

Setting the Tone:

Acquisition and processing of lexical tone in East-Limburgian dialects of Dutch

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: Adobe Stock.

ISBN: 978-94-6093-269-4
NUR 616

Copyright © 2018: Stefanie Ramachers. All rights reserved.

Setting the Tone:
Acquisition and processing of lexical tone
in East-Limburgian dialects of Dutch

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 31 januari 2018
om 12.30 uur precies

door
Stefanie Theodora Maria Rosa Ramachers
geboren op 11 november 1987
te Geleen

Promotoren: Prof. dr. Paula Fikkert, Prof. dr. Carlos Gussenhoven

Copromotor: Dr. Susanne Brouwer

Manuscriptcommissie:

Prof. dr. James McQueen

Prof. dr. Marc van Oostendorp

Prof. dr. Leonie Cornips (Meertens Instituut, Universiteit Maastricht)

Dr. Yiya Chen (Universiteit Leiden)

Dr. Mitsuhiro Ota (University of Edinburgh, Verenigd Koninkrijk)

The research reported here was supported by the Netherlands
Organization for Scientific Research under project number 322-75-001.

*“Da Menschen Sprache gebrauchen, wie es kein Tier vermag,
dürften wir auch die menschliche Natur ein wenig besser begreifen,
wenn wir das Wesen der Sprache verstehen.”*

- Paul Ibbotson & Michael Tomasello, 22.02.2017, Spektrum.de

Acknowledgments

A journey has come to an end... It started in my second year as a student at Radboud University when I chose a minor in general linguistics. One of the courses I attended was about first language acquisition. It opened up a whole new world to me. And here I am, ten years later, (hopefully) about to obtain a PhD in first language acquisition. During this journey, a lot of people supported and encouraged me.

First of all, I would like to thank my supervisors: Paula Fikkert, Carlos Gussenhoven, and Susanne Brouwer.

Paula – you encouraged me to apply for an individual NWO PhD project grant by extending a research proposal I had written during one of your master's courses. Who knows what my career would have looked like if I had not attended this course? Actually, you encouraged me to do a lot of things, for example to participate in the scientific pitch competition Radboud Talks. But you also hit the brakes if necessary. I could not have wished for a more accessible, caring, encouraging, optimistic, and supportive supervisor. Thanks for your invaluable contribution to my development as an independent researcher and as a proud lecturer.

Carlos – our meetings always started with a description of the status quo of my project. What was your last experiment again? Indeed, we did not meet very often. Nevertheless, you were always there to discuss a variety of topics relating to prosodic typology (what do we need pitch-accent languages for?) or stimulus characteristics (don't you hear the difference in voice quality?). I also enjoyed your attempts to speak Limburgian (or any other tone language). Thank you for sharing your bottomless expertise with me.

Susanne – although you joined my project at a later stage, it feels like we have been working together for years. I really appreciate your hands-on, pragmatic approach and openness. We make a great team. We both need food, at any time, a laugh, at any time, and yoga, preferably each day (even in a hotel room when visiting a conference). I had a great time visiting *SPPL2017* in

London with you. Thank you for cheering me up and boosting my confidence. You made me feel at ease, which is something I value a lot. I sincerely hope our paths will keep crossing in the future.

Next, I would like to thank my reading committee. James McQueen, Marc van Oostendorp, Leonie Cornips, Yiya Chen, and Mits Ota: Thanks a lot for your contributions. They enabled me to put the cherry on the cake. You all approached my thesis from a different angle, which makes sense, since my thesis turned out to address many different scientific discussions (some of them more elaborately than others) with respect to typology, multilingualism, methodology, etc.

I did not embark on the PhD journey on my own, of course. A few times a month, our PI Group First Language Acquisition met to discuss experimental design issues, (preliminary) results, articles, upcoming conferences, oral or poster presentations, etc. Thanks to all of you: Titia Benders, Christina Bergmann, Susanne Brouwer, Helen Buckler, Nienke Dijkstra, Paula Fikkert, Laura Hahn, Evelien Krikhaar, Vanja de Lint, Ingeborg Roete, Nadine de Rue, Tineke Snijders, Antje Stöhr, Sho Tsuji, and Maarten Versteegh.

From the very beginning of my PhD project in 2011, I have also been a lecturer at the department of German Language and Culture. I developed and taught my own courses in German linguistics, supervised my first bachelor's theses, joined the curriculum committee, and I even got nominated for the faculty's teaching award in 2015. I couldn't have wished for a more valuable experience next to my PhD research. Teaching, seeing students grasping a concept and becoming enthusiastic about language, is an enormous source of satisfaction. A big thank you to all my students, and to my (former) colleagues for nourishing my other 'personality' (the Germanist): Yvonne Delhey, Arash Farhidnia, Ad Foolen (my mentor and, quite frankly, the person who woke up the linguist in me. Ad, I cannot thank you enough!), Annemarie Frye, Giel van Gemert, René Gerritsen (my *Pendelbus*-buddy and the person who makes me feel proud of being able to recite the *Nibelungenlied*), Liesbeth van de Grift, Sonja Häffner (the perfect roomy and sparring partner when it comes to

grammatical issues in German, *Aachener Printentee* and sushi), Jana Hermann, Sabine Jentges (*lächle, morgen wird's noch schlimmer*, thank you for constantly putting flowers, cookies and sweets in my office), Godelief de Jong (our *fashionista*, always there in times of need), Christa van Mourik (the other perfect roomy and the perfect *regelneef*), Johan Oosterman, Quirine Peters – van Winsen (a.k.a. Q.), Erika Poettgens, Margit Rem, Paul Sars, Mieke Steins, Klaar Vernailen, and Jantine Vingerhoets.

