



Universiteit
Leiden
The Netherlands

Versatility of phonemic pitch in affective iconicity and perceptual reorganisation

Zheng, T.

Citation

Zheng, T. (2025, November 19). *Versatility of phonemic pitch in affective iconicity and perceptual reorganisation*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/4283265>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4283265>

Note: To cite this publication please use the final published version (if applicable).

Chapter 6

General discussion

This dissertation investigates the multifaceted roles of phonemic pitch in two distinct contexts: (1) the affective iconicity of linguistic pitch in Standard Chinese (SC) tones and (2) the developmental hemispheric lateralisation of linguistic pitch during perceptual reorganisation in Dutch- and Japanese-learning infants. In the first context, the study explores: (**R1**) whether the linguistic pitch of SC lexical tones biases emotional responses (arousal and valence) in adult speakers, (**R2**) how segmental phonemes influences emotional responses, (**R3**) whether pitch dominates segmental phonemic cues in impacting emotional responses, and (**R4**) the potential mechanisms through which lexical pitch affects emotional responses. In the second context, the study examines cross-linguistic differences in infants' brain responses to lexical pitch, comparing learners of a pitch accent language (Japanese) and a lexical stress (Dutch) language. Specifically, it addresses research questions (**R5**) whether Dutch-learning infants can distinguish lexical pitch contrasts, and (**R6**) how Dutch- and Japanese-learning infants differ in their ability to discriminate lexical pitch contrasts.

By addressing these research questions, this work provides insights into the relationships between linguistic pitch, emotional responses, language experience, brain development, and hemispheric lateralisation. The following two sections summarise each chapter, highlighting key findings, contributions, limitations, and directions for future research.

6.1 Research questions and findings

Chapter 2 explored the potential pitch iconic effects of lexical tones on affective interpretations in bisyllabic words in Standard Chinese (**R1**). Our findings provide compelling evidence that affective iconicity in Standard Chinese (SC) lexical tones systematically influences both

the general lexicon (as analysed across three corpora) and spoken nonce words. Consistent with our hypotheses, the FF tonal sequence, characterised by extensive pitch variation in range and slope, was consistently associated with higher arousal compared to RR across both written words (Studies 1 and 2) and spoken nonce words (Study 3). In spoken nonce words, HH was more likely to be rated as low arousal compared to the other three tonal sequences: RR, RL (LL), and FF. Additionally, a significant influence of tonal sequences on valence was observed exclusively in nonce words. Specifically, FF and RL tonal sequences were more likely to be rated as having negative valence, whereas RR and HH sequences were associated with positive valence.

The pitch iconic effects of lexical tones on affective interpretations in bisyllabic words in SC mirror the intonational pitch variations in emotional expressions, suggesting a universal tendency for pitch cues to shape emotional communication. Specifically, higher overall pitch level, higher average pitch height, wider pitch range, and steeper pitch slope are expected to be associated with higher arousal (Bänziger & Scherer, 2005; Scherer et al., 2003; Thompson, 2018). Furthermore, higher pitch level, wider pitch range, and upward pitch contour direction are likely associated with positive valence (Belyk & Brown, 2014; Kamiloğlu et al., 2020; Yap et al., 2014).

Chapter 3 examined how lexical tones, vowels, consonants, and their interactions contribute to affective iconicity. We first replicated widely observed phonemes-emotional meaning correspondences (**R2**). For vowels, /u/ was associated with negative valence and /i/ with positive valence across both vowel-only and CV(L) monosyllabic syllables, consistent with previous studies. Regarding consonants, /t/ was more likely to be rated as high arousal, while /n/ was associated with low arousal. At the monosyllabic level, lexical tones also showed

systematic effects: T2 and T4 were associated with higher arousal compared to T1 and T3, while T1 and T2 were more likely to be associated with positive valence compared to T3 and T4. In line with our hypothesis, lexical tones emerged as stronger predictors of emotional arousal and valence compared to phonemes. Specifically, when lexical tones and phonemes predicted opposing directions for valence and arousal, lexical tones were more decisive. The results revealed significant associations between tonemes and both arousal and valence, consonants and arousal, and vowels and valence, supporting the idea that native Standard Chinese speakers integrate both phonemes and tonemes when interpreting the affective meanings of auditory stimuli. However, no significant consonant-valence or vowel-arousal correspondences were observed in this study.

