production laid important groundwork for further empirical studies in this area.

The first study of tone acquisition using a group sample was conducted by Li and Thompson (1977) in which a cohort of 17 Mandarin-Chinese-learning children were observed between the ages of 1.5 and

3 years over a period of 7 months. Sessions consisted of a picture-naming task designed to elicit tone productions. From records of tones produced over time, Li and Thompson observed a clear progression in the emergence and refinement of early tones. Based on their data, they arrived at a four-stage model of tone acquisition. At Stage 1, single-word utterances prevail and the early tone inventory is relatively sparse. High tones (55) and falling tones (51) are the primary tones produced. Stage 2 is characterized by children continuing to produce single-word utterances, and although all the tones had surfaced by this stage in Li and Thompson's cohort, there remained tone confusion between Mandarin Tones 35 and 214 (rising and dipping tones). Evidence of confusion between these tones persisted into Stage 3, also marked by the advent of multiword utterances. At this stage, however, children demonstrated a primitive understanding of how phonological context can alter the realization of a tone. Maturation of the tone system was evident at Stage 4 where sentences were produced, with each word receiving correct tone assignment. It should be noted that Li and Thompson did not present age ranges for each stage. However, a conclusion drawn from their study was that the tone system is acquired precociously relative to consonants and vowels in production.

Using a similar method designed to chart age-related developments in tone production, Hua and Dodd (2000) investigated tone production via picture naming in a study of 129 monolingual Mandarin learners between the ages of 1.5 and 4.5 years in a cross-sectional design. In contrast to Li and Thompson (1977), and Hua and Dodd adopted a different approach, attempting to compare tone, consonant, and vowel productions. The authors explored productive speech errors committed by the participants. By the time the children were 4.5 years old, 90% were able to articulate the entire inventory of 21 syllable-initial consonants. Additionally, while vowels emerged very early in development such that even the youngest children (1.5 years) were able to produce simple vowels, vowel reduction was evident throughout all groups. In comparison to vowels and consonants, tone errors were extremely low in frequency and only observed in the youngest age group (1.5 years). Thus, the authors concluded that tone is the very first type of phonetic segment to be fully acquired and syllable-final consonants were the last to be acquired. Using a different tone inventory to ask a similar set of questions about tone production, So and Dodd (1995) investigated tone production with

a larger tone inventory (Cantonese). Although Cantonese has a comparatively dense tone space relative to Mandarin Chinese, Cantonese learners showed a similar pattern of early tone acquisition with very few errors in tones or vowels registered after 2 years of age. In contrast, consonants were only reliably produced between by 5 years of age. Finally, early acquisition of tone relative to vowels and consonants was further corroborated in a large-scale population-based study from children aged 2 to 12 years (To, Cheung, & McLeod, 2013). In this study, tones were largely acquired by 2.5 years of age and the majority of vowels and consonants were acquired by 5 and 6 years of age, respectively. Taken together, these studies arrived at a common conclusion that tone is early acquired followed by vowels and then by consonants and that the early acquisition of tone is observable across different tone inventories.

While there appears to be general consensus that relative to vowels and consonants, lexical tones are produced early with considerable accuracy, judgments of tone errors were typically made by human observers in a categorical fashion (correct or incorrect). Innovations in speech analysis tools have afforded greater precision in evaluation of early tone productions. Acoustic analyses allow for children's productions to be compared with adult productions as well as for children's productions to be filtered such that the pitch contour underlying a tone production becomes the dominant cue in accuracy judgments. A move toward acoustic analyses of tone productions has revealed a different conclusion, specifically, that tone contrasts may be early to emerge in production but that the process actually takes a relatively long time to mature as compared to vowels and consonants. Wong, Schwartz, and Jenkins (2005) analyzed speech tokens of child productions (at 3 years) collected in a picture-naming task. All speech tokens were processed with a filter to remove segmental information and to preserve only the pitch contours to be judged by native Mandarin speakers. Results revealed that children were significantly less accurate than adults in producing target tones in spite of the fact that they could accurately perceive most tones. Even at 3 years of age, child productions were qualitatively different from that of adults.

Similar to adult ratings employed by Wong et al. (2005), acoustic analyses on the same data demonstrated significant differences between children's productions and adult productions (Wong, 2012a). These differences appeared regardless of whether tone language learners were raised in a tone-lan-

guage-speaking environment or as immigrants to a nontone language environment (Wong, 2012b). Follow-up analyses with older children revealed that tone mastery is still incomplete by 4–5 years of age with little developmental change in tone production observed over the preschool years (Wong, 2013).

The suite of studies by Wong (Wong, 2012a, 2012b, 2013; Wong et al., 2005) has contributed valuably to prior research by revealing a different maturational profile of tone production when children's productions are subject to more detailed scrutiny. In Wong's studies, ratings by blind adult raters as well as acoustic analyses of the pitch contour point to a relatively lengthy course of acquisition, spanning at least the first 5 years of life. This finding stands in contrast to earlier studies that claim very early acquisition of stable tone productions (Hua & Dodd, 2000; Li & Thompson, 1977; So & Dodd, 1995). Furthermore, Wong's studies suggest that tone does not necessarily mature in advance of vowels and consonants in production (Wong, 2012a, 2013). It is noteworthy that Wong's studies primarily focused on pitch cues, such as  $F_0$  height and contour, but did not include any analyses of secondary cues of tone. Thus, it is unclear whether children are able to produce secondary cues reliably and the extent to which these cues may have contributed to the perception of accuracy in prior studies.