Recruiting and testing babies, toddlers and adults in Limburg was the greatest challenge I encountered during my project. There was no babylab in Limburg, and neither was there a participant database, meaning that I had to personally arrange a testing location, testing equipment, and participants. Lots of people helped me along the way. My first stop for support was at the Huis voor de Kunsten Limburg, where I met *streektaalfunctionaris* Ton van de Wijngaard. He advised me to contact the Roermond city council. Councillor Wim Kemp put me into contact with *Stichting Kinderopvang Roermond*. Employees from SKR then brought me to daycare center *Ot en Sien*, where I was kindly welcomed by managers Wendy Janssen and Jolanda Schenk, and assistant manager Janine Janssen. They gave me the opportunity to set up my own little lab in their office and test my participants there, and to recruit participants at *Ot en Sien*. Soon, it turned out that I could not collect enough participants at only one daycare center, so the people at *Ot en Sien* brought me into contact with their colleagues at *GGD Limburg Noord*. I want to thank Petra Tegelbeckers, Manon Steinschuld, and the other employees at *GGD Limburg Noord* for another warm welcome and for facilitating participant recruitment among their client pool. For my word learning experiment, I needed older children, which I recruited at a later stage at daycare center *Donderstenen*. Thank you, Loes van de Borgh, for helping me out in these final stages of testing! Participant recruitment has been an ongoing business throughout the project. I am grateful for each and every parent and child who lend me their trust and support.

Next to finding a location and a subject pool, I had to start setting up my testing equipment. I want to thank Joop Kerkhoff, Foto Grijpink Nijmegen, Harrie Harings, Han Sloetjes, Johan Weustink, Paul Trilsbeek, and Bas van Gestel for their technical support, even if some of them couldn't help me in the end, which sometimes caused serious amounts of frustration. Thanks also go to my dad, for constructing a wooden frame that resembled the set-up in the Nijmegen Babylab, and to my mum, for driving me and my equipment back and forth and for being my first participant to check whether everything was working. All in all, it took me about half a year to get my lab up and running.

Lab support in Nijmegen was provided by Angela Khadar, Margret van Beuningen, and Bob Rosbag. Of course, I also want to mention the interns and student assistants that spent some time in our research group and contributed to my research, in one way or another: Maaike ten Buuren, Catheleyne Creusen, Renske van der Cruijssen, Arjan Cuppens, Daniela Dincheva, Dean Hermans, Imme Lammertink, Chrissy Laurentzen, Lisette Maes, Lotte Minheere, Clementine Pizzey-Gray, Natascha Roos, and Romy Roumans. A big thank you also goes to my friend and fellow-linguist Mieke Bussemakers for re-coding part of my data, and to speech therapist Willeke Wassenberg for being the cheerful Limburgian and Dutch voice in my experiments.

Many of my adult participants were tested in public libraries in Nijmegen and Roermond. In Nijmegen, we were kindly welcomed by Martijn Mommers from Bibliotheek 'De Mariënborg'. In Roermond, the discrimination and word learning experiments with adults were run in Bibliotheek 'Bibliorura'. Thanks go to Suzan Crapels, Marion Knippenberg, Lidy Vos, Ivo Wiermans, and other employees who facilitated our work. A big thank you also goes to all my adult participants in Nijmegen and Roermond.

A big thanks also goes to my friend and graphic designer Iris Dellink, who designed beautiful information brochures for my participants and helped me create my experimental stimuli.

For statistical support, I am indebted to Susanne Brouwer, Roeland van Hout, and Frans van der Slik.

I'm also thankful for helpful discussion and support from colleagues at other departments and universities, both in the Netherlands and abroad. I would like to mention Frens Bakker, Elma Blom, Tessel Boerma, Ao Chen, Leonie Cornips, Kirsten van den Heuij, Caroline Junge, Brigitta Keij, Liquan Liu, Karen Mattock, Sharon Unsworth, and Patrick Wong for conversations about prosodic acquisition, recruiting participants, coding issues, language background questionnaires for bilinguals, and lots more. In setting up my word learning experiment, Leher Singh and Daniel Swingley were so kind to share their expertise with me.

I also want to thank dialect-association Veldeke Limburg for inviting me to contribute to their yearbook twice. These were great opportunities for me to write down what I was doing for a lay audience. Thanks also to Jan Sjure for mentioning my research in his column 'Druimentaere' in De Limburger/Limburgs Dagblad and for encouraging his family to participate.

For their invaluable support, I want to express my gratitude to Lidwien van den Heuvel and Eefje Dilven. You turned me into another person. *Ik stel mezelf altijd grote doelen, dan mis ik ze niet*, it says on a postcard which I received from my friend and fellow-PhD student at the Classics department Eveline Rutten. Thank you, Eveline, for joining me on the journey. I hope yours will end soon, too. I also found friendship at the English department; Lindsay Bruijn-Janssen, Chris Cusack, and Nicole Verberkt, thanks for being understanding, kind, funny, and out-of-the-box.

Christa and Nienke, thank you very much for being my paranymphs! I am sure that you will be of great support before, during, and after my defense.

Of course, there are also many non-work related people that contributed to my quality of life in the past six years. Thanks go to all my friends and family for understanding my busyness and for making me feel proud of what I do. I hope that I can spend more time with you now that this journey is completed.

And than, finally, there's you, Frank, *leeveke*. Thank you for being you, and for letting me be me.

Contents

1. General Introduction	15
1.1 Tone languages	17
1.2 The case of Limburgian	19
1.3 Research questions and outline of the thesis	22
2. The word prosodic contrast in Limburgian	35
2.1 The different functions of pitch across languages	35
2.2 Word-level pitch in Limburgian	39
2.2.1 The linguistic context in the province of Limburg	40
2.2.2 The dialect of Roermond and its relation to other varieties.....	43
2.2.3 Word-level pitch in Roermond Dutch	46
2.2.4 The tone-intonation interaction in Roermond Dutch	50
2.3 The phonological nature of word-level pitch in Roermond Dutch	52
3. No perceptual reorganization for Limburgian tones? A cross-linguistic investigation with 6- to 12-month-old infants ...	69
3.1 Introduction	70
3.1.1 The case of lexical tone	71
3.1.2 Perceptual reorganization for lexical tones	73
3.1.3 Limburgian dialects of Dutch.....	76
3.1.4 Lexical tone in Roermond Dutch.....	78
3.1.5 The present study	81
3.2 Method	82
3.2.1 Participants.....	83
3.2.2 Stimuli	84
3.2.3 Apparatus	86
3.2.4 Procedure	87
3.2.5 Data pre-processing.....	88
3.3 Results	89
3.4 Discussion	92
4. Distinctive function, distinctive processing? Native and non-native perception of Limburgian lexical tones by Limburgian and Dutch adults	111
4.1 Introduction	112
4.1.1 How different functions of pitch affect pitch processing	113
4.1.2 Theoretical models for the naïve perception of non-native pitch contrasts.....	118

