



How Do Infants Disaggregate Referential and Affective Pitch?

René Kager*

Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, Netherlands

Infants are faced with a challenge of disaggregating functions of pitch in the ambient language into affective, pragmatic or referential (the latter in tone languages only). This mini review discusses several factors that might facilitate the disaggregation of referential and affective pitch in infancy: acoustic characteristics of infant-directed speech, recognition of vocal affect, facial cues accompanying affective prosody, and lateralization of affective and referential prosody in the brain. It proposes two hypotheses concerning the role of audiovisual cues and brain lateralization

Keywords: infant speech perception, pitch processing, infant language representation, lexical tone acquisition¹, Lexical tone perception

This article discusses potential factors to facilitate the disaggregation of referential and affective pitch in infancy: acoustic characteristics of infant-directed speech, recognition of vocal affect, facial cues accompanying affective prosody, and lateralization of affective and referential prosody in the brain. It proposes two hypotheses concerning the role of audiovisual cues and brain lateralization.

Among the many acoustic cues in speech, fundamental frequency (perceived as pitch) is arguably the one that, cross-linguistically, has the widest range of linguistic and para-linguistic uses (Gussenhoven, 2004). Universally, pitch signals *affective* use (for example, express happiness by high average pitch and wide pitch range) and *pragmatic* use (for example, marking a question by rising pitch is a universal tendency). Exclusively in tone languages, pitch supports *referential* use by contrasting word meanings (for example, Cantonese /fan/“divide” carries a high-level tone; “angry” a mid-rising tone). Infants born into tone languages (a term which includes “pitch accent” languages; Hyman, 2009) are faced with a challenge of discovering how pitch patterns in the ambient language distinguish different word meanings—hence, they must *disaggregate* pitch in the input into non-referential and potentially referential information. Infants learning a (non-tone) lexical stress language must discover that pitch has no direct, but only indirect referential significance as one of the cues associated with stress (next to other cues, e.g., duration). Detection of the referential significance of pitch poses a critical challenge for infants when they are learning their first words. Yet several studies suggest infants discover the presence/absence of lexical tones before their first birthday. Tone-learning infants retain their initial ability to discriminate tones, while infants exposed to a non-tone language lose it between 6 and 9 months (Mattock and Burnham, 2006; Mattock et al., 2008; Yeung et al., 2013; Liu and Kager, 2014; Götz et al., 2018), before losing the ability to learn tone-to-word associations, which they still possess at 9 months (Yeung et al., 2014), by 18 months (Singh et al., 2014; Hay et al., 2015; Burnham et al., 2018; Liu and Kager, 2018). How are infants able to disaggregate pitch into non-referential affective and referential linguistic information?

Infants’ environments are rich in affective content, as infant-directed speech (IDS) is characterized by exaggerated pitch contours reflecting “free vocal expression of emotion” (Trainor et al., 2000), which attracts infants’ attention (Cooper and Aslin, 1990; Werker et al., 1994), yet does not a priori facilitate tone acquisition, as it may partially obscure contrastive shapes of tones (Papoušek and Hwang, 1991; Kitamura et al., 2002). Pitch exaggeration in IDS may be partly compensated by tonal hyper-articulation (Liu et al., 2007; Xu Rattanasone et al., 2013; Tang et al., 2017), yet to what extent precisely is an open

OPEN ACCESS

Edited by:

Leher Singh,
National University of Singapore,
Singapore

Reviewed by:

Carolyn Quam,
Portland State University,
United States
Feng-Ming Tsao,
National Taiwan University, Taiwan

*Correspondence:

René Kager
r.w.j.kager@uu.nl

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 04 May 2018

Accepted: 10 October 2018

Published: 31 October 2018

Citation:

Kager R (2018) How Do Infants
Disaggregate Referential and Affective
Pitch? *Front. Psychol.* 9:2093.
doi: 10.3389/fpsyg.2018.02093

issue. In order to facilitate disaggregation of referential and affective pitch, young infants may draw on their ability to recognize vocal and visual expression of affect.