The predominant effect of lexical pitch over phonemes in affective interpretations lends further support to the notion that lexical-prosodic features facilitate the interpretation of iconic meanings (Dingemanse et al., 2016). This finding aligns with previous research demonstrating the dominance of emotional prosody over verbal content in emotion recognition (Filippi et al., 2017; Lin et al., 2020, 2021). The shared phonetic feature between lexical tones and emotional prosody, pitch, a highly salient perceptual dimension, suggests that these systems may rely on shared or transferable mechanisms in emotional processing. Specifically, the prominence of pitch in both lexical tones and emotional prosody highlights its central role in conveying affective meaning, whether through linguistic or paralinguistic channels.

Chapter 4 investigated the adaptive significance of lexical tones in Standard Chinese in affective iconicity. Using the three corpus datasets from Chapter 2, we found that individual tones significantly

influenced the arousal and valence ratings of bisyllabic Standard Chinese words. Both the lexical tones of the initial and second syllables predicted the arousal ratings of words. Specifically, high-arousing words were characterised by a falling tone (T4) in both syllables, while low-arousing words were more likely to be characterised by a rising (T2) tone or a low-dipping tone (T3). In contrast, the valence ratings of the bisyllabic words were predominantly influenced by the lexical tone of the first syllable. Negative words were more likely to initiate with a falling tone (T4), while positive words tended to start with a rising tone (T2). These findings align with our hypotheses and suggest a potential mechanism of affective iconicity due to adaptive significance (Adelman et al., 2018). Specifically, lexical tones with falling contours are associated with negative, high-arousing words, signalling the urgency and importance of the stimulus or information, thereby facilitating communication efficiency. Such an interpretation of the lexical tone effects on emotional expression lends further support to the adaptive significance proposal by Adelman et al. (2018), which is based on findings from segmental phonemes.

Chapter 5 explored the developmental trajectory of pitch processing in Dutch and Japanese infants, revealing cross-linguistic similarities and differences in hemispheric lateralisation influenced by the acoustic properties and linguistic relevance of pitch. Across three experiments, we observed that Dutch infants at both 4 and 10 months could distinguish pitch changes within speech stimuli, though with some individual differences in the consistency of their discrimination (Experiment 1). In Experiments 2 and 3, Dutch infants showed right-lateralised responses to both pure tones and pseudowords at 4 months. By 10 months, their responses to pure tones remained right-lateralised, but for pseudowords, they shifted to a bilateral pattern. Japanese

infants exhibited bilateral responses to pure tones at 4 months, but by 10 months, their responses became right-lateralised. For pseudowords, they exhibited left-lateralised responses at both ages, with stronger activation at 10 months than at 4 months. These findings align with our hypotheses, highlighting the influence of both acoustic and linguistic factors in early pitch processing. However, the left-lateralised response to pseudowords in 4-month-old Japanese infants may indicate an early developmental sensitivity to linguistically relevant pitch contrasts. This difference between Dutch and Japanese infants, particularly in the older group, reflects how pitch processing is shaped by its linguistic function, emphasising the influence of language experience.

6.2 Contributions, limitations, and future directions

This dissertation provides new insights into the affective and developmental implications of linguistic (phonemic) pitch, revealing its multifaceted roles across acoustic, cognitive, and linguistic domains. The findings emphasise that pitch processing is not solely determined by acoustic properties but also shaped by cognitive processing, linguistic experiences, and relevance.