#### *Order of Acquisition of Individual Tones*

Thus far, our review of studies has focused on the acquisition of a tone system in its entirety. However, as with tone perception, mastery of the tone system is not uniform across the entire inventory. Rather, there are observable differences in the timing of acquisition of particular tones. Integrating across multiple studies of tone production (Wong, 2012a, 2012b, 2013; Wong et al., 2005), Wong proposed an order of acquisition from Tone 51 to Tone 55 to Tone 35 to Tone 214 (from earliest to latest acquired). A second consistent finding is that within this ordering, there appears to be a sharp decrement in accuracy for Tone 214 relative to Tones 55, 35, and 51 (see Wong, 2013; Wong et al., 2005).

There are several proposed reasons for the asynchronous acquisition of tones. One account derives from examining muscular involvement in production of different tones (Wong, 2012a, 2013). The earliest acquired tone (Tone 51) demands the least in terms of motor control of the laryngeal muscles

(Wong, 2013) and is primarily defined by relaxation of the cricothyroid (CT) muscle. At other end of the continuum of tone acquisition, Tone 214, which is particularly late acquired, appears to demand greater motor sequencing and control, requiring initial activation of a different muscle, the sternohyoid (SH) muscle, followed by a decrease in SH activity and an increase in CT activity (see Wong, 2012a, for a detailed description of laryngeal muscular involvement for each tone). There is of course no conclusive evidence demonstrating that differences in muscular involvement cause the asynchronous pattern of acquisition observed across tones. However, it is possible that the physiological complexity of tone productions influences the precision with which they are articulated in early childhood, as demonstrated for vowel and consonant productions (see Smith, 2006, for a review).

A second explanation for differences in timing of acquisition of individual tones relates to their linguistic and psychoacoustic properties. This perspective is based largely on the late acquisition of Tone 214, which may be inherently unstable because of tone sandhi rules. While there are several tone sandhi rules, there is a rule that applies specifically to Mandarin Tone 214. When two syllables, each associated with Tone 214, are produced in succession, the first syllable is substituted for Tone 35. As a result, the first syllable of a disyllabic unit undergoes a whole-tone substitution from 214 to 35, which could cause confusion between these tones for young learners. In addition, Tone 214 is often produced in a truncated form (with the contour 21). The full tone (i.e., with the 214 contour) is usually only used when a target word is produced in isolation or at the end of an utterance. As such, a child might hear the full tone produced less frequently than each of the other tones. Wewalaarachchi and Singh (2014) reported that resolution of tone sandhi rules in word recognition emerges relatively late between 4 and 5 years of age, even though children were familiar with the words that underwent tone sandhi as early as 2–3 years of age. It is therefore possible that uncertainty about the form of Tone 214 may linger in young children based on its susceptibility to substitution on account of sandhi rules. Finally, Chen (2013) reports low sensitivity to tones involved in sandhi (35 and 214) both in tone language and nontone language speakers alike. This invites the alternative possibility that there may be universal perceptual biases that render these tones less distinctive and therefore more challenging to produce.

In summary, Clumeck's (1980) initial observation that while tone contrasts are evident in infancy, this does not justify the conclusion that infants truly possess sophisticated knowledge of tone was indeed a prescient insight. A collection of subsequent studies investigating tone productions using adult judgments and acoustic analyses, most notably those by Pui-san Wong, have corroborated the early presence of tone contrasts in children's early productions and their appearance preceding consonants and vowels, yet these studies have also demonstrated that tones take several subsequent years to mature. These studies lead us back to Clumeck's original observation that tone categories emerge early but take several years to fully "crystallize" and to reach adult-like targets.

### **From Research to Theory: Implications for Theories of Language Development**

As reviewed earlier, several aspects of tone acquisition are unique in comparison to consonant–vowel acquisition in the domains of both perception and production. This suggests that we cannot easily generalize from learners of consonant–vowel systems to learners of tone systems in some aspects of first language acquisition. In light of this, three ways in which prevailing models of language development could be modified to accommodate lexical tone are proposed.

First, psycholinguistic research has historically drawn a critical distinction between the acquisition and processing of segmental and suprasegmental information. Segmental variation is often equated with phonemic variation. In contrast, suprasegmental variation is often defined at the level of the syllable or phrase and is commonly construed to be nonphonemic variation. However, in tone language systems, lexical tone is a suprasegmental cue that fulfills a phonemic function, akin to segments. Consequently, arriving at a correct linguistic interpretation of segmental and suprasegmental cues, a hallmark of language development that is integrated into existing theoretical models (e.g., Curtin et al., 2011), is arguably more complex for the tone language learner. Specifically, for tone language learners, the set of cues that may be useful in distinguishing segmental and suprasegmental variation differs from those that may be useful for nontone language learners. Current and future models could be enhanced by incorporating cross-linguistic variation in how phonemic variation is defined in languages, encompassing phonemic

variation based on segmental (i.e., vowels/consonants) as well as suprasegmental (i.e., tone) cues.

