Perception and lexical encoding of tone in a restricted tone language: Developmental evidence from Limburgian

Stefanie Ramachers[1], *Susanne Brouwer*[1], *& Paula Fikkert*[1]

[1]Centre for Language Studies, Radboud University, Nijmegen
s.ramachers@let.ru.nl, s.brouwer@let.ru.nl, p.fikkert@let.ru.nl

Abstract

Within the large family of tone languages, differences exist with respect to the importance and phonetic realization of tones. It remains unclear how these differences influence the acquisition and processing of lexical tone. Limburgian, spoken in the south of the Netherlands, is assumed to have lexical tone, but it has a lower functional load than for example Mandarin Chinese. Moreover, lexical tone in Limburgian is subject to an intriguing amount of surface variation [1].

We compared performance of native Limburgians to native non-tonal Dutch controls on a series of behavioral experiments: Discrimination of lexical tone in the first year of life (Exp. 1) and in adulthood (Exp. 2), and the encoding of lexical tone during word learning in toddlers and adults (Exp. 3). Our results partly deviate from previous research on tone languages like Mandarin, but are on a par with recent findings with Japanese children, suggesting that the acquisition and processing of lexical tone indeed seems to be influenced by functional load and phonetic variability.

Index Terms: lexical tone, restricted tone language, acquisition, discrimination, word learning, word recognition

1. Introduction

Psycholinguistic research on the acquisition and processing of lexical tones has typically investigated well-known tone languages such as Mandarin, Cantonese, or Thai (for a review, see [2]). These studies yielded mixed results, possibly due to the use of different procedures and different tone contrasts varying in acoustic salience. There is limited research dedicated to the acquisition of lexical tones in more *restricted* tone systems like Japanese and Swedish [3-8].

This study investigates Limburgian, which can be considered a restricted tone language. Speakers of Limburgian also have command of non-tonal Standard Dutch, which is the official national language. Limburgians can thus be considered bidialectal [9]. As in Norwegian and Swedish, in many Limburgian dialects a word prosodic contrast can signal lexical and morphological differences. In the dialect of Roermond, for example, *haas* [ha:s] with falling pitch (accent 1) means 'hare', whereas *haas* with falling-rising pitch (accent 2) means 'glove'. In a small number of frequent nouns in this dialect, pitch differences also serve a grammatical function with accent 1 systematically indicating plurality. For example, *knien* [kni:n] with accent 2 means 'rabbit', but pronounced with accent 1 it means 'rabbits' (see Fig. 1A and 1B; subscripts indicate accents 1 and 2). Relative to Mandarin, Limburgian has few tonal minimal pairs. Fournier [10] counted around 80 tonal minimal pairs in a dictionary of Roermond Limburgian. Moreover, there is only a two-way contrast. Also, the distribution of tones is subject to syllabic restrictions, in that the contrast can only be realized on minimally bimoraic syllables with primary stress [11].

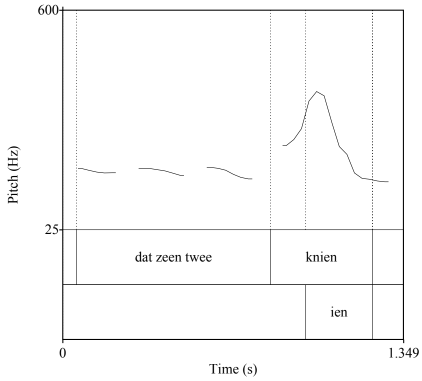

Figure 1A: *F0 contour of the Limburgian sentence* dat zeen twee KNIEN$_1$ *'those are two rabbits'. The rhyme of the target word carries accent 1.*

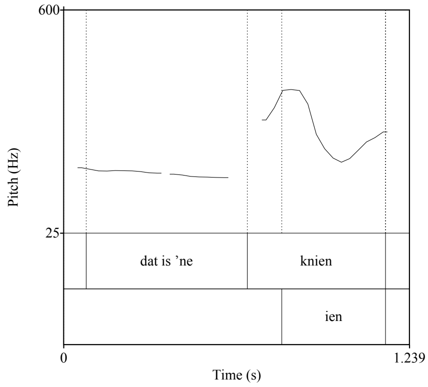

Figure 1B: *F0 contour of the Limburgian sentence* dat is 'ne KNIEN$_2$ *'that is a rabbit'. The rhyme of the target word carries accent 2.*

