A new approach to the learning of tonal categories in tone and non-tone languages
Aoju Chen
Institute for Language Sciences, Utrecht University
aoju.chen@uu.nl

In the autosegmental-metrical (AM) theory, a mainstream framework for intonational phonology, the prosodic system of a language encompasses a hierarchical structure, composed of both a finite number of pitch patterns residing in words which are grouped into prosodic phrases, and form-meaning mappings between changes in pitch patterns and changes in meaning [1, 2]. Recent research on prosodic development in children acquiring a non-tone language has adopted the AM view on prosody and examined development trajectories of pitch accents, phrasing and intonational form-meaning mappings, focusing on what children can do at which age. Similarly, research on children acquiring a tone language is mostly concerned with production of lexical tones and limited form-meaning mappings at different ages, albeit making no reference to the AM framework [3-5]. It has been found that children become increasingly attuned to native pitch patterns and less sensitive to non-native pitch patterns between 6 and 9 months and they have developed the inventory of lexical tones in a tone language and the inventory of pitch accents in a non-tone language at about 12 months. However, exactly HOW children acquire pitch accents in non-tone languages and lexical tones in tone languages remains to be investigated. In this position paper, I propose a new approach to address this question.

In this approach, adopting the AM framework, I assume that lexical tones have the same phonological status as pitch accents, i.e. being the discrete building blocks of the intonational pattern of an utterance, and refer to both as tonal categories. Furthermore, I take the view that although tone and non-tone languages differ in the density of tonal distribution (i.e. tone residing in every word vs. in some words in an utterance) and tonal function, across languages tonal categories can be distinguished in three dimensions, i.e. direction of pitch change (e.g. fall vs. rise), pitch height (e.g., high level vs. mid-level), and alignment of the highest or lowest pitch (e.g., early fall vs late fall) [6, 7]. But languages can differ in the weighting between dimensions [7, 8]. For example, the most important dimension in tonal perception is direction of pitch change in Mandarin but pitch height in English. Essential to acquiring native tonal categories is thus to find out which dimension(s) of pitch variation is/are relevant to formation of native tonal categories.

Hence, I propose to answer the HOW question by studying (1) innate attunement to pitch height and direction of pitch change; (2) the role of prenatal exposure in formation of tonal categories; (3) the role of distributional learning as a mechanism for the learning of tonal categories in the last trimester of gestation (26-39 weeks) and in the first months after birth; and (4) the role of visual cues, in particular, head and neck movements, in the learning of tonal categories at 4 to 12 months. In what follows, I briefly explain the rationale behind each research direction.

**Innate attunement:** The auditory system of human and other mammals is sensitive to variation in pitch, duration and intensity. For example, both humans and rats tend to group sequences of sounds in terms high-low in pitch and intensity and short-long in terms of duration, known as the Iambic-Trochaic law [9, 10]. The innate sensitivity to the prosodic parameters raises the exciting possibility that children may use it to uncover the relevance of pitch height and direction of pitch change for distinguishing tonal categories.

**Prenatal exposure:** Infants appear to possess some knowledge of the dimensions along which tonal categories differ in their native language already at 4-5 months, most probably the dimensions of pitch height and direction of pitch change [11-14]. Interestingly, infants’ sensitivity to pitch height, duration and intensity appears to already be molded into preliminary language-specific preferences at birth, presumably through prenatal exposure to speech [15]. This finding suggests that the learning of tonal categories may start in the third trimester of gestation when the fetus can hear and process low-frequency sounds with almost intact prosody transmitted through maternal abdomen [16, 17].

**Distributional learning:** Infants aged 6-8 months can learn non-native phonemes through distributional learning, i.e. a type of statistical learning that involves tracking distributional properties in the input. They interpret an acoustic parameter with bimodal distribution (i.e. the most frequent sounds are from the two ends of an acoustic continuum) as an indicator for the relevance of this acoustic parameter in categorising sounds, not an acoustic parameter with unimodal distribution (i.e. the most frequent sounds are from the middle of the continuum), and learn to discriminate novel sounds along the parameter with bimodal distribution only [18]. Distribution learning heightens 2-3 months olds’ perception of non-native phonemes [19] and appears to support infants’ learning of non-native tonal categories at 11-12 months [20]. The questions arise as to whether and how early distributional learning is available as a mechanism for the learning of tonal categories.

**Visual cue:** Infants usually not only hear but also see other people talking when interacting with them. Speech is typically accompanied by gestures; co-speech gestures are enhanced in infant-directed speech [21]. Visual information contributes to the learning of sounds in infancy if it contains sufficient category-related information [22, 23]. Visual cues facilitate the perception of lexical tones in adults unfamiliar with the tones [24, 25]. Lexical tones differ perceivably in head and neck movements when produced in isolation [26, 27]. For example, in Mandarin, dropping of head can signal the low tone. Visual cues may thus potentially support infants in working out the relevant prosodic dimension in which tonal categories differ in the auditory modality.

To sum up, I have outlined a new approach to the study of acquisition of tonal categories, bridging the divide in past research on children’s production and perception of lexical tones in tone languages and pitch accents in non-tone languages and tackling the question of how children come to produce and perceive tonal categories in their prosodic system. In the talk, I will also suggest ideas for how to test hypotheses arising from this approach.
References