Second, a potential complication to acquiring tones is the issue of multifunctionality. In tone languages, lexis is just one of the multiplex of functions served by pitch variation. Such languages also use pitch variation to mark pragmatic functions, such as marking question and statement distinctions, assignment of prosodic focus, and conveyance of emotional prosody (Liu & Xu, 2005; Wang, Li, & Fang, 2005; Yuan, 2011). One would expect conflicting tone and intonation relations to cause lexical ambiguity and potential confusion. Indeed, tone–intonation relations can be difficult for adult speakers of a tone language to negotiate and reconcile (Yuan, 2011) and one might suspect such difficulties would only be compounded earlier in development. Developmental accounts of how young children partition sources of pitch variation into discrete linguistic and communicative functions have not been advanced, yet this requirement presents a potential learning challenge to the young tone language learner that remains unexplored. It is possible that it is this very challenge that accounts for the protracted course according to which tone knowledge appears to stabilize. In sum, theories of language development should optimally incorporate a mechanism to disaggregate vocal pitch in tone languages into its constituent functions.

Finally, in several respects, the course of acquisition charted for language development does not accommodate findings obtained from tone language learners. For instance, one of the most widely documented developmental transitions in infancy—attunement to native phonological categories—appears to assume a different trajectory with regard to lexical tone (Liu & Kager, 2014; Yeung et al., 2013). In particular, the notion that infants learning a nontone language undergo a decline in sensitivity to lexical tones, followed by a resurgence in sensitivity to tone, points to a unique developmental profile that has not been documented for vowels and consonants in monolingual infants. These findings suggest that the time course linked to normative phonological development may not extend to lexical tone. Theories of typical language development merit expansion to accommodate unique aspects of tone languages that may shift the timetable of phonological attunement to tone.

Although lexical tone is a highly frequent feature of human languages, theories of language development have traditionally side-stepped the role of tone in defining the acquisition of words and sounds. This may be attributable to historical

traditions that have long prevailed within developmental psycholinguistics. For example, foundational theories of phonological systems, such as generative phonology (Chomsky & Halle, 1968) and distinctive feature theory (Jakobson, 1962) were derived for English. These theories were engineered to describe all languages and indeed, they were successfully applied across a diversity of languages such as Russian, Hebrew, Japanese, French, and Spanish. However, such theories are tightly bound to consonants and vowels and do not account for lexical tone. As such, proposed theoretical adaptations, most notably Goldsmith's (1976) theory of autosegmental phonology and Wang's (1967) proposal for distinctive features for tone (see Clements, Michaud, & Patin, 2011, for a discussion of tone features), significantly advanced our understanding of human languages by integrating tone. Introducing theoretical crosswinds to traditional psycholinguistics, such theories inspired a body of experimental research on the role of tone in early language acquisition. This body of experimental research has shown us that the acquisition of tone is distinct from more commonly studied phonetic segments and that lexical tone acquisition may be subject to a unique set of constraints in the early language learner. The uniqueness of tone in language acquisition is discussed below, followed by a discussion of how exceptional properties of tone acquisition can be integrated into formal models of language development.

*The Uniqueness of Lexical Tone Acquisition: Evidence for Precocity and Protraction*

In the development of production and perception, tone appears to stand out as exceptional among other components of phonology. In particular, tone appears to assume a comparatively lengthier and arguably more complex course of acquisition. Tone categories seem not only to emerge precociously in several domains of early language acquisition (Li & Thompson, 1977; Singh et al., 2014; Singh et al., 2015; To et al., 2013; Yeung et al., 2013), but they also seem to stabilize more gradually than consonant and vowel categories (Liu & Kager, 2014; Singh et al., 2014; Singh et al., 2015; Wewalaarachchi & Singh, 2014; Wong, 2013). Perhaps the most striking contrast with vowel and consonant sensitivity is that tone sensitivity has been shown to remain strong for nontone learning infants over a lengthier period of development than vowels and consonants (Liu & Kager, 2014).

The findings of Liu and Kager's (2014) study complement a groundswell of evidence from of adult studies demonstrating persistent tone sensitivity in speakers of nontone languages. For example, non-tone-speaking adults have been shown to discriminate lexical tones with reasonably high levels of accuracy even though tones do not draw lexical distinctions in their language (although the basis for tone judgments has been shown to differ for tone and nontone language learners; Francis, Ciocca, Ma, & Fenn, 2008; Hallé, Chang, & Best, 2004). In contrast, several studies have reported poor discrimination of non-native segmental contrasts in adults (e.g., Miyawaki et al., 1975; Werker, Gilbert, Humphrey, & Tees, 1981; see Best & Tyler, 2007, for a detailed review on attenuation of non-native segmental discrimination). Comparisons across studies have revealed that adults can make great strides in learning non-native tone contrasts in the absence of prior tone language experience (So, 2006; Wang, Spence, Jongman, & Sereno, 1999) in contrast to the relatively modest gains observed when adults are charged with learning non-native segments (e.g., Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively, & Pisoni, 1991). In a direct comparison of adults' representation and perception of non-native tones and non-native segments, Zhang, Samuel, and Liu (2012) reported an advantage for tones over segments.