6.2.1 Phonemic pitch in affective iconicity

Our exploration of affective iconicity highlights how iconicity adapts to specific linguistic systems, particularly Standard Chinese (SC), and its universal role in human communication.

Lexical tones and emotional responses

We demonstrated that lexical tones significantly shape emotional responses, including arousal and valence. Lexical tones at both the monosyllabic and bisyllabic levels, regardless of lexical meaning,

predict emotional arousal. Valence predictions, however, were observed primarily in bisyllabic nonce words. There are a few possibilities explaining this difference.

One is that iconicity itself is pre-semantic. The observation that both tone–arousal and tone–valence associations occur in nonce words supports the notion that iconicity is a pre-semantic phenomenon at the acoustic level (rather than relying on semantic content), which is possibly a foundational mechanism in early human communication systems (Sučević et al., 2015; Westbury, 2005). These findings suggest that iconicity operates at the acoustic level, independent of semantic content, with pitch features such as overall level, average height, range, contour direction, and slope driving emotional associations. For instance, the falling contour, large pitch range, and steep slope of falling tone (T4) may explain its association with high arousal and negative valence, while the nearly static pitch and high level of high-level tone (T1) align with its low-arousal, positive valence tendency. Similarly, the rising contour and moderate pitch range of the rising tone (T2) appear linked to positive valence. Notably, bisyllabic falling-falling (T4T4) tonal sequences evoked higher arousal than rising-rising (T2T2) tonal sequences, despite there being no significant difference in the monosyllabic context. This discrepancy suggests that pitch dynamics, syllable structure, and tonal articulation might interact, warranting further exploration with these considerations. In addition, our stimuli in the corpora and nonce words are limited; future research may consider larger datasets and diverse phonemes to thoroughly investigate tone–arousal and tone–valence association.

Another possibility is that while both associations may be pre-semantic, arousal is more pronounced than valence for iconicity at the

level of lexical tone, given their different effect sizes in Chapter 2. This aligns with findings by Aryani et al. (2018), which demonstrated that phonemic segments more strongly influence arousal than valence. As two key emotional dimensions, arousal is a more ancient and universal dimension tied to physiological responses, whereas valence is a more abstract dimension influenced by cultural and semantic contexts (Darwin, 1998; Russell, 2003). The differences between the two dimensions might lead to varying degrees of iconic associations with tone. For example, tone–valence association is easily shadowed by lexical meanings and conventions. Further replication and exploration are necessary to test these possibilities. Cross-linguistic studies could further investigate whether naïve speakers with no knowledge of SC would perceive affective iconicity in tonal sequences. Additionally, experiments with homophonic SC bisyllabic words, where tones vary but lexical meanings remain constant, might clarify whether lexical tone alone influences valence ratings.

Affective iconicity and the adaptation hypothesis

Our findings revealed distinct patterns of tone–arousal and tone–valence associations and supported the adaptation account of iconicity. We observed positional effects of lexical tones on emotional arousal and valence ratings: the first-syllable tone of bisyllabic words predicted valence, while both syllables' tones predicted arousal. This positional effect of lexical tones on emotional valence rating replicates findings from Indo-European languages, where the initial phoneme often predicts valence, and extends the adaptation hypothesis to include tonemes.

These results align with the distinctions for arousal and valence (Citron, 2012; Darwin, 1998) and their neural processing mechanisms

(Fischler & Bradley, 2006; Herbert et al., 2008; Kissler et al., 2007, 2009; Schacht & Sommer, 2009). Moreover, Aryani et al. (2019) revealed a potential neural mechanism underlying affective iconicity in an fMRI study. They found that affective iconic words, compared to their non-iconic counterparts, elicited additional brain activity in the left amygdala, a region associated with the multimodal representation of emotions. This effect was modulated by sound processing in the left superior temporal gyrus and language processing in the left inferior frontal gyrus.