The ability to interpret speech prosody as having affective value emerges early in life. Pitch contours in IDS presumably carry innately specified affective meanings to young infants, eliciting attention, arousal, approval, and disapproval (see Fernald, 1992, for a review). Neonates show an increase in eye opening responses to happy vocal stimuli as compared to other expressions (angry, sad, neutral), however only for their native language (Mastropieri and Turkewitz, 1999), suggesting prenatal influence on perception of vocal affect. By 5 months, infants reliably discriminate affect, detecting changes in vocal affect from sad to happy (Walker-Andrews and Grolnick, 1983); 7-month-olds show different ERP responses to affective (happy or angry) vs. neutral prosody (Grossmann et al., 2005). Yet infants' ability to *discriminate* affect may not provide a reliable basis for affective-referential pitch disaggregation; perhaps it should be matched by an ability to *understand* emotion in speech. However, this ability is not developed until 4–5 years (Quam and Swingley, 2012). School-aged children (around age 10) experience difficulties integrating vocal affect with lexical content (Friend, 2000, 2001, 2003; Friend and Bryant, 2000; Morton and Trehub, 2001; Morton et al., 2003). Since the ability to understand emotion in speech develops so slowly, it is worth exploring how affective-referential pitch disaggregation during the first year of life might be supported not only by auditory/vocal cues, but also by visual/facial cues.

By 4–6 months of age, infants in spite of their reduced visual processing can discriminate their native language from other languages partly by relying on visual cues accompanying gestures such as vocalic lip rounding (Weikum et al., 2007). In comparison, visual cues to tonal gestures are weak and unreliable to native listeners (Chen and Massaro, 2008; Hannah et al., 2017). Young infants (4-month-olds) can detect different emotions (happy, angry, sad) when presented with facial-vocal cues (Flom and Bahrick, 2007), an ability emerging prior to affect detection based on unimodal cues (Walker-Andrews, 1997). In light of infants' early sensitivity to facial-vocal cues to affect, the hypothesis can be proposed that affective-referential pitch disaggregation draws on facial affective cues accompanying vocal affect. By labeling pitch information as affective, infants may focus their linguistic attention to residual pitch information that has no clear affective interpretation, which includes referential information.

A neurological marker of affective-referential pitch disaggregation may be obtained in the hemispheric specialization for linguistic and affective pitch. A functional asymmetry between the right hemisphere (RH; dominant in processing pitch changes and emotional vocalization) and the left hemisphere (LH; dominant in processing speech, in

particular segmental information) occurs in neonates (Dehaene-Lambertz, 2000; Peña et al., 2003). Native listeners process linguistically relevant lexical pitch dominantly in LH (Wang et al., 2001); affective pitch dominantly in RH (Edmondson et al., 1987). Yet hemispheric lateralization of linguistic and affective pitch processing remains a controversial issue (Wong, 2002; Zatorre and Gandour, 2008). Turning to infant studies, early RH specialization for pitch processing is found in neonates (Arimitsu et al., 2011); 3-month-old Japanese infants show stronger RH responses to natural speech, which includes pitch contours, as compared to prosodically flattened speech (Homae et al., 2006). The processing of lexical pitch is lateralized to LH in Japanese infants between 4 and 10 months (Sato et al., 2010; see Minagawa-Kawai et al., 2011 for discussion). Plausibly, the disaggregation of affective-referential pitch involves a functional specialization of the brain's hemispheres: general (affective and linguistic) pitch processing starts out in RH, while disaggregation amounts to a lateralization of linguistic pitch processing to LH. Infants' detection of affect, guided by vocal-facial cues, provides the key ability. A second hypothesis is proposed to this effect: the more emotional speech is, the more dominant RH becomes in speech processing; conversely, less emotional speech implies a decreased role for RH in pitch processing, enabling a partial shift of pitch processing to LH, the dominant hemisphere for speech processing. This predicts that (the perceived amount of) facial affect influences the locus of pitch processing in the infant brain.