In addition to detecting tone contrasts, nontone language learners have also been shown to be able to integrate contrastive tones into newly learned words after relatively brief amounts of training (e.g., Wong & Perrachione, 2007) and were as sensitive (in the aggregate) to tone contrasts as adults learning a tone language (Francis et al., 2008). In a recent study by Antoniou and Wong (2015), nontone language speakers were better able to integrate tones into newly learned words than a non-native consonant contrast, although it should be noted that this study employed a small set of contrasts. In sum, across a number of studies, there appears to be an emerging consensus that tones may assume a unique representational status. This conclusion is based on developmental research suggesting that tone sensitivity appears particularly strong both in phoneme discrimination and word recognition tasks and on adult studies, where basal and acquired tone sensitivity has been demonstrated to be relatively high even if tone is not encoded in the native language. In combination, these findings invite the possibility that tone may be privileged in terms of perception and learnability relative to phonetic

segments. We offer two candidate reasons for this potential asynchrony.

One possibility for the early emergence of tone and for the maintenance of tone sensitivity relates to the provenance of tone: vocal pitch. As noted by Hallé et al. (2004) and others, tone contrasts surface in every language in the form of intonational contrast, which are also marked primarily by pitch movements. For example, pragmatically meaningful distinctions, such as questions and statements, are reliably distinguished by pitch across tone and nontone languages alike (Bolinger, 1978). Likewise, the use of pitch to signal vocal affect is robust across languages (Lieberman, 1967). This may contribute to a functional and adaptive sensitivity toward vocal pitch in human languages that is independent of whether it is integrated into the native language as a source of lexical contrast. Indirect support for this hypothesis comes from evidence for an apparent loss of sensitivity to suprasegmental features other than pitch (i.e., stress) that are truly absent from the native language (Dupoux, Pallier, Sebastian, & Mehler, 1997), inviting the possibility that the broad communicative scope of pitch may preserve pitch (and tone) sensitivity in tone and nontone speakers.

The centrality of pitch in auditory processing is evidenced at various stages in development. At the earliest phases of auditory perception, pitch, along with stress and rhythm, are preferentially available to infants prior to birth (Fifer & Moon, 1988). Over the months following birth, infants are astutely tuned into various sources of pitch variation (Fernald & Kuhl, 1987). The primacy of pitch early in development is followed by a continuing sensitivity to pitch throughout the life span in tone and nontone language speakers alike (e.g., Lee & Nusbaum, 1993). A foundation for lexical tone sensitivity that is rooted in pitch perception is perhaps most clearly attested by demonstrations that nontone-speaking adults are able to profit from training on lexical tone contrasts, but that their success at this task was associated with their general auditory sensitivities to pitch (e.g., Wong & Perrachione, 2007). Overall, the influence of pitch appears remarkably potent early in development and appears resistant to decay even when it is downweighted in the native language. It is possible that lexical tone rides on a dimension of speech—vocal pitch—that begins as and remains an influential source of variation in language processing to draw communicatively relevant distinctions and that this accounts in part for the unique profile of tone acquisition.

A second possibility for the precocious analysis of tone—not mutually exclusive from the first—may derive from its overall acoustic salience within a tone language. Specifically, tone is encoded at the level of the syllable above the level of the segment and is obligatorily applied to every syllable in a tone language. In contrast, consonants, which appear to be emerge relatively late in perception and in production (e.g., Hua & Dodd, 2000; Werker & Tees, 1984), are comparatively transient and are optional elements of syllables. As such, tones may be easier to track and analyze compared to phonetic segments, such as consonants, by virtue of their acoustic properties and integrality to the syllable. It is possible that acoustic salience contributes to the early analysis of tone as a linguistic feature, as suggested in prior studies demonstrating early uptake of tone information (Hua & Dodd, 2000; Yeung et al., 2013).

The primacy of pitch in human perception and the relative salience of tones relative to other phonetic segments are two candidate (and possibly related) accounts for precocity in tone acquisition, yet further research is required to draw firm conclusions on precisely why tone appears exceptional in the course of acquisition. What remains evident is that lexical tone exhibits a unique course of acquisition in comparison to other aspects of phonology.

#### *Formal Models of Language Acquisition: Where Does Lexical Tone Fit?*

Lexical tone has currently not yet been integrated into existing developmental models of language development. However, a recent initiative has been launched to adapt a prevailing model of psycholinguistic processing—the TRACE model (McClelland, Mirman, & Holt, 2006)—to tone languages. Due to its emphasis on interactivity between tiers of linguistic representation (phonetic features, phonemes, words), the TRACE model is uniquely well suited to the natural variation that exists within human languages in the acoustic composition of phonemes as well as in the relations between features, phonemes, and words. An adaptation of TRACE, TTRACE, was recently proposed by Tong, McBride, and Burnham (2014) to account for lexical tone in speech processing. In contrast to its progenitor, TRACE, TTRACE builds in “tonemes” (phonetic segments corresponding to lexical tones) at the phoneme tier. Furthermore, TTRACE expands on the repertoire of features used to define phonemes in TRACE by including pitch characteristics that define lexical tone. Finally, a central postulate of

TTRACE is a unique set of relations between features, phonemes, and words that reflects the underlying structure of tone language systems.