With respect to the functional load of tone, we assume Limburgian to be comparable to Swedish [12-13], Norwegian [14-15], and Japanese [16]. As in most other Franconian dialects [17], and comparable to Swedish [18], Limburgian lexical tones co-occur with intonation tones in syllables with main stress. As a consequence, the Limburgian tones show up in different shapes as a function of information status, sentence type, and position in the intonational phrase [11]. Studies on lexical tone acquisition in Swedish [3,4], Japanese [3,5,8], and Sesotho [19] suggest that the reliability of the mapping between underlying tones and their surface realizations as well as the functional load has an impact on the speed of acquisition of tone. The Limburgian tone system might thus pose challenges to its learners. To see if this indeed holds true, we tested the perception and lexical encoding of tone in Limburgian infants, toddlers, and adults.

2. Experiment 1

In our first experiment, we compared 6- to 12-month-old Limburgian infants to Dutch infants to investigate language- and/or age-related effects on infants' ability to discriminate the Limburgian tone contrast [20].

2.1. Participants

Thirty-nine bidialectal Limburgian infants were included in the analysis: 11 six-month-olds, 10 nine-month-olds, and 18 twelve-month-olds. Infants who were exposed to any East-Limburgian dialect spoken by at least one caregiver were allowed to participate. Eighty-three monolingual Dutch infants were included in the analysis, of which 28 six-month-olds, 29 nine-month-olds, and 26 twelve-month-olds. None of the Dutch infants had substantial experience with Limburgian or another tone language.

2.2. Stimuli

Stimuli were two pseudo-words that only differed in tone: $taag_1$ [taːç] and $taag_2$ [taːç]. They were recorded by a female native speaker of East-Limburgian. Five tokens per tone type were selected. Independent samples *t*-tests showed that accent 1 and accent 2 stimuli did not differ significantly with respect to duration (p = .86). However, they did differ significantly with respect to the timing of the pitch peak (p = .001) and the pitch range (p = .002) as observed in the nucleus of the syllable.

2.3. Procedure

Infants' discrimination was tested using the hybrid visual habituation procedure [21]. The experiment was divided into a habituation phase and a test phase. During habituation, infants listened to repetitions of 4 tokens of *taag* carrying either accent 1 or accent 2. Once the infants were habituated, indicated by an attention decrease of 35%, they were presented with 8 trials featuring alternations between 4 old tokens and 1 new token with the habituated tone (*same* trials) and 4 trials with alternations between the 4 old tokens of the habituated tone and 1 new token with the other tone (*switch* trials). Successful discrimination is signaled by a significant difference in looking time between same and switch trials.

2.4. Results

The results of a mixed ANOVA with Trial Type (same vs. switch) as the within-subjects variable and Language (Limburgian vs. Dutch) and Age (6 vs. 9 vs. 12 months) as between-subjects variables yielded a main effect of Trial Type, indicating that both Limburgian and Dutch infants discriminated the Limburgian tones throughout their first year of life (F (1,116) = 15.90, p < .001, η_p^2 = 0.12). No other main effects or interactions reached significance (all p's >.01).

2.5. Discussion

The success of the Dutch infants could be due to the degree of similarity to the native prosodic system. The Dutch infants might have perceived the Limburgian accents as native intonation patterns. The fact that the Limburgian infants were not more sensitive than their Dutch peers might have to do with the fact that tone in Limburgian is relatively less prominent and children need more exposure to increase their sensitivity. The question arose whether we would find a significant difference in tone sensitivity in adult speakers of Limburgian and Dutch.

3. Experiment 2

To investigate adults' tone discrimination ability, a categorical AXB-discrimination task was run. Participants heard sequences of three stimuli and had to judge whether the second token (X) most resembled the first (A) or third (B) token. We were interested in language-related differences in tone discrimination ability, as well as in the 'categoricalness' of tone perception in Limburgian listeners [22].

3.1. Participants

Twenty adult speakers of East-Limburgian (*M* age 49 yrs) and 19 adult speakers of Dutch (*M* age 45 yrs) were tested. Dutch adults had no substantial experience with Limburgian, and none of them reported command of a tonal second language. They all reported normal hearing and no speech, language, or attention deficits, and we kept the relative amount of musically trained individuals comparable across groups.