4.1.3	The present study	120
4.2	Method.....	123
4.2.1	Participants.....	123
4.2.2	Stimuli	124
4.2.3	Apparatus and procedure.....	129
4.3	Results	129
4.3.1	Introduction to the data analysis	129
4.3.2	Between-category variation analysis.....	130
4.3.3	Between-category variation vs. within-category variation analysis	134
4.4	Discussion	136
4.4.1	Do Limburgian and Dutch adults differ in their perception of Limburgian lexical tones?.....	136
4.4.2	Is Dutch adults' perception of Limburgian tones influenced by their native prosodic system?.....	140
4.4.3	Remarkable findings pertaining to the Limburgian listeners	146
4.4.4	Directions for future research.....	149
4.5	Conclusion	150
5.	Does native prosody affect pitch processing during word learning? Lexical encoding of tones by Limburgian and Dutch toddlers and adults.....	165
5.1	Introduction	166
5.1.1	Integration of pitch into lexical representations	168
5.1.2	The present study	173
5.2	Experiment 1.....	176
5.2.1	Participants.....	176
5.2.2	Apparatus	178
5.2.3	Procedure	179
5.2.4	Stimuli	184
5.2.5	Data pre-processing and analysis	187
5.2.6	Results.....	189
5.3	Experiment 2	192
5.3.1	Participants.....	192
5.3.2	Apparatus, procedure, and stimuli.....	194
5.3.3	Data pre-processing and analysis	194
5.3.4	Results.....	195
5.4	Discussion	198
5.4.1	The lexical encoding of pitch in Limburgian and Dutch	

toddlers.....	198
5.4.2 The lexical encoding of pitch in Limburgian and Dutch adults	205
5.5 Conclusion.....	206
6. General Discussion.....	225
6.1 Summary.....	225
6.2 Theoretical considerations.....	230
6.2.1 Perceptual reorganization for lexical tones	230
6.2.2 Perceptual assimilation of lexical tones.....	232
6.2.3 Differences between tone languages and their influence on pitch processing.....	236
6.2.4 The influence of variation due to bidialectalism.....	238
6.3 Methodological considerations.....	240
6.3.1 Comparability across studies.....	240
6.3.2 Choosing the distracter object: Familiar or novel?	242
6.3.3 Assessing the linguistic input of bidialectal children.....	243
6.4 Conclusion.....	245
7. Samenvatting in het Nederlands.....	255
7.1 Inleiding.....	255
7.2 Experiment 1: Toonperceptie bij Limburgse en Nederlandse baby's	257
7.3 Experiment 2: Toonperceptie bij Limburgse en Nederlandse volwassenen	258
7.4 Experiment 3: De lexicale representatie van toon in Limburgse en Nederlandse peuters en volwassenen	261
7.5 Conclusie.....	265
Curriculum Vitae	267

1. General Introduction

Learning to efficiently perceive the sounds of a language is a complex process that lies at the heart of language acquisition. This complexity emerges from the fact that the speech signal contains an incredible amount of variation, but not all this variation is linguistically meaningful. For example, in many Limburgian dialects spoken in the south of the Netherlands, a word like *zeeve* [ze:və] ('seven') can be pronounced with a mean pitch of 350Hz (spoken by a female speaker) or with a mean pitch of 220Hz (spoken by a male speaker). These differences in mean pitch do not lead to different words and are therefore called PHONETIC. However, it does matter whether the first syllable in *zeeve* is pronounced with level pitch or with falling pitch. With level pitch, *zeeve* means 'seven', but with falling pitch, it means 'to sift'. In these Limburgian dialects, differences in pitch contours are thus linguistically meaningful (i.e., PHONEMIC), in that they are used to distinguish words.

The challenge for children is to learn to separate variation that has an impact on lexical meaning (i.e., phonemic variation) from variation that does not (i.e., phonetic variation). In light of this complex task, it is astonishing that infants seem to have learned which phonetic contrasts are relevant for distinguishing meaning in their native language even before producing their first words (for a review, see Werker & Tees, 1992). In particular, between 4 and 12 months of age, infants begin to attend to sound patterns relevant to their native language structure. Native vowel perception appears to be in place between 6 and 8 months of age (e.g., Polka & Werker, 1994; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Tsuji & Cristiá, 2014). For example, Polka and Werker (1994) tested English-learning infants for their ability to discriminate native and non-native vowel contrasts. The infants were tested on two non-native German contrasts, which are not phonemic in English (i.e., /u/-/y/ and /u/-/ʏ/), and on a native English phonemic control contrast (i.e., /i/-/a/). Whereas 4-month-olds discriminated both the native English and the non-

16 Setting the Tone

native German vowel contrasts, 6- to 8-month-olds only discriminated the native control contrast, showing a decline in the perception of non-native vowel contrasts between 4 and 6-8 months of age. Moreover, the ability to perceive non-native vowel contrasts continues to decline between 6-8 months and 10-12 months. Attenuation of non-native consonant contrasts appears later, at around 10 to 12 months of age (e.g., Kuhl et al., 2006; Tsushima et al., 1994). This change or relocation of attention with respect to sound contrasts has been called PERCEPTUAL TUNING or PERCEPTUAL REORGANIZATION (henceforth PR). During their first year of life, infants shift from language-general to language-specific perception. By their first birthday, they seem to know which sounds are part of their native language.

However, acquiring a language not only involves the capacity to accurately perceive phonetic and phonemic contrasts. It also entails the ability to assign appropriate interpretations to different sorts of phonetic variation. Some contrasts are associated to representations of words, stored in our long-term memory, and some are not. A plethora of research has explored the developmental perception of phonetic and phonemic contrasts in the first year of life and the way they are processed during word learning and recognition at later ages (e.g., Jusczyk, 1997; Kuhl, 2004; Stager & Werker, 1997; Swingley & Aslin, 2000; White & Morgan, 2008). This research has mainly focused on segmental contrasts.