Future research, particularly neuroimaging studies, could further investigate the neural pathways involved in pitch iconicity and their interactions with emotional and linguistic processing regions. Examining the mechanisms underlying pitch iconicity may offer new insights into the neurocognitive foundations of tonal and affective processing. Specifically, tones may engage distinct neural pathways that integrate both linguistic and emotional processing, potentially involving interactions between auditory and affective brain regions. Additionally, exploring whether speakers of tonal and non-tonal languages process pitch iconicity differently could shed light on the brain's plasticity in adapting to diverse linguistic environments. Such findings would have broader implications for understanding how language experience shapes neural responses to emotionally salient speech cues.

Lexical tones dominate over phonemes in emotional predictions

Lexical tones demonstrated a stronger influence than phonemes in predicting emotional responses. While phoneme-based associations such as “/i/–positive,” “/u/–negative,” “/t/–high arousal,” and “/n/–

low arousal” were replicated, tonal predictions often overrode phonemic ones, particularly in cases of conflict (e.g., /i4/ versus /u1/ were rated as negative versus positive). Lexical-prosodic information has been considered a facilitator in interpreting iconic meanings across languages (Dingemanse et al., 2016; Perlman, Clark, et al., 2015; Stel et al., 2012; Thompson, 2018). Our findings further underscore the prioritisation of suprasegmental features over segmental ones in affective interpretation, emphasising the critical role of prosody in tonal languages.

This dominance of lexical tones suggests that tonemes may function as a dual-purpose tool for conveying both lexical and emotional meaning. Investigating the relative contributions of segmental versus suprasegmental features in tonal and non-tonal languages could provide deeper insights into how linguistic systems harness pitch for communication.

As discussed in the *Lexical tones and emotional responses* section, we propose that iconicity operates at the acoustic level, possibly independent of semantic content. This suggests that pitch, beyond its primary role in distinguishing lexical meaning in tonal languages, retains a residual capacity to convey emotion. Even with this residual capacity, pitch can exert a stronger influence than phonemes in emotional expressions, highlighting its salience in conveying affective meanings. However, further studies replicating this phenomenon across diverse language users and dialects are needed to provide more robust evidence.

Pitch iconicity and the frequency code hypothesis

Some of our findings on pitch iconicity in SC lexical tones provide further support for the frequency code hypothesis, which posits

universal principles of pitch-iconicity based on acoustic properties (Gussenhoven, 2016; Hinton et al., 2006; Ohala, 1984). According to this hypothesis, higher and/or rising *f₀* is associated with smallness, submission, and politeness, whereas lower and/or falling *f₀* conveys largeness, dominance, and aggression.

Particularly, we assume that the *f₀* difference between T2/T4 and T3 (high versus low-dipping) likely contributed to their arousal contrast, and differences in *f₀* between T1/T2 and T3/T4 (high/rising versus low-dipping/falling) may explain their valence choice differences. However, it is important to note that these arousal and valence differences among lexical tones are more likely influenced by the tones' pitch characteristics, such as pitch contour directions, pitch ranges, and pitch slope.

It is important to bear in mind that while the frequency code may contribute to the sound-symbolic link between pitch and affective connotations, the relationship is not entirely straightforward, as shown in studies such as Morton (1977), Perlman (2024), and Winter et al. (2021). Future research with more nuanced experimental designs is needed to clarify how specific pitch features function in affective iconicity and how these functions relate to the frequency code hypothesis.