3.2. Stimuli

Test stimuli were four pseudo-words that could be pronounced with accent 1 and accent 2: Two monosyllables, *taag* [taːç] and *moon* [moːn], and two disyllables, *keeve* [keːvə] and *perger* [pɛʀɣəʀ]. The disyllabic non-words had trochaic stress. Thus, accents 1 and 2 were realized on the penultimate syllable. As the prosodic context influences the realization of the tones, accents 1 and 2 surfaced as a falling tone and a level tone, respectively, in the disyllabic stimuli. For each pseudo-word, four acoustically different tokens were selected. As in Exp. 1, accent 1 and accent 2 tokens of the stimuli only differed in F0 measures and not in duration.

3.3. Procedure

During the test phase of the experiment, participants proceeded through a total number of 192 trials, 96 test trials and 96 filler trials. Among the 96 test trials, we introduced two conditions. The first condition constituted 64 between-category variation trials. In this condition, tonal minimal pairs (which shared the same segmental structure) served as stimuli to test lexical tone discrimination. Crucially, half of the between-category trials featured monosyllabic pseudo-words and half of the trials featured disyllabic pseudo-words. In the disyllabic words, the tones occurred in non-final position. As such, they do not signal a linguistically meaningful contrast for speakers of Dutch [23]. As a consequence, we expected Dutch listeners to show poorer discrimination performance in these trials as compared to monosyllabic trials. The X token was always physically different from the categorically matching A or B token, so that listeners could not make a simple acoustic identity judgment and had to ignore acoustic differences that are not phonetically relevant [24-26].

The second condition constituted 32 within-category variation trials. The three stimuli in these triads showed the same tonal pattern with only subtle differences in their actual instantiation of the pitch contour. These trials were added to see if pitch perception by Limburgians is more categorical than pitch perception by Dutch listeners. If their perception is indeed driven by lexical tone categories, Limburgian listeners were expected to be worse in perceiving within-category than between-category pitch differences (e.g., [27-28]). In these trials, the X token was physically identical to the A or B token.

The 96 filler trials featured segmental minimal pairs to distract participants' attention away from the purpose of the experiment, i.e., lexical tone discrimination.

Participants were allowed to pause three times during the experiment, i.e. after every 48 trials. The experiment lasted approximately 25 minutes. The dependent measure was accuracy (proportion of correct responses).

3.4. Results

3.4.1. Between-category variation trials

A mixed ANOVA with Condition (monosyllabic, disyllabic) and presentation Order (accent 1 > accent 2, accent 2 > accent 1) as within-subjects variables and Language (Limburgian, Dutch) as a between-subjects variable was conducted, showing a significant main effect of Language, $F(1,34) = 43.70$, $p < .001$, $\eta_p^2 = .56$, indicating that Limburgian listeners ($M = .86$, $SE = .02$) outperformed Dutch listeners ($M = .69$, $SE = .02$). The analysis also yielded a significant main effect of Condition, $F(1,34) = 12.64$, $p = .001$, $\eta_p^2 = .27$, indicating that both participant groups performed better in disyllabic trials ($M = .81$, $SE = .02$) than in monosyllabic trials ($M = .75$, $SE = .02$). However, these main effects had to be interpreted in light of a significant Condition x Order x Language interaction, $F(1,34) = 8.52$, $p = .006$, $\eta_p^2 = .20$. To break down the three-way-interaction, repeated-measures ANOVAs with Condition and Order as the within-subjects variables were run for each language group separately.

Limburgians performed significantly better in disyllabic trials ($M = .90$, $SD = .08$) compared to monosyllabic trials ($M = .81$, $SD = .13$), $t(17) = -3.27$, $p = .005$, Cohen's $d = -0.77$, but only if first presented with an accent 2 token. The repeated-measures ANOVA on the Dutch participants' data revealed a main effect of Condition, $F(1,17) = 5.43$, $p = .03$, $\eta_p^2 = .24$. Dutch subjects performed better on disyllabic ($M = .72$, $SE = .02$) compared to monosyllabic trials ($M = .66$, $SE = .03$). No other significant main effects or interactions were attested (all p's > .05).