More recently, scholars have come to acknowledge that the existing body of research does not reflect the fact that many of the world's languages – according to some estimates approximately 60-70% (Yip, 2002) – use LEXICAL TONE. In tone languages, pitch patterns are used to distinguish words (recall that *zeeve* in Limburgian can represent two different words depending on the pitch contour of the first syllable). This is similar to the function of phonemes at the segmental level. Probably the best-known example of a tone language is Mandarin Chinese. In Mandarin, the syllable /ma/ can represent four words, depending on its tonal make-up. With high level pitch, it means 'mother', with rising pitch, it means 'hemp', with low dipping pitch, it means 'horse', and with

high falling pitch, it means ‘scold’. Research on tone languages is essential to inform and advance the field of first language acquisition, since existing theories of language development do not necessarily generalize to the world’s population (Singh & Fu, 2016).

1.1 Tone languages

Psycholinguistic research on the acquisition and processing of lexical tones has typically investigated well-known tone languages such as Mandarin, Cantonese, or Thai (for an extensive review of tone acquisition studies, see Singh & Fu, 2016). However, within the family of tone languages, large differences exist. It remains unclear how these differences might impact upon the acquisition and processing of lexical tone.

First, tone languages differ with respect to the functional load of tone, which has been proposed to depend on the tonal inventory (i.e., the number of tones, and, related to this, their information value), the distributional restrictions of tones, or ‘tonal density’ (Gussenhoven, 2004a; i.e., the percentage of syllables that require a tone feature), the importance of tones for lexical disambiguation (i.e., the number of words that are distinguished solely by lexical tone), and the extent to which f_0 is the only cue to the tonal distinction (i.e., do duration or voice quality play a role?) (e.g., Kristoffersen, 2000; Pierrehumbert & Beckman, 1988; Tong, Francis, & Gandour, 2008; Wang, Bene, Jongman, & Sereno, 2004; Wu, Tu, & Wang, 2012). If tone languages were mapped on a scale according to their tonal density, Mandarin Chinese would be on one end of the continuum because almost every syllable, except those with so-called neutral tone, carries one of four lexical tones (Duanmu, 2000). In Swedish, Tokyo Japanese, or Limburgian, however, the distribution of lexical tones is more restricted, depending on syllable characteristics (e.g., stress and syllable weight). Typically, the lower the functional load, the more restricted a tone system is considered to be (Voorhoeve, 1973). The functional load of word-level pitch patterns has been assumed to influence sensitivity to word-

18 Setting the Tone

level pitch in adult speakers of these languages. The higher the functional load, the more sensitive speakers are to word-level pitch contrasts (e.g., Goss, 2015; Schaefer & Darcy, 2014; Wang et al., 2004; Wu et al., 2012).

A second difference within the family of tone languages lies in the complexity of the post-lexical use of pitch, called INTONATION. Intonation is found in all of the world's languages (Gussenhoven, 2004a; Yip, 2002) and is used at the level of utterances, for example to signal questions and statements. Typically, tone languages do not have complex intonation systems and, as a consequence, the pronunciation of a word with a certain lexical tone does not greatly differ depending on whether it is part of an interrogative or declarative utterance. In Standard Chinese, for example, different intonations only cause changes in pitch register, not in pitch contours (Wu, 2000). However, some more restricted tone systems, like Swedish and Limburgian, do have complex intonation systems. In these languages, intonation tones interact with lexical tones, causing many different surface realizations (i.e., contours) of a lexical tone (e.g., Bruce, 1977; Gussenhoven, 2000). If a learner is confronted with many surface realizations, he will need to abstract away from all these variations in order to recover the underlying categories. The consistency of the relationship between underlying representations and surface forms appears to delay the acquisition of phonetic features, including pitch contours (e.g., Demuth, 1995; Ota, 2003; Quam & Swingley, 2014).

In summary, there is limited research dedicated to the acquisition of lexical tones in more *restricted* tone systems. The few existing studies looked at perception (e.g., Japanese infants: Sato, Sogabe, & Mazuka, 2010) or production of lexical tones (e.g., Japanese and Swedish children: Ota, 2003, 2006; Swedish children: Plunkett & Strömqvist, 1992, and references therein; Swedish children and adults: Schmid, 1986).

1.2 The case of Limburgian

This thesis seeks to contribute to the field of lexical tone acquisition and processing by investigating an understudied group of restricted tone languages: Limburgian dialects of Dutch. The Limburgian dialects of Dutch belong to the Central Franconian dialect continuum, which covers the provinces of Limburg in the Netherlands and Belgium as well as the north of the German Rhineland-Palatinate and the southwest of North-Rhine Westphalia. We use the term dialect to refer to a regional linguistic variety that differs from the standard language at the phonological, morphosyntactic and lexical level (O’Grady, Archibald, Aronoff, & Rees-Miller, 2001).

As in many varieties of North Germanic (Norwegian, Swedish, and some variants of Danish), 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, also called *Stoßton*, ‘pushtone’) means ‘hare’, whereas *haas* with falling-rising pitch (accent 2, also called *Schleifton*, ‘dragging tone’) 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: *knien* [kni:n] with accent 2 means ‘rabbit’, but pronounced with accent 1 it means ‘rabbits’ (see Figures 1A and 1B).

20 Setting the Tone

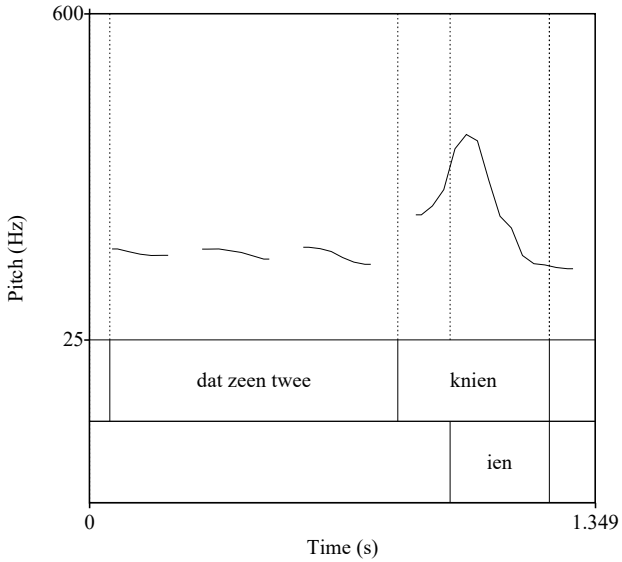


Figure 1A. F0 contour of the Limburgian sentence *dat zeen twee KNIEN* 'those are two rabbits'. The rhyme of the target word carries accent 1.