Broader implications and future directions

The investigation of tonemes as affective carriers enriches our understanding of iconicity by extending it from segmental to suprasegmental features. Cross-linguistic studies are crucial to determining whether the affective iconicity of tonal sequences, rooted in pitch variations, is universal. Additionally, research on the interaction of tone and intonation could shed light on how tonal

languages integrate prosodic and lexical functions to convey meaning efficiently. Affective iconicity may also play a role in reducing cognitive load by allowing listeners to infer emotional meaning directly from acoustic cues. This effect could be particularly relevant in tonal languages, where pitch is a dominant linguistic feature. Furthermore, iconicity has been shown to facilitate word learning, which may be especially beneficial for children and second-language learners (e.g., Asano et al., 2015; Imai et al., 2008, 2015; Lockwood et al., 2016; Nielsen & Dingemanse, 2021). By advancing the study of pitch iconicity, this research not only deepens our understanding of tonal languages but also contributes to broader theories of language acquisition, sound symbolism, and emotional communication. Future work should explore how pitch-based iconicity interacts with other linguistic and cognitive processes, shedding further light on the fundamental links between speech, emotion, and meaning.

6.2.2 Phonemic pitch in perceptual reorganisation

Investigating perceptual reorganisation and its associated hemispheric lateralisation in infants is crucial for understanding the developmental trajectory of language acquisition and the neural mechanisms underlying this process. Between 6 and 12 months, infants transition from being universal listeners, capable of discriminating most phonemic contrasts, to becoming specialists attuned to their native language(s). This perceptual reorganisation process forms the foundation for acquiring phonology. Understanding how and why infants lose sensitivity to nonnative contrasts provides critical insights into the timing and mechanisms of language learning and acquisition.

Hemispheric lateralisation and perceptual reorganisation of pitch

Across the lifespan, pitch remains an indispensable feature of auditory perception, serving multiple functions in linguistic processing, emotional and social communication, and musical perception. The early shift from broad pitch sensitivity to language-specific specialisation reflects the brain's ability to adapt to environmental input, a hallmark of perceptual reorganisation (Gervain, 2020; Werker, 2018).

Our findings provide further evidence on the role of pitch in hemispheric lateralisation by comparing Japanese and Dutch infants. While the acoustic properties of pitch remain constant across languages, its linguistic functions differ, serving as a phonemic contrast in Japanese but not in Dutch. This cross-linguistic comparison sheds light on how perceptual properties interact with language experience and linguistic relevance or function to shape phonemic pitch processing. Perceptual reorganisation reflects experience-dependent plasticity, wherein the brain selectively enhances sensitivity to relevant phonetic cues while deprioritising those that are not functionally meaningful in the infant's linguistic environment (Minagawa-Kawai et al., 2011).

At 10 months, Japanese and Dutch infants demonstrated diverging neural responses to pitch-based speech stimuli, supporting the hypothesis that pitch processing pathways are shaped by the linguistic relevance of acoustic features. Specifically, pitch cues deemed "language-relevant" appear to be routed from the primary auditory cortex to higher-level processing areas in the left temporal cortex, whereas general acoustic properties are predominantly processed in

the right temporal cortex (Aslin, 2012). This functional specialisation underscores the role of language experience in shaping developmental trajectories. Our findings suggest that Japanese infants develop a phonologically oriented sensitivity to pitch, reflecting the importance of pitch-accent distinctions in their native language, while Dutch infants show no such specialisation, treating pitch as a general auditory feature.

We found an early emergence of left-lateralised processing for linguistic pitch in 4-month-old Japanese infants, aligning with previous studies (e.g., Ren et al., 2024; Telkemeyer et al., 2011). This early lateralisation may reflect a developmental readiness to acquire pitch contrasts in a linguistically relevant way. Moreover, by 10 months, left-hemisphere activation became stronger, suggesting increasing specialisation for processing pitch contrasts relevant to Japanese. This developmental shift likely results from both accumulating linguistic experience and the critical role of pitch in Japanese. These findings suggest a hierarchical developmental trajectory between phonological and lexical processing. One possibility is that phonological pitch processing emerges early, providing a foundation for later lexical development. Alternatively, phonological pitch processing may gradually strengthen over time, eventually integrating with lexical-level processing. Further systematic investigations across different language backgrounds are needed to clarify these developmental mechanisms.