3.4.2. Between- vs. within-category variation

A mixed ANOVA with Trial Type (between vs. within-category variation) as a within-subjects variable and Language as a between-subjects variable was conducted. A significant main effect of Language was detected, $F(1,35) = 15.31$, $p < .001$, $\eta_p^2 = .30$, indicating that Limburgian listeners outperformed ($M = .84$, $SE = .02$) Dutch listeners ($M = .75$, $SE = .02$) overall. The analysis also yielded a main effect of Trial Type, $F(1,35) = 8.90$, $p = .005$, $\eta_p^2 = .20$, demonstrating that participants scored better in within-category trials ($M = .83$, $SE = .02$) than in between-category trials ($M = .77$, $SE = .01$). Moreover, we found a significant Trial Type x Language interaction, $F(1,37) = 17.74$, $p < .001$, $\eta_p^2 = .30$, showing that the main effect of Trial Type was carried by the Dutch participants: they performed significantly better in within-category trials ($M = .81$, $SD = .11$) than in between-category trials ($M = .69$, $SD = .07$; $t(17) = -4.65$, $p < .001$, Cohen's $d = -1.10$). However, Limburgians performed equally well in between-category ($M = .85$, $SD = .10$) and within-category trials ($M = .84$, $SD = .07$; $t(18) = .49$, $p > .05$, Cohen's $d = 0.11$).

3.5. Discussion

Our results demonstrated that Limburgian adults are significantly more accurate than Dutch adults in between-category variation trials, which is most likely due to the lexical distinctiveness of tone in Limburgian. This is in line with previous research suggesting that tone language speakers are at an advantage in discriminating lexical tones compared to naïve non-tone language speakers (e.g., [27-31]).

A second important result is that Limburgian listeners performed equally well in between- and within-category trials. This is not in line with prior studies (e.g., [27-28, 32-33]). A possible explanation for this finding might be the distinctive properties of the Limburgian tonal system: Limburgian listeners have to deal with an intriguing amount of linguistically meaningful pitch variation due to the interaction of tonal and intonational tones, which is typically absent in previously studied tone languages. In fact, the tonal surface variation that Limburgian listeners are exposed to cannot be considered irrelevant phonetic variation, since it signals meaningful information at the post-lexical level. This surface variation might cause them to exhibit a greater sensitivity to (linguistic) pitch in general.

A third result that was against our expectations was that Dutch participants performed significantly better on disyllables than on monosyllables. We propose that the higher discrimination accuracy in disyllables could be attributed to acoustic salience. In trochaic disyllables, where accents 1 and 2 surface as a fall and as a level tone, respectively, the difference between the tones may have been more apparent than in monosyllables, where accents 1 and 2 only differ in the final rise.

The question therefore no longer is whether Limburgian and Dutch listeners perceive the difference between accents 1 and 2, but whether they treat this difference as lexically relevant. Do they attend to word-level pitch variation when learning and recognizing novel words? This question was addressed in our third experiment.

4. Experiment 3

Limburgian and Dutch 2,5- to 4-year-olds as well as adults took part in a word learning experiment [34]. Following the procedure employed by [35] and [36], participants learned two novel word-object mappings. After training, word recognition was tested in correct pronunciation (CP) trials and mispronunciation (MP) trials featuring a pitch change. Based on previous studies [35-38], we expected that Limburgian listeners would notice tonal MPs, but Dutch listeners would not.

4.1. Participants

Twenty-three bidialectal Limburgian children (M age = 40.9 months) and 35 monolingual Dutch toddlers (M age = 36.8 months) were included in the analysis. None of the Dutch toddlers had substantial exposure to Limburgian.

In addition, we tested 14 Limburgian adults (M age 53.6 yrs) and 22 Dutch adults (M age 23 yrs). All Limburgian and Dutch participants reported some degree of non-native command of one or more non-tonal languages, but none of them had experience with a tone language. The number of musically trained individuals was again held constant across groups.

4.2. Stimuli

We created two pseudo-words: *taaf* [ta:f] and *moon* [mo:n], which were recorded both with accent 1 and accent 2. The target stimuli were equally compatible with Limburgian and Dutch. In total, 12 tokens of each word with each tone were selected. Independent *t*-tests revealed that accent 1 and accent 2 tokens differed significantly from each other with respect to minimum f0, maximum f0 and f0 range. Carrier sentences were recorded in Limburgian and in Dutch.

The visual target stimuli consisted of four unknown plush toy objects of an animate character in different colors.

4.3. Procedure

The procedure employed was the intermodal preferential looking paradigm [39]. In two subsequent blocks, participants learned two novel word-object mappings. Subsequently, it was tested how they reacted to a pitch change in the newly learned word. Which word carried what tone and which word was taught first was counterbalanced across subjects. For Limburgian participants, the entire experiment was in Limburgian. For Dutch participants, the experiment was in Dutch. Across language contexts, only the tokens of the target stimuli *taaf* and *moon* were the same.