In sum, affective-referential pitch disaggregation by infants may be accomplished by a combination of two (possibly innate) abilities, matching the two hypotheses stated above: (a) recognition of affect in pitch contours and integration of audiovisual (vocal-facial) cues on affect; (b) hemispheric specialization for pitch processing, where RH acts as “emotion attractor” and LH as “language attractor.” Integrating research on early tone perception, audiovisual affect recognition, and hemispheric specialization may open a new perspective on how infants manage to detect the presence/absence of lexical tone in their native language.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The Consortium on Individual Development (CID) is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO Grant No. 024.001.003).

REFERENCES

- Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., et al. (2011). Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Front. Psychol.* 2:202. doi: 10.3389/fpsyg.2011.00202
- Burnham, D., Singh, L., Mattock, K., Woo, P. J., and Kalashnikova, M. (2018). Constraints on tone sensitivity in novel word learning by monolingual and bilingual infants: tone properties are more influential than tone familiarity. *Front. Psychol.* 8:2190. doi: 10.3389/fpsyg.2017.02190
- Chen, T. H., and Massaro, D. W. (2008). Seeing pitch: visual information for lexical tones of Mandarin-Chinese.

- J. Acoust. Soc. Am.* 123, 2356–2366. doi: 10.1121/1.2839004
- Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595. doi: 10.2307/1130766
- Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *J. Cogn. Neurosci.* 12, 449–460. doi: 10.1162/089892900562264
- Edmondson, J. A., Chan, J. L., Seibert, G. B., and Ross, E. D. (1987). The effect of right-brain damage on acoustical measures of affective prosody in Taiwanese patients. *J. Phon.* 15, 219–233.
- Fernald, A. (1992). “Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective,” in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds J. H. Barkow, L. Cosmedes, and J. Tooby (New York, NY: Oxford: Oxford University Press), 391–428.
- Flom, R., and Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: the role of intersensory redundancy. *Dev. Psychol.* 43:238. doi: 10.1037/0012-1649.43.1.238
- Friend, M. (2000). Developmental changes in sensitivity to vocal paralinguistic. *Dev. Sci.* 3, 148–162. doi: 10.1111/1467-7687.00108
- Friend, M. (2001). The transition from affective to linguistic meaning. *First Lang.* 21, 219–243. doi: 10.1177/014272370102106302
- Friend, M. (2003). What should I do? Behavior regulation by language and paralinguistic in early childhood. *J. Cogn. Dev.* 4, 161–183. doi: 10.1207/S15327647JCD0402_02
- Friend, M., and Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill Palmer Q.* 46, 342–369.
- Götz, A., Yeung, H. H., Krasotkina, A., Schwarzer, G., and Höhle, B. (2018). Perceptual reorganization of lexical tones: effects of age and experimental procedure. *Front. Psychol.* 9:477. doi: 10.3389/fpsyg.2018.00477
- Grossmann, T., Striano, T., and Friederici, A. D. (2005). Infants’ electric brain responses to emotional prosody. *Neuroreport* 16, 1825–1828. doi: 10.1097/01.wnr.0000185964.34336.b1
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Hannah, B., Wang, Y., Jongman, A., Sereno, J. A., Cao, J., and Nie, Y. (2017). Cross-modal association between auditory and visuospatial information in Mandarin tone perception in noise by native and non-native perceivers. *Front. Psychol.* 8:2051. doi: 10.3389/fpsyg.2017.02051
- Hay, J. F., Graf Estes, K., Wang, T., and Saffran, J. R. (2015). From flexibility to constraint: the contrastive use of lexical tone in early word learning. *Child Dev.* 86, 10–22. doi: 10.1111/cdev.12269
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280. doi: 10.1016/j.neures.2005.12.006
- Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Lang. Sci.* 31, 213–238. doi: 10.1016/j.langsci.2008.12.007