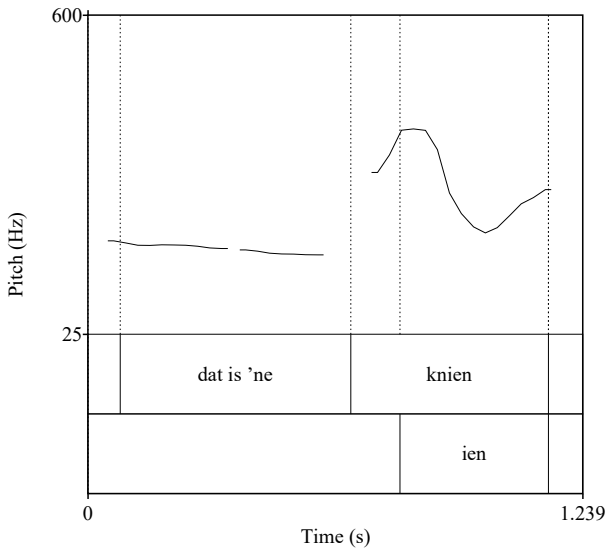


Figure 1B. F0 contour of the Limburgian sentence *dat is 'ne KNIEN* 'that is a rabbit'. The rhyme of the target word carries accent 2.

Studying the acquisition and processing of lexical tone in Limburgian¹ is interesting for two main reasons. First, Limburgian can be considered restricted with respect to the functional load of tone. Relative to Mandarin Chinese, there are few tonal minimal pairs. Moreover, there is only a two-way contrast, which has been described as privative. According to Gussenhoven and Peters (2008, p. 88) “the word accent contrast (...) amounts to a contrast between the absence of lexical tone (accent 1) and its presence (accent 2)”. Also, the distribution of tones is subject to syllabic restrictions. In Dutch Limburgian dialects and in the dialect of Cologne, the contrast can only be realized on minimally bimoraic syllables with primary stress, meaning that an unbound multisyllabic morpheme can only carry one accent. With respect to the functional load of tone, we assume Limburgian to be comparable to languages like Swedish (Gussenhoven, 2004b; Riad, 2013), Norwegian (Kristoffersen, 2000; Steien & van Dommelen, 2016; Wetterlin, 2007) and Japanese (Kubozono, 1993).

A second reason why it is interesting to study Limburgian is that it shows a complex intonation system, which is rare for a language that also has lexical tone (Gussenhoven & van der Vliet, 1999). As in most other Franconian dialects (Köhnlein, 2016), and comparable to Swedish (Bruce, 1977), Limburgian lexical tones co-occur with intonation tones in syllables with main stress. The tone-intonation interaction in Limburgian is more complex than in Swedish due to the larger number of intonation patterns (Gussenhoven, 2004b). The occurrence of different intonation contours causes the Limburgian lexical tone contrast to show up in very different shapes (Gussenhoven, 1999). In *zeeve* (‘seven’), accent 2 appears in pre-final position and is realized as a level tone, whereas in *knien* (‘rabbit’), it is in final position and pronounced as a

¹ Henceforth, the term Limburgian is used to refer to those Limburgian dialects that have lexical tone. Note that Limburgian is usually an umbrella term for all dialects spoken in the Limburg region. Many of those dialects have lexical tone, but some do not. The linguistic situation in Limburg will be discussed in more detail in Chapter 2.

22 Setting the Tone

falling-rising contour, while the pronunciation with interrogative intonation is different again. In this way, every pitch contour represents a unique combination of intonational and lexical meaning. This is different from Standard Chinese, where different intonations cause changes of register rather than of contours (Wu, 2000). As a result of the large amount of surface variation in the realization of lexical tones, the Limburgian tone system can be considered rather opaque.

Due to this opacity and the relatively low functional load, the Limburgian tone system might pose challenges to its learners. By studying the acquisition and processing of lexical tone in Limburgian, this thesis seeks to contribute to our understanding of the influence of the functional load of tone and the complexity of the prosodic system on the acquisition and processing of lexical tone in particular and the native prosodic system in general. Throughout this thesis, participants acquiring and/or speaking Limburgian (mostly next to Dutch, see Chapter 2 for elaboration) will be compared to a control group of non-tonal, monolingual Dutch speakers to see whether differences in the functionality of pitch in tonal Limburgian and non-tonal Dutch bring about differences in pitch processing. In the following, note that *Limburgian* and *Dutch* are used to refer to the linguistic background of our participants rather than to their nationality, since all our participants were Dutch in terms of their nationality.

The research questions as well as the thesis outline will be discussed in the next section.

1.3 Research questions and outline of the thesis

This thesis aims to answer the following main research questions:

- 1a. How does Limburgian infants' lexical tone perception develop in the first year of life?
- 1b. Does Limburgian infants' perception of lexical tones develop differently from that of infants acquiring non-tonal Dutch?

- 2a. Are Limburgian adults sensitive to word-level pitch?
- 2b. Do Limburgian and Dutch adults differ in their sensitivity to word-level pitch?
- 2c. Is Dutch listeners' perception of Limburgian lexical tones influenced by their native intonation system?
- 3a. Do child and adult speakers of Limburgian encode pitch information as part of their lexical entries when learning novel words?
- 3b. Do child and adult speakers of Limburgian behave differently from Dutch peers with respect to their sensitivity to pitch in a word learning context?

These research questions will be answered in three experimental chapters summarized below. First, **Chapter 2** provides the basic theoretical background to the experimental studies in chapters 3-5. It first gives an introduction into the different uses of pitch in the world's languages (2.1). Subsequently, it provides an extensive discussion of the complex (socio)linguistic context in the Dutch province of Limburg that our target population is exposed to (2.2.1, 2.2.2), and finally it provides a discussion on the use of word- and utterance-level pitch in Limburgian (2.2.3, 2.2.4), and on the nature of phonological representations of word-level pitch (2.3).