The learning phase consisted of four trials of 30 s each. In the first and the third trial, the target object was labeled twenty times in total in a declarative carrier sentence. In trials two and four, a distracter object appeared and was talked about for an equal amount of time, but did not receive a label.

The subsequent test phase consisted of four test trials and four filler trials. In test trials, the target and the distracter toy appeared side by side. Participants were asked to "Look at the [target]." Target naming always occurred at 2500 ms to enable participants to inspect both objects prior to naming and to establish a baseline preference. In two of the test trials, the label for the target object was correctly pronounced (CP trials), while in the other two, the label was mispronounced (MP trials). This MP involved a change in pitch: A word taught with accent 1 was mispronounced with accent 2 and vice versa. The presence of a nameless distracter offered participants the possibility of considering the mispronounced version of the target label to be a novel label for the unlabeled distracter. To make sure children would remain engaged in the task, four filler trials involved correct pronunciations of four well-known words. For adults, we included 16 filler trials to distract their attention away from the purpose of the experiment. Between blocks participants watched a 1-minute video featuring farm animals. The second block introduced a new pair of novel objects and a new object label, but the structure was the same.

Target recognition was inferred from the presence of a naming effect. To calculate the naming effect, the increase in the proportion of target looking between the 2500 ms pre-naming window and the post-naming window (2000 ms for children, 1000 ms for adults) of a test trial was calculated, resulting in a difference score (e.g., [38]).

4.4. Results

Both Limburgian and Dutch toddlers showed a significant positive naming effect in CP trials, meaning that word learning was successful ($p < .001$). A three-way mixed ANOVA with Condition (CP vs. MP) and Tone (Accent 1 vs. Accent 2) as within-subjects factors and Language (Limburgian vs. Dutch) yielded a significant main effect of Condition ($F(1,56) = 8.53$, $p = .005$, $\eta_p^2 = .13$), showing a significantly larger naming effect in CP trials ($M = .21$, $SD = .20$) than in MP trials ($M = .09$, $SD = .24$). No other effects or interactions were found (all p's > .1). To investigate the strength of the MP, the naming effect in MP trials was compared to zero by means of a one-sample *t*-test. The test revealed a significant positive naming effect ($M = .09$, $SD = .24$; $t(57) = 2.81$, $p < .01$, Cohen's $d = .37$).

The same analyses were run on the adult data. Adults successfully learned the new words, but the ANOVA did not yield any main effects or interactions (all p's > .05). As in the CP trials, the naming effect in MP trials was significantly above zero ($M = .34$, $SD = .22$; $t(38) = 9.53$, $p < .001$, Cohen's $d = 1.53$).

4.5. Discussion

Our main finding is that both Limburgian and Dutch children pay attention to pitch changes in newly learned words. However, children still preferred the target object over the distracter object upon hearing a pitch change, indicating that a tonal MP did not hinder word recognition to a great extent. For our adult listeners, a pitch change did not seem to hinder word recognition at all. Previous studies investigating Mandarin toddlers found no naming effects in tonal MP conditions [36,38], whereas recent studies with Japanese toddlers did find a naming effect [5,8], suggesting that pitch changes are more detrimental to word recognition in Mandarin than in Limburgian and Japanese. The low frequency of tonal minimal pairs, plus the great amount of surface variation might mitigate the reliance on pitch when learning novel words in Limburgian and Japanese. Similar explanations have been put forward in studies investigating the role of lexical stress in English [40-41] and Italian [42]. Dutch toddlers could have interpreted the Limburgian pitch patterns as post-lexical intonation [36-37,43].

5. Conclusions

Our data suggest that the distinctive properties of the Limburgian lexical tone system influence its acquisition and processing. Our discrimination data show that Limburgian infants do not yet show greater sensitivity to word-level pitch than non-tonal Dutch peers. This suggests that more input is needed to develop this sensitivity, possibly due to the low functional load of tone in Limburgian. Limburgian adults, however, show a striking sensitivity to very subtle, within-category differences, which is possibly caused by their daily experience with a high amount of surface variability. These factors combined could cause the weak reliance on pitch when learning novel words. Future studies could investigate the impact of tone changes on the recognition of *known* words in Limburgian, as well as the lexical encoding of tone in other restricted tone languages like Swedish and Norwegian.