Although TTRACE has recently been formulated and awaits validation via further empirical testing, it represents an important advance in linguistic theory. Specifically, it calls into question the tacit assumption that existing models of language processing are generalizable across human languages and not limited to a relatively small class of typologically similar languages on which they were originally founded. Tong et al. (2014) provide much-needed direction in furthering our understanding of the diversity of systems inherent in human language. At present, as acknowledged by Tong et al., TTRACE does not yet make developmental predictions nor does it posit effects of age on tone language acquisition. However, the rapidly expanding body of empirical work in lexical tone acquisition provides a bedrock of evidence from which testable predictions of models such as TTRACE may naturally derive.

### Future Directions in Tone Language Acquisition

From the body of work reviewed thus far, a clearer view of tone acquisition is emergent. However, there remain many open questions. A proposal for future directions in the empirical study of tone follows below.

#### *The Phonological Specification of Tone*

From our review of the literature, it is clear that tone learners show early phonological attunement for tone. There is also a stable pattern of integration of native tone in word-learning tasks, which is evident as early as 11 months (Singh & Foong, 2012), throughout the 2nd year (Gao et al., 2011; Singh et al., 2014), and sustained into the 3rd year of life (Singh et al., 2015). However, the specificity of early tone categories has yet to be determined. Research by Yeung et al. (2013) has given us first insight into this by demonstrating distinct responses to Cantonese tones by Mandarin and Cantonese learners as early as 4 months of age. However, Mattock and Burnham's (2006) reports of sensitivity to Thai tones in Cantonese learners suggest comparatively nonspecific tone categories. Further research is needed to reconcile these different findings. The broader aim of this would be to progress beyond documenting the process of attunement to lexical tone—which is established to some degree—to

determine the degree of phonological precision ascribed to early tone representations. Such studies could involve within-subjects analyses of tone language learners of their sensitivity to native and nonnative tones to determine the phonological specificity of early tone categories.

#### *Decoupling Lexical Tone and Intonation*

An important and related question relates to the interaction of lexical tone and intonation. Tone–intonation relations need to be resolved in order for learners to acquire a vocabulary that represents words with appropriate phonological definition. As previously highlighted, pitch changes signal intonational or prosodic cues in all languages, but serve the additional function of determining lexical identity in tone languages. Infants appear to be biased toward storing suprasegmental cues in memory, demonstrating retention and integration of pitch even if they are only exposed to nontone languages (Singh et al., 2008) in addition to storing several other types of surface details that do not determine meaning (see Newman, 2008, for a review). Differentiating the multiple functions of pitch in a tone language is an added challenge that tone language learners must overcome to become proficient. However, developmental research in this area is currently lacking. Future studies could consider how infants and children decouple intonational variation from lexical tone by comparing tone learners' responses to intonational phrases that diverge from (or converge with) the acoustic profile of lexical tones. As a first foray into this question, Singh and Chee (2016) compared recognition of words in rising tones (e.g., Tone 35) conveyed in rising intonational phrases (i.e., in questions) versus in falling intonational phrases (e.g., statements) and vice versa. Their findings demonstrated that it was not until 4–5 years of age that children were able to resolve tone–intonation conflicts (Singh & Chee, 2016). This particular conflict has been shown to be challenging for adult speakers of a tone language to negotiate (Yuan, 2011) and it stands to reason that children may encounter difficulty in this respect. Furthermore, it is possible that the processing of differentiating pitch into discrete functions contributes to the late stabilization of tone categories in native learners, although such an account merits further investigation.

#### *Sensitivity to Variation in Cues to Tone Identity*

As has been highlighted in the above review, while tone is primarily identified and driven by

pitch-related cues (Chandrasekaran et al., 2010; Fu & Zeng, 2000; Gandour, 1983, 1984), there are also many secondary cues to tone identity, most notably, duration and amplitude (Blicher et al., 1990; Whalen & Xu, 1992). In fact, it is unlikely that pitch is ever used in isolation to convey lexical tone. However, developmental studies of tone acquisition have focused almost exclusively on pitch cues to tone. Incidentally, pitch cues to tone are commonly obscured in particular speech styles that feature prominently in infant-directed speech, such as singing or whispered speech. Therefore, it is incumbent upon the language learner to learn how and when to integrate cues to tone other than pitch. While mature listeners can readily compensate for the absence of pitch cues (e.g., Liu & Samuel, 2004), there have been no developmental studies to determine how and when this ability emerges in development. Mature listeners clearly possess a fine-grained, multidimensional view of tone (e.g., Connell, Hogan, & Rozsypal, 1983) that is necessary for successful tone perception. However, developmental research has not probed the nature of the developing tone percept in children nor has the pathway to a complex adult-like percept been explored. Further studies aimed at identifying the determinants of tone perception and integration other than pitch would allow us to define early tone representations with greater precision. Furthermore, such an approach toward the empirical study of tone would appropriately recognize the truly multidimensional nature of tones, allowing us to understand how cues to tone other than pitch are perceived and integrated into words.