Chapters 3 to 5 then address how infants and children growing up with Limburgian as well as adult speakers of Limburgian perceive and interpret the tonal make-up of words by reporting a series of behavioural experiments.

In **Chapter 3**, discrimination of the Limburgian tones in 6- to 12-month-old Limburgian and Dutch infants is tested using a variant of the visual habituation procedure (Horowitz, Paden, Bhana, & Self, 1972; Houston, Horn, Qi, Ting, & Gao, 2007; Stager & Werker, 1997). The aims of this study are twofold. First, we want to provide insights into the development of lexical pitch perception in Limburgian infants growing up with a restricted tone language. Secondly, we compare Limburgian infants to Dutch infants learning a non-tonal language that uses post-lexical intonation. Given the differences in the

24 Setting the Tone

functionality of pitch in particular and the differences in complexity of the prosodic systems in general (see Chapter 2), we investigate the assumption that Limburgian children's sensitivity to lexical pitch develops differently from that of Dutch children. In accordance with previous results on native tone and pitch-accent perception in the first year of life (Sato et al., 2010; Tsao, 2017; Yeung, Chen, & Werker, 2013), we expect that 6- to 12-month-old native Limburgian infants successfully discriminate the Limburgian tones throughout their first year of life. For the Dutch infants, two outcomes are possible. On the one hand, they could show an age-related performance decline (as the non-tone language listeners in Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; and Yeung et al., 2013). On the other hand, Dutch infants could maintain their discrimination of the Limburgian tones, as the Dutch infants in Liu and Kager (2014) and Chen and Kager (2016) also showed persistent sensitivity to certain Chinese contrasts. Our results suggest that both Limburgian and Dutch infants are sensitive to the Limburgian tonal contrast throughout their first year of life. We propose that Dutch infants might perceive the Limburgian tones as native intonation patterns (i.e., statement vs. question).

The study reported in **Chapter 4** expands on the findings reported in Chapter 3 and investigates whether the differences in the functionality of pitch in Limburgian and Dutch cause differences in lexical tone discrimination in adult speakers of these languages. It also seeks to explain the successful discrimination observed in the Dutch infants in Chapter 3. Is the Dutch listener's sensitivity to non-native lexical tones influenced by their native intonation system? Moreover, testing adults can reveal whether the infants' behavior reflects the adult prosodic system. To investigate adults' tone discrimination ability, a categorical AXB-discrimination task was run with Limburgian and Dutch adults. Participants heard sequences of three stimuli and had to judge whether the second token (X) most resembled the first (A) or third (B) token. As expected based on the different functions of pitch in the languages under investigation, Limburgian adults outperformed Dutch adults in general by showing better discrimination of Limburgian lexical tones. Nevertheless,

Dutch participants performed above chance, thus showing persistent sensitivity to Limburgian tones.

To investigate the possibility that Dutch listener's perception of Limburgian tones is influenced by native intonation categories, we manipulated the position of the tones in our pseudo-word stimuli. In utterance-final position, Dutch listeners could perceive the tones as instances of native post-lexical intonation categories (i.e., statement vs. question intonation), but this is not as easily achieved in non-final position (e.g., Braun & Johnson, 2011; see Chapter 4 for more details). As a consequence, we expected Dutch listeners to show poorer discrimination performance in non-final contexts than in final contexts. Our findings, however, showed the opposite pattern, which might be explained by the greater acoustic salience of the Limburgian tone contrast in non-final position.

To investigate differences in the sensitivity to word-level pitch in Limburgian and Dutch adults more closely, we also asked whether word-level pitch perception in Limburgians is categorical. If their perception is driven by native lexical tone categories, Limburgian listeners are expected to be worse at perceiving within-category pitch differences (e.g., the difference between two tokens with accent 1) than between-category differences (e.g., the difference between a token with accent 1 and a token with accent 2). To efficiently categorize f_0 contours into tones, tone language speakers must ignore irrelevant (i.e., within-category) tonal variations (e.g., Hallé, Chang, & Best, 2004; Xu, Gandour, & Francis, 2006). In other words, they are assumed to perceive the incoming signal through a native category filter. However, if no lexical tone categories are distinguished in the native phonological system, as is the case in Dutch, there are no categories that could influence perception. Dutch listeners should thus perceive all pitch patterns as equally (un)important melodic variations, leading to similar performance in within- and between-category discrimination. This would be in line with the non-tone language

speakers in Hallé et al. (2004), Hoffmann, Sadakata, Chen, Desain, and McQueen (2014) and Xu et al., (2006).

Surprisingly, however, our data showed that Limburgian adults performed equally well in between- and within-category discrimination. We suggest that this could be explained by the Limburgians' daily exposure to the many surface realizations of the Limburgian tones as a consequence of the tone-intonation interaction. Their experience with a rich prosodic system might enhance their sensitivity to pitch in general. As opposed to the Limburgians' performance, the performance of our Dutch listeners was *not* constant across within- and between-category variation. Dutch participants performed better in within-category discrimination than between-category discrimination. We propose that the Dutch listeners might have been using *different* listening strategies in between- and within-category variation triads due to particular characteristics of the stimuli.

After Chapters 3 and 4, the question no longer was whether speakers of Limburgian and Dutch are able to perceive Limburgian tones, but rather whether they *interpret* these word-level pitch variations as lexically relevant. To address this question, in **Chapter 5** the influence of word-level pitch on word learning and word recognition is tested in 2,5- to 4-year-old children and adults. When acquiring a lexicon, tone-language learners need to learn to ascribe lexical relevance to pitch changes and store tones as part of their word entries in the mental lexicon. Conversely, non-tone-language learners have to learn to disregard pitch changes that occur within words, despite the fact that they might still discriminate these pitch changes in a purely perceptual task. Recent work suggests that child and adult speakers of tone languages behave differently from non-tone language speakers in exploiting contrastive pitch contours when learning words. In accordance with native prosody, tone language speakers attend to pitch information and exploit it during lexical access, whereas non-tone language speakers (mostly) do not, or at least to a lesser extent (Braun, Galts, & Kabak, 2014; Hay, Graf Estes, Wang, & Saffran, 2015; Quam & Swingley, 2010; Singh, Hui, Chan, & Golinkoff, 2014). Note,