- Kitamura, C., Thanavistuth, C., Burnham, D., and Luksaneeyanawin, S. (2002). Universal pitch modifications in infant directed speech: a prelinguistic longitudinal study in a tonal and non-tonal language. *Infant Behav. Dev.* 24, 372–392. doi: 10.1016/S0163-6383(02)00086-3
- Liu, H. M., Tsao, F. M., and Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Dev. Psychol.* 43:912. doi: 10.1037/0012-1649.43.4.912
- Liu, L., and 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, L., and Kager, R. (2018). Monolingual and bilingual infants’ ability to use non-native tone for word learning deteriorates by the second year after birth. *Front. Psychol.* 9:117. doi: 10.3389/fpsyg.2018.00117
- Mastropieri, D., and Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Dev. Psychobiol.* 35, 204–214.
- Mattock, K., and 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., and 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
- Minagawa-Kawai, Y., Cristià, A., and Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: a developmental scenario. *Dev. Cogn. Neurosci.* 1, 217–232. doi: 10.1016/j.dcn.2011.03.005
- Morton, J. B., and Trehub, S. E. (2001). Children’s understanding of emotion in speech. *Child Dev.* 72, 834–843. doi: 10.1111/1467-8624.00318
- Morton, J. B., Trehub, S. E., and Zelazo, P. D. (2003). Sources of inflexibility in 6-year-olds’ understanding of emotion in speech. *Child Dev.* 74, 1857–1868. doi: 10.1046/j.1467-8624.2003.00642.x
- Papoušek, M., and Hwang, S. F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: comparison with registers of adult conversation and foreign language instruction. *Appl. Psycholinguist.* 12, 481–504. doi: 10.1017/S0142716400005889
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F. et al. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705. doi: 10.1073/pnas.1934290100
- Quam, C., and Swingle, D. (2012). Development in children’s interpretation of pitch cues to emotions. *Child Dev.* 83, 236–250. doi: 10.1111/j.1467-8624.2011.01700.x
- Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513. doi: 10.1162/jocn.2009.21377
- Singh, L., Hui, T. J., Chan, C., and Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Dev. Sci.* 17, 94–109. doi: 10.1111/desc.12097
- Tang, P., Xu Rattanasone, N., Yuen, I., and Demuth, K. (2017). Phonetic enhancement of Mandarin vowels and tones: infant-directed speech and Lombard speech. *J. Acoust. Soc. Am.* 142, 493–503. doi: 10.1121/1.4995998
- Trainor, L. J., Austin, C. M., and Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11, 188–195. doi: 10.1111/1467-9280.00240
- Walker-Andrews, A. S. (1997). Infants’ perception of expressive behaviors: differentiation of multimodal information. *Psychol. Bull.* 121, 437–456.
- Walker-Andrews, A. S., and Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behav. Dev.* 6, 491–498. doi: 10.1016/S0163-6383(83)90331-4
- Wang, Y., Jongman, A., and Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain Lang.* 78, 332–348. doi: 10.1006/brln.2001.2474
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., and Werker, J. F. (2007). Visual language discrimination in infancy. *Science* 316:1159. doi: 10.1126/science.1137686
- Werker, J. F., Pegg, J. E., and McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behav. Dev.* 17, 323–333. doi: 10.1016/0163-6383(94)90012-4
- Wong, P. C. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Res. Bull.* 59, 83–95. doi: 10.1016/S0361-9230(02)00860-2
- Xu Rattanasone, N., Burnham, D., and Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *J. Phon.* 41, 332–343. doi: 10.1016/j.wocn.2013.06.001
- Yeung, H. H., Chen, K. H., and Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *J. Memory Lang.* 68, 123–139. doi: 10.1016/j.jml.2012.09.004
- Yeung, H. H., Chen, L. M., and Werker, J. F. (2014). Referential labeling can facilitate phonetic learning in infancy. *Child Dev.* 85, 1036–1049. doi: 10.1111/cdev.12185
- Zatorre, R. J., and Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1087–1104. doi: 10.1098/rstb.2007.2161

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kager. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.