6. Acknowledgments

We thank all parents, children, and adult participants for contributing to our experiments. We also thank Stichting Kinderopvang Roermond, GGD Limburg-Noord, de Bibliotheek, and all student assistants and interns for their help during recruitment and testing. Thanks also to Carlos Gussenhoven for his assistance in creating the target stimuli and for helpful discussion throughout the project.

7. References

[1] C. Gussenhoven and P. van der Vliet. "The phonology of tone and intonation in the Dutch dialect of Venlo," *Journal of Linguistics*, vol. 35, pp. 99-135, 1999.

[2] L. Singh and C.A. Fu. "A new view of language development: The acquisition of lexical tone," *Child Development*, vol. 87, pp. 834–854, 2016.

[3] M. Ota. "The development of lexical pitch accent systems," *Canadian Journal of Linguistics*, vol. 48, pp. 357-383, 2003.

[4] M. Ota. "Children's production of word accents in Swedish revisited," *Phonetica*, vol. 63, pp. 230-246, 2006.

[5] M. Ota, N. Yamane, and R. Mazuka. "The effects of lexical pitch accent on infant word recognition in Japanese," *Frontiers in Psychology*, doi: 10.3389/fpsyg.2017.02354, 2018.

[6] K. Plunkett and S. Strömqvist. "The acquisition of Scandinavian Languages," in *The crosslinguistic study of the language acquisition,* 3, D. Slobin, Ed. New Jersey/London: Lawrence Erlbaum, 1992, pp. 457-556.

[7] Y. Sato, Y. Sogabe, and R. Mazuka. "Development of hemispheric specialization for lexical pitch-accent in Japanese infants," *Journal of Cognitive Neuroscience*, vol. 22, pp. 2503-2513, 2010.

[8] H.W. Yamamoto and E. Haryu. "The role of pitch pattern in Japanese 24-month-olds' word recognition," *Journal of Memory and Language*, vol. 99, pp. 90-98, 2018.

[9] L. Cornips. "Socio-syntax and variation in acquisition: problematizing monolingual and bidialectal acquisition," *Three factors and beyond: Socio-syntax and language acquisition. Special issue of Linguistic Variation*, vol. 14, pp. 1-25, 2014.

[10] R. Fournier. *Perception of the tone contrasts in East-Limburgian dialects*. Utrecht: LOT, 2008.

[11] C. Gussenhoven. "The lexical tone contrast of Roermond Dutch in Optimality Theory," in *Prosody: Theory and Experiment.* M. Horne, Ed. Amsterdam: Kluwer, 2000, pp. 129-167.

[12] C. Gussenhoven. "Tone in Germanic: Comparing Limburgian with Swedish," in *From Traditional Phonology to Modern Speech Processing*. G. Fant, H. Fujisaki, J. Cao, and Y. Xu, Eds. Beijing: Foreign Language Teaching and Research Press, 2004, pp. 129-136.

[13] T. Riad. *The phonology of Swedish.* Oxford: Oxford University Press, 2013.

[14] G. Kristoffersen. *The phonology of Norwegian.* Oxford: Oxford University Press, 2000.

[15] A. Wetterlin. *The Lexical Specification of Norwegian Tonal Word Accents.* Doctoral dissertation, University of Konstanz, 2007.

[16] H. Kubozono. *The organization of Japanese prosody*. Tokyo: Kurioso, 1993.

[17] B. Köhnlein. "Contrastive foot structure in Franconian tone-accent dialects," *Phonology*, vol. 33, pp. 87-123, 2016.

[18] G. Bruce. *Swedish word accents in sentence perspective*. Lund: Gleerup, 1977.

[19] K. Demuth. "The acquisition of tonal systems," in *Phonological acquisition and phonological theory.* J. Archibald, Ed. Hillsdale: Lawrence Erlbaum, 1995, pp. 111-134.

[20] S. Ramachers, S. Brouwer, and P. Fikkert. "No perceptual reorganization for Limburgian tones? A cross-linguistic investigation with 6- to 12-month-old infants. *Journal of Child Language,* doi: 10.1017/S0305000917000228, 2017.

[21] D. Houston, D. Horn, R. Qi, J. Ting, and S. Gao. "Assessing speech discrimination in individual infants," *Infancy*, vol. 12, pp. 119-145, 2007.

[22] S. Ramachers. *Setting the Tone: Acquisition and processing of lexical tone in East-Limburgian dialects of Dutch.* Doctoral dissertation. Utrecht: LOT, 2018.