however, that none of the existing studies tested the influence of word-level pitch during word learning and recognition in more restricted tone systems. Following the procedure of Quam and Swingley (2010) and Singh et al. (2014), Chapter 5 explores the sensitivity to mispronunciations involving pitch changes in newly learned words in Limburgian and Dutch children and adults. By studying the influence of word-level pitch on word learning, we hope to be able to shed more light onto the lexical representations of the Limburgian tones. As a consequence of the different functions of pitch in the languages under investigation in this thesis, Limburgians were expected to behave differently from Dutch participants. We expected that the Limburgians' recognition of newly learned words would be hindered, at least to some extent, by a change in pitch. Dutch participants were expected to ignore pitch changes, as these are not lexically relevant in Dutch. Strikingly, both Limburgian and Dutch toddlers showed sensitivity to tonal mispronunciations. Adult participants showed very strong naming effects both upon hearing correct pronunciations and mispronunciations, preventing us to draw any conclusions about their sensitivity to pitch changes within words.

The last chapter of this thesis, **Chapter 6**, provides a summary of the conclusions, a discussion of the findings and suggestions for future research.

References

- Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Language Research*, 30, 323-350.
- Braun, B., & Johnson, E. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, 39, 585-594.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, 25(5), 426-439.
- Demuth, K. 1995. The acquisition of tonal systems. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 111-134). Hillsdale: Lawrence Erlbaum.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
- Goss, S. (2015). *The effects of internal and experience-based factors on the perception of lexical pitch accent by native and nonnative Japanese listeners*. PhD Dissertation, Ohio State University.
- Gussenhoven, C. (1999). Discreteness and gradience in intonational contrasts. *Language and Speech*, 42(2-3), 283-305.
- Gussenhoven, C. (2000). The lexical tone contrast of Roermond Dutch in Optimality Theory. In M. Horne (Ed.), *Prosody: Theory and Experiment* (pp. 129-167). Amsterdam: Kluwer.
- Gussenhoven, C. (2004a). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

- Gussenhoven, C. (2004b). Tone in Germanic: Comparing Limburgian with Swedish. In G. Fant, H. Fujisaki, J. Cao, & Y. Xu (Eds.), *From Traditional Phonology to Modern Speech Processing* (pp. 129-136). Beijing: Foreign Language Teaching and Research Press.
- Gussenhoven, C., & Peters, J. (2008). De tonen van het Limburgs. *Nederlandse Taalkunde*, 13(1), 87-114.
- Gussenhoven, C. & van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.
- Hallé, P. Chang, Y.-C. & Best, C. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- 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(1), 10-22.
- Hoffmann, C., Sadakata, M., Chen, A., Desain, P., & McQueen, J. (2014). Within-Category Variance and Lexical Tone Discrimination in Native and Non-Native Speakers. *Proceedings TAL 2014*, 45- 49.
- Horowitz, F. D., Paden, L., Bhana, K., & Self, P. (1972). An infant control procedure for studying infant visual fixations. *Developmental Psychology*, 7, 90.
- Houston, D., Horn, D., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, 12(2), 119-145.
- Jusczyk, P. (1997). *The discovery of spoken language*. MIT Press.
- Köhnlein, B. (2016). Contrastive foot structure in Franconian tone-accent dialects. *Phonology*, 33, 87-123.

30 Setting the Tone

- Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.
- Kubozono, H. (1993). *The organization of Japanese prosody*. Tokyo: Kurosio.
- Kuhl, P. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews/ Neuroscience*, 5, 831-843.
- Kuhl, P., Stevens, E., Hayashi, A., Degushi, 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.
- Kuhl, P., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by six months of age. *Science*, 255, 606-608.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133, 385-394.
- Mattock, K., & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization. *Infancy*, 10, 241-265.
- 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.
- O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (Eds.). (2001). *Contemporary Linguistics*. Boston: Bedford/St. Martin's.
- Ota, M. (2003). The development of lexical pitch accent systems. *Canadian Journal of Linguistics*, 48(3/4), 357-383.
- Ota, M. (2006). Children's production of word accents in Swedish revisited. *Phonetica*, 63, 230-246.

- Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Linguistic inquiry monograph 15. Cambridge: MIT Press.
- Plunkett, K., & Strömquist, S. (1992). The acquisition of Scandinavian Languages. In D. Slobin (Ed.), *The crosslinguistic study of the language acquisition 3* (pp. 457-556). New Jersey/London: Lawrence Erlbaum.
- 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.
- Quam, C., & Swingley, 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.
- Quam, C., & Swingley, D. (2014). Processing of lexical stress cues by young children. *Journal of Experimental Child Psychology*, 123, 73-89.
- Riad, T. (2013). *The phonology of Swedish*. Oxford University Press.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, 22(11), 2503-2513.
- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, 5(4), 489-522.
- Schmid, B. (1986). A comparative study of children's and adult's acquisition of tone accents in Swedish. *Language Learning*, 36, 185-210.
- Singh, S., & Fu, C. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, 87(3), 834-854. doi:10.1111/cdev.12512.

32 Setting the Tone

- Singh, L., Hui, T.J., Chan, C., & Michnick Golinkoff, R. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, *17*(1), 94-109.
- Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381-382.
- Steien, G.B., & van Dommelen, W.A. (2016). The production of Norwegian tones by multilingual non-native speakers. *International Journal of Bilingualism*. doi: 10.1177/1367006916673218.
- Swingley, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.
- Tong, Y., Francis, A., & Gandour, J. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, *23*(5), 689-708.
- Tsao, F.-M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in Psychology*, doi:10.3389/fpsyg.2017.00558.
- Tsuji, S., & Cristiá, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, *56*(2), 179-191. doi:10.1002/dev.21179.
- Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C.T. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: language-specific developmental changes in speech perception abilities. *1994 International Conference on Spoken Language Processing* (pp. 1695-1698). Yokohama, Japan.
- Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*, *4*(1), 1-22.