The hemispheric response to pitch in non-speech stimuli primarily reflected processing based on acoustic properties (Aslin, 2012), with slower pitch variations predominantly processed in the right hemisphere. However, Dutch infants exhibited a shift from right-lateralised (at 4 months) to bilateral responses (at 10 months) for

speech stimuli. The right-hemispheric dominance at 4 months likely reflects an early-stage reliance on acoustic properties for pitch processing. By 10 months, the shift to bilateral activation may indicate an emerging sensitivity to prosodic structure, aligning with adult-like bilateral pitch processing to linguistic prosody in Dutch (e.g., Witteman et al., 2011, 2014). This further supports the idea that infants' neural specialisation for pitch develops in response to their native language environment.

Moreover, previous studies have shown increased involvement of the frontal region in tone categorisation as infants mature (Ren et al., 2024), and stronger connectivity between the frontal and temporal regions following exposure to speech stimuli (Homae et al., 2011). These findings suggest that other top-down modulations, shaped by prior knowledge, attention, and/or executive functions, also play a role in pitch processing. Future research should incorporate measures of both frontal and temporal lobe activity to better understand their interaction in hemispheric specialisation and perceptual reorganisation.

In summary, the dual acoustic-functional properties of lexical pitch accent provide valuable insight into the developmental trajectory of perceptual reorganisation and the corresponding hemispheric lateralisation of auditory processing. By comparing Dutch and Japanese infants, we explored whether early pitch processing is influenced by language-specific phonological categories or remains a general auditory skill at this developmental stage. The observed similarities and differences highlight the role of linguistic, experiential, and cognitive contexts in shaping early auditory processing pathways. Specifically, Japanese infants may develop a phonologically-oriented sensitivity to pitch, while infants in Dutch show no such specialisation.

Methodological considerations and future directions

From a methodological perspective, fNIRS remains one of the most effective non-invasive neuroimaging techniques for studying infant brain responses due to its good temporal and spectral resolution, suitability for testing awake infants (Wilcox & Biondi, 2015). The adoption of fNIRS has facilitated the investigation of pitch processing variations across language groups and auditory contexts. Although there is no consensus on fNIRS preprocessing practices, we employed well-justified pipelines and further refined them by applying different parameters critical for haemoglobin signal extraction. However, key preprocessing steps, such as motion correction and baseline correction, require further refinement to improve consistency and comparability across studies. Standardising these techniques will enhance the reliability of fNIRS research on infant speech perception and neural specialisation.

Current studies in Dutch and Japanese infants suggest that multiple factors influence pitch processing, leading to variations in hemispheric lateralisation. Future research could explore more salient pitch contrasts, such as the distinction between T1 and T4 in Standard Chinese, to better understand how infants from different language backgrounds process pitch variations during perceptual reorganisation. Additionally, investigating how pitch saliency and relevance vary across language contexts, shaped by both pitch function and cognitive processing, could provide deeper insights into early pitch perception mechanisms.

6.3 Conclusion

This dissertation highlights two interrelated roles of pitch processing: affective iconicity in Standard Chinese (SC) speakers and perceptual reorganisation in Japanese (JP) and Dutch (NL) infants. Despite their differences, both cases illustrate how pitch perception emerges from the interplay between acoustic properties and linguistic experience.

In the affective iconicity context, SC speakers systematically associate pitch with emotional valence and arousal, demonstrating that pitch perception extends beyond linguistic functions to iconic sound-meaning mappings. Meanwhile, in the developmental context, JP and NL infants process phonemic pitch differently, with JP infants showing greater left-hemisphere involvement, which reflects the linguistic relevance of pitch in their native languages.

These findings underscore pitch as a dynamic cue shaped by experience and communicative function, whether as a lexical contrast (in JP and NL) or an emotional signal (in SC). By bridging insights from affective iconicity and hemispheric lateralisation, this work deepens our understanding of how the brain adapts to the auditory environment, refining pitch processing based on both linguistic experience and cognitive demands.