[23] B. Braun and E. Johnson. "Question or tone 2? How language experience and linguistic function guide pitch processing," *Journal of Phonetics,* vol. 39, pp. 585-594, 2011.

[24] C. Best, G. McRoberts, and E. Goodell. "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," *Journal of the Acoustical Society of America,* vol. 109, pp. 775-794, 2001.

[25] W. Strange and V. Shafer. "Speech perception in second language learners: Re-education of selective perception," in *Phonology and second language acquisition*. J. Hansen Edwards and M. Zampini, Eds. Amsterdam: John Benjamins, 2008, pp. 153-191.

[26] M. Tyler, C. Best, A. Faber, and A. Levitt. "Perceptual assimilation and discrimination of non-native vowel contrasts," *Phonetica,* vol. 71, pp. 4-21, 2014.

[27] P. Hallé, Y. Chang, and C. Best. "Identification and discrimination of Mandarin lexical tones by Mandarin Chinese vs. French listeners," *Journal of Phonetics,* vol. 32, pp. 395-421, 2004.

[28] Y. Xu, J. Gandour, and A. Francis. "Effects of language experience and stimulus complexity on the categorical perception of pitch contrasts," *Journal of the Acoustical Society of America,* vol. 120, pp. 1063-1074, 2006.

[29] V. Schaefer and I. Darcy. "Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones," Laboratory Phonology, vol. 5, pp. 489-522, 2014.

[30] X. Tong, S. Lee, M. Lee, and D. Burnham. "A tale of two features: Perception of Cantonese lexical tone and English lexical stress in Cantonese-English bilinguals," *PLoS ONE,* vol. 10, e0142896. doi:10.1371/journal.pone.0142896, 2015.

[31] R. Wayland and S. Guion. "Perceptual discrimination of Thai tones by naïve and experienced learners of Thai," *Applied Psycholinguistics,* vol. 24, pp. 113-129, 2003.

[32] C. Hoffman, M. Sadakata, A. Chen, P. Desain, and J. McQueen. "Within-category variance and lexical tone discrimination in native and non-native speakers," *TAL 2014 Proceedings,* pp. 45-49, 2010.

[33] Z. Qin and P. Mok. "Discrimination of Cantonese tones by Mandarin, English and French speakers," *Proceedings of the Psycholinguistic Representation of Tone,* Hong Kong: Causal Productions, pp. 50-53, 2011.

[34] S. Ramachers, S. Brouwer, and P. Fikkert. "How native prosody affects pitch processing during word learning in Limburgian and Dutch toddlers and adults," *Frontiers in Psychology,* vol. 8, doi: 10.3389/fpsyg.2017.01652, 2017.

[35] C. Quam and D. Swingley. "Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning," *Journal of Memory and Language,* vol. 62, pp. 135-150, 2010.

[36] L. Singh, T.J. Hui, C. Chan, and R. Golinkoff. "Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation," *Developmental Science,* vol. 17, pp. 94-109, 2014.

[37] J. Hay, K. Graf Estes, T. Wang, and J. Saffran. "From flexibility to constraint: the contrastive use of lexical tone in early word learning," *Child Development,* vol. 86, pp. 10-22, 2015.

[38] L. Singh, H. Goh, and T. Wewalaarachchi. "Spoken word recognition in early childhood: comparative effects of vowel, consonant and lexical tone variation," *Cognition,* vol. 142, pp. 1-11, 2015.

[39] R. Golinkoff, K. Hirsh-Pasek, K. Cauley, and L. Gordon. "The eyes have it: lexical and syntactic comprehension in a new paradigm," *Journal of Child Language,* vol. 14, pp. 23-45, 1987.

[40] A. Cutler. "Forbear is a homophone: lexical prosody does not constrain lexical access," *Language and Speech,* vol. 29, pp. 201-220, 1986.

[41] C. Quam and D. Swingley. "Processing of lexical stress cues by young children," *Journal of Experimental Child Psychology,* vol. 123, pp. 73-89, 2014.

[42] S. Sulpizio and J. McQueen. "When two newly acquired words are one: New words differing in stress alone are not automatically represented differently," *Proceedings of Interspeech 2011,* pp. 1385-1388, 2011.

[43] B. Braun, T. Galts, and B. Kabak. "Lexical encoding of L2 tones: the role of L1 stress, pitch accent and intonation," *Second Language Research,* vol. 30, pp. 323-350, 2014.