#### *Coevolution of Tone Perception and Production*

Speech perception and production are presumed to be related constructs, exemplified by the fact that English speakers with a more fine-grained discrimination of phonetic contrast are also likely to produce the same phonemes with a greater degree of acoustic contrast (e.g., Fox, 1982). The extent of linguistic transfer between perception and production points to an important area of research in first language development. The current review of tone production and perception points to several similarities between domains, such as early emergence and late stabilization of tone categories. However, a more precise determination of the nature of the production–perception interface would inform our understanding of the way in which parallel development across these domains may be linked. While some studies have been conducted in this area, they

have focused on atypical populations (e.g., Xu et al., 2010). An examination of production–perception correspondences has not been established in typically developing children. It is possible that tone learners stand to gain from a transfer of discrimination skills between perception and production as has been demonstrated in adults (Wang, Jongman, & Sereno, 2003). A more precise understanding of production–perception links in the developing tone language learner would provide insight into possible mediators of tone acquisition.

#### **Conclusion**

In summary, languages of the world exhibit a natural diversity, which is not reflected in mainstream empirical research or in theoretical foundations of language development. In particular, lexical tone—although predominant in human languages—is quite rarely studied in language development. A recently evolving focus on lexical tone leads us to the conclusion that core processes of tone acquisition, such as the emergence of tone as a phonological category, the integration of tone information in word knowledge, and the stabilization of tone representation within the mature mental lexicon, follow a unique developmental course. These empirical findings invite reflection on theories of early language acquisition—in terms of both timing of events and constraints on early language processing. In sum, empirical studies of tone acquisition in childhood provide strong motivation for an expanded narrative on early phonological and lexical development that incorporates greater linguistic diversity.

#### **References**

- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*, 155–182. doi:10.1177/00238309030460020601
- Antoniou, M., & Wong, P. C. M. (2015). *Varying irrelevant phonetic features hinders learning of the feature being trained*. Poster presented at the annual meeting of the Psychonomic Society, Chicago, IL.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 141–204). York Timonium, MD: York Press.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In*

- honor of James Emil Flege (pp. 13–34). Amsterdam, Netherlands: John Benjamins.
- Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, 18, 37–49.
- Bolinger, D. (1978). Intonation across languages. In J. H. Greenberg (Ed.), *Universals of human language: Vol. 2. Phonology* (pp. 471–524). Stanford, CA: Stanford University Press.
- Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., . . . Webster, D. (2014). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, 1–33. doi:10.1017/S0142716414000496
- Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. *Journal of the Acoustical Society of America*, 128, 456–465. doi:10.1121/1.3445785
- Chao, Y. R. (1930). A system of tone letters. *Le Maître Phonétique*, 45, 24–27. Retrieved from [http://en.cnki.com.cn/Article\\_en/CJFDTOTAL-FYZA198002000.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-FYZA198002000.htm)
- Chao, Y. R. (1951). The Cantian idiolect: An analysis of the Chinese spoken by a twenty-eight-month-old child. In W. J. Fischel (Ed.), *Semitic and oriental studies* (pp. 27–44). Berkeley, CA: University of California Press.
- Chen, A. (2013). *Universal biases in the perception of Mandarin tones, from infancy to adulthood* (Unpublished doctoral dissertation). Utrecht University, Utrecht, Netherlands.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper & Row.
- Clements, G. N., Michaud, A., & Patin, C. (2011). Do we need tone features? In E. Hume, J. Goldsmith, & L. Wetzels (Eds.), *Tones and features* (pp. 3–24). Berlin, Germany: De Gruyter Mouton.
- Clumeck, H. V. (1977). *Studies in the acquisition of Mandarin phonology* (Unpublished doctoral dissertation). University of California, Berkeley, CA.
- Clumeck, H. (1980). The acquisition of tone. In G. H. Yeni-Komshian, J. E. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 1. Production* (pp. 257–275). New York, NY: Academic Press.
- Connell, B. A., Hogan, J. T., & Rozsypal, A. J. (1983). Experimental evidence of interaction between tone and intonation in Mandarin Chinese. *Journal of Phonetics*, 11, 337–351.
- Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, 39, 492–504. doi:10.1016/j.wocn.2010.12.002
- Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 10944–10949. doi:10.1073/pnas.0610848104
- Diller, A. (2004). Resources for Thai language research. In A. V. N. Diller, J. A. Edmondson, & Y. Luo (Eds.), *The Tai-Kadai languages* (pp. 31–82). London, UK; New York, NY: Routledge.
- Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing “deafness” in French? *Journal of Memory and Language*, 36, 406–421. doi:10.1006/jmla.1996.2500
- Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology*, 23, 157–191. doi:10.1017/S095267570600087X
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306. doi:10.1126/science.171.3968.303
- Fernald, A. (1992). Meaningful melodies in mothers’ speech to infants. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262–282). Cambridge, UK: Cambridge University Press.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293. doi:10.1016/0163-6383(87)90017-8
- Fifer, W. P., & Moon, C. (1988). Auditory experience in the fetus. In W. P. Smotherman & S. R. Robinson (Eds.), *Behavior of the fetus* (pp. 175–188). Caldwell, NJ: Telford Press.
- Fox, R. A. (1982). Individual variation in the perception of vowels: Implications for a perception-production link. *Phonetica*, 39, 1–22. doi:10.1159/000261647
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36, 268–294. doi:10.1016/j.wocn.2007.06.005
- Fu, Q. J., & Zeng, F. G. (2000). Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing*, 5, 45–57. doi:10.1179/136132800807547582.
- Gandour, J. T. (1978). The perception of tone. In V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 41–76). New York, NY: Academic Press.
- Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 147–175.
- Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12, 235–261. Retrieved from <http://www.jstor.org/stable/23767002>
- Gao, J., Shi, R., & Li, A. (2011). *Lexical tone representations of Mandarin-speaking toddlers*. Poster presented at the International Congress of Phonetic Sciences, Hong Kong.
- Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, 11, 43–79. doi:10.1017/S030500090005584
- Goldsmith, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis*, 2, 23–68.
- Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32, 395–421. doi:10.1016/S0095-4470(03)00016-0

- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581–616. doi:10.1016/S0024-3841(00)00003-6
- Haudricourt, A. G. (1954). De l'origine des tons en Vietnamien. *Journal Asiatique*, 242, 69–83.
- Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*, 86, 10–22. doi:10.1111/cdev.12269
- Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, 32, 262–274. doi:10.1016/j.infbeh.2009.03.004
- Hombert, J. M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55, 37–58. doi:10.2307/412518
- Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language*, 27, 3–42. Retrieved from [http://journals.cambridge.org/abstract\\_S030500099900402X](http://journals.cambridge.org/abstract_S030500099900402X)
- Hyman, L. M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, 31, 213–238. doi:10.1016/j.langsci.2008.12.007
- Jakobson, R. (1962). Phonetics and phonology. In S. Rudy (Ed.), *Selected writings* (Vol. 1, pp. 464–504). Hague, Netherlands: Mouton.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831–843. doi:10.1038/nrn1533
- Kuhl, P. K. (2011). Early language learning and literacy: Neuroscience implications for education. *Mind, Brain, and Education*, 5, 128–142. doi:10.1111/j.1751-228X.2011.01121.x
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–F21. doi:10.1111/j.1467-7687.2006.00468.x
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–608. doi:10.1126/science.1736364
- Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants* (3rd ed.). Chichester, UK: Wiley-Blackwell.
- Leather, J. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics*, 11, 373–382.
- Lee, C.-Y. (2000). *Lexical tone in spoken word recognition: A view from Mandarin Chinese* (Unpublished doctoral dissertation). Brown University, Providence, RI.
- Lee, L., & Nusbaum, H. C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception & Psychophysics*, 53, 157–165. doi:10.3758/BF03211726
- Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4, 185–199. doi:10.1017/S0305000900001598
- Lieberman, P. (1967). *Intonation, perception and language*. Cambridge, MA: MIT Press.
- Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62, 70–87. doi:10.1159/000090090.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133, 385–394. doi:10.1016/j.cognition.2014.06.004
- Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech*, 47, 109–138. doi:10.1177/00238309040470020101
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076–2087. doi:10.1121/1.410149
- Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89, 874–886. doi:10.1121/1.1894649
- Matisoff, J. A. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman (Ed.), *Consonant types and tone* (Southern California Occasional Papers in Linguistics No. 1, pp. 71–95). UCLA, Los Angeles, CA.
- Matisoff, J. A. (2001). Genetic versus contact relationship: Prosodic diffusibility in South-East Asian languages. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics* (pp. 291–327). New York, NY: Oxford University Press.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10, 241–265. doi:10.1207/s15327078in1003\_3
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106, 1367–1381. doi:10.1016/j.cognition.2007.07.002
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465–494. doi:10.1006/cogp.1999.0721
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, 101–111. doi:10.1016/S0010-0277(01)00157-3
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10, 363–369. doi:10.1016/j.tics.2006.06.007
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178. doi:10.1016/0010-0277(88)90035-2
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of

- linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, 18, 331–340. Retrieved from <http://link.springer.com/article/10.3758/BF03211209>
- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13, 407–420. doi:10.1111/j.1467-7687.2009.00898.x
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756–766. doi:10.1037/0096-1523.24.3.756
- Newman, R. S. (2008). The level of detail in infants' word learning. *Current Directions in Psychological Science*, 17, 229–232. doi:10.1111/j.1467-8721.2008.00580.x
- Polka, L., & Bohn, O. S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41, 221–231. doi:10.1016/S0167-6393(02)00105-X
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421–435. doi:10.1037/0096-1523.20.2.421
- Quam, C., & Swingle, D. (2010). Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language*, 62, 135–150. doi:10.1016/j.jml.2009.09.003
- Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Rattanasone, N. X., & Best, C. T. (2014). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Attention, Perception, & Psychophysics*, 77, 571–591. doi:10.3758/s13414-014-0791-3
- Singh, L., & Chee, M. (2016). Rise and fall: Effects of tone and intonation on spoken word recognition in early childhood. *Journal of Phonetics*, 55, 109–118. doi:10.1016/j.wocn.2015.12.005
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, 124, 128–142. doi:10.1016/j.cognition.2012.05.008
- Singh, L., Goh, H. H., & Wewalaarachchi, T. D. (2015). Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation. *Cognition*, 142, 1–11. doi:10.1016/j.cognition.2015.05.010
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3, 365–394. doi:10.1207/S15327078IN0303\_5
- Singh, L., Tam, H. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, 17, 94–109. doi:10.1111/desc.12097
- Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, 4, 157–178. doi:10.1080/15475440801922131
- Smith, A. (2006). Speech motor development: Integrating muscles, movements, and linguistic units. *Journal of Communication Disorders*, 39, 331–349. doi:10.1016/j.jcomdis.2006.06.017
- So, C. K. (2006). *Effect of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English* (Unpublished doctoral dissertation). Simon Fraser University, Burnaby, Canada.
- So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language*, 22, 473–495. doi:10.1017/S0305000900009922
- Thurgood, G. (2002). Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica*, 19, 333–363. doi:10.1075/dia.19.2.04thu
- To, C. K. S., Cheung, P. S. P., & McLeod, S. (2013). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech Language and Hearing Research*, 56, 103–122. doi:10.1044/1092-4388(2012/11-0080)
- Tong, X., McBride, C., & Burnham, D. (2014). Cues for lexical tone perception in children: Acoustic correlates and phonetic context effects. *Journal of Speech Language and Hearing Research*, 57, 1589–1605. doi:10.1044/2014\_JSLHR-S-13-0145
- Tsao, F.-M. (2008). The effect of acoustical similarity on lexical-tone perception of one-year-old Mandarin-learning infants. *Chinese Journal of Psychology*, 50, 111–124. Retrieved from <http://www.airitilibrary.com/Publication/allDetailedMesh?docid=10139656-200806-50-2-111-124-a>
- Tse, J. K.-P. (1978). Tone acquisition in Cantonese: A longitudinal case study. *Journal of Child Language*, 5, 191–204. doi:10.1017/S0305000900007418
- Tuaycharoen, P. (1977). *The phonetic and phonological development of a Thai baby: From early communicative interaction to speech* (Unpublished doctoral dissertation). School of Oriental and African Studies, University of London, London, UK.
- Wang, W. S. Y. (1967). Phonological features of tone. *International Journal of American Linguistics*, 33, 93–105. Retrieved from <http://www.jstor.org/stable/1263953>
- Wang, H., Li, A., & Fang, Q. (2005). F0 contour of prosodic word in happy speech of Mandarin. In J. Tao, T. Tan, & R. W. Picard (Eds.), *Affective computing and intelligent interaction* (pp. 433–440). Berlin, Germany: Springer. doi:10.1007/11573548\_56
- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113, 1033–1043. doi:10.1121/1.1531176
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, 106, 3649–3658. doi:10.1121/1.428217

- Werker, J. F. (2003). Baby steps to learning language. *Journal of Pediatrics*, 143, 62–69. doi:10.1067/S0022-3476(03)00403-7
- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349–355. doi:10.2307/1129249
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63. doi:10.1016/S0163-6383(84)80022-3
- Wewalaarachchi, D., & Singh, L. (2014). Influences of tone sandhi on word recognition in preschool children. *Proceedings of INTERSPEECH*, 2568–2571.
- Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25–47. doi:10.1159/000261901
- Wong, P. (2012a). Monosyllabic Mandarin tone productions by 3-year-olds growing up in Taiwan and in the United States: Interjudge reliability and perceptual results. *Journal of Speech, Language, and Hearing Research*, 55, 1423–1437. doi:10.1044/1092-4388(2012/11-0273)
- Wong, P. (2012b). Acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions. *Journal of Phonetics*, 40, 141–151. doi:10.1016/j.wocn.2011.10.005
- Wong, P. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *Journal of the Acoustical Society of America*, 133, 434–443. doi:10.1121/1.4768883
- Wong, P. C. M., & Diehl, R. L. (2003). Perceptual normalization for inter- and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research*, 46, 413–421. doi:10.1044/1092-4388(2003/034)
- Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28, 565–585. doi:10.1017/S0142716407070312
- Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language & Hearing Research*, 48, 1065–1079. doi:10.1044/1092-4388(2005/074)
- Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., . . . Han, D. (2010). Tone perception and production in pediatric cochlear implants users. *Acta Oto-Laryngologica*, 131, 395–398. doi:10.3109/00016489.2010.536993
- Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68, 123–139. doi:10.1016/j.jml.2012.09.004
- Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.
- Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *Journal of the Acoustical Society of America*, 130, 4063–4069. doi:10.1121/1.3651818
- Zhang, X., Samuel, A. G., & Liu, S. (2012). The perception and representation of segmental and prosodic Mandarin contrasts in native speakers of Cantonese. *Journal of Memory and Language*, 66, 438–457. doi:10.1016/j.jml.2011.12.006