- Wang, Y., Behne, D., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, 25, 449-466.
- Werker, J.F., & Tees, R.C. (1992). The organization and re-organization of human speech perception. *Annual Review of Neuroscience*, 15, 377-402.
- Wetterlin, A. (2007). *The Lexical Specification of Norwegian Tonal Word Accents*. Doctoral dissertation, University of Konstanz.
- White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59, 114-132.
- Wu, Z. (2000). From traditional Chinese phonology to modern speech technology: Realization of tone and intonation in spoken Chinese. *Proceedings of the International Conference on Spoken Language Processing (ICSLP 2000)*, Vol. 1, B1-B12.
- Wu, X., Tu, J.-Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623-641.
- Xu, Y., Gandour, J., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch contrasts. *Journal of the Acoustical Society of America*, 120(2), 1063-1074.
- Yeung, H.H., Chen, K.H., & Werker, J. (2013.). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

2. The word prosodic contrast in Limburgian

2.1 The different functions of pitch across languages

In the human perception of speech, acoustic differences can be perceived as differences in pitch. The acoustic measure that most closely corresponds to this perceptual phenomenon is fundamental frequency, or f_0 , which in turn reflects the frequency of vibration of the vocal folds. Languages vary in the functions they attribute to pitch differences. Linguists distinguish between two of the major uses of linguistic pitch as TONE and INTONATION.

Intonation is found in all of the world's languages (Gussenhoven, 2004a; Singh & Fu, 2016; Yip, 2002). It is used to distinguish sentence types, like statements versus questions, to indicate whether a speaker has finished or intends to continue speaking, or to show which parts of an utterance present new or important information. Intonation is also referred to as POST-LEXICAL use of pitch because pitch information is not part of a word's representation in the mental lexicon.

Tone is the term used to describe the use of pitch patterns to distinguish words, similar to the function of phonemes at the segmental level. Accordingly, this is LEXICAL use of pitch, as pitch information is part of a word's mental representation. By some estimates, the number of languages using tone is as high as 60-70% (Yip, 2002). These languages are spoken by more than 50% of the world's population (Fromkin, 1978). Languages that use pitch to distinguish words are called TONE LANGUAGES. Tone languages are widespread in Africa, South-East Asia and Central- and South-America. In Europe, Australia, and the Pacific, they are particularly scarce or absent (See Figure 1).

36 Setting the Tone

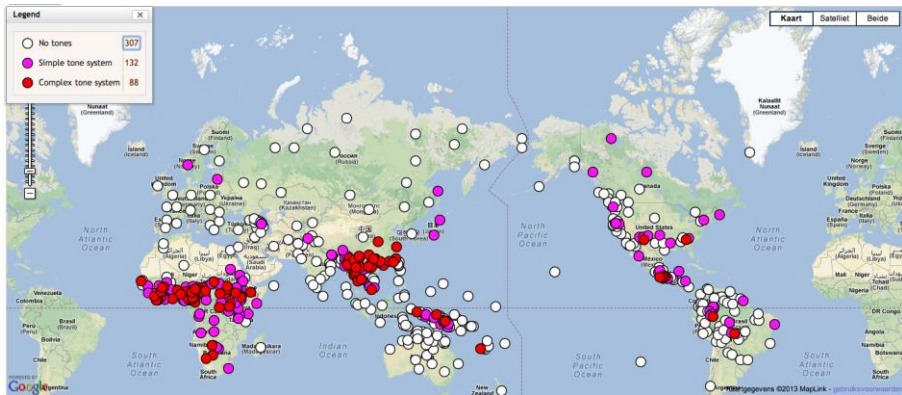


Figure 1. Distribution of tone languages according to the World Atlas of Language Structures Online. Retrieved May 3, 2013, from <http://wals.info/feature/13A#2/19.3/152.9>

Probably the best-known example of a tone language is Mandarin Chinese. In Mandarin, the syllable /ma/ can represent four words, depending on its tonal make-up. With high level pitch, it means ‘mother’, with rising pitch, it means ‘hemp’, with low dipping pitch, it means ‘horse’, and with high falling pitch, it means ‘scold’.

Some tone languages make very extensive use of lexical tone. In Mandarin Chinese, a very dense tone language, almost every syllable has to be pronounced with one of four tones (Duanmu, 2000). Other tone languages are more restricted in their use of lexically contrastive pitch. These languages, for example Tokyo Japanese, have been referred to as RESTRICTED TONE LANGUAGES (Voorhoeve, 1973). Whereas from a typological point of view the phonetic and functional distinction between *tone* languages and *intonation* languages is rather unproblematic, the distinction between tone languages and *restricted* tone languages is not clear-cut (e.g., Gussenhoven, 2004a).

Nevertheless, many phonologists have recognized an intermediate type of language with word level prosody that shares properties of tone languages and languages with word stress, referred to as a PITCH-ACCENT LANGUAGE. According to Yip (2002, p. 257), there is no clear distinction between tone and pitch-accent languages. Pitch-accent languages are best considered a subtype of

tone languages. Hyman (2009) too argues against pitch-accent languages as a third prosodic type of language next to languages with tone and languages with stress. Hyman states that languages that appear to be somewhere in between tone languages and stress languages (e.g., Tokyo Japanese, Somali, Swedish, Limburgian) do not form a coherent class. For a typological distinction to be valid, there should be a set of features or some prototypicality that is distinctive of languages that fall into neither the tone group nor the stress group. Until today, no such criterion has been formulated. Werth (2011, p. 245) lists a number of properties that tone languages and languages that often have been called pitch-accent languages share, again questioning the distinction between the two. In both language types (1) phonological distinctiveness can be signaled by tone, (2) tone contrasts can be equipollent or privative², (3) the tone bearing units are moras or syllables, (4) tone contrasts can be subject to context-dependent neutralization, (5) the distribution of tones can be subject to syllable structure constraints, (6) tone can co-exist with stress, and (7) the functional load can be low due to few minimal pairs. All languages that have been called pitch-accent languages can be analyzed as tonal, since all these languages have in common that pitch, be it to a greater or to a lesser extent, is necessary for determining the identity of a word (Hyman, 2009; Werth, 2011). In this thesis, we will take the term ‘tone language’ to refer to languages that use pitch to signal word-level meaning differences, unless when referring to studies that used a different term.

Despite the fact that we consider all languages that use pitch as a distinctive feature in at least a few words to be tone languages, we still have to recognize that, within the family of tone languages, large differences exist. First, tone languages differ with respect to the functional load – or usefulness – of tone. Although intuitively the meaning of the concept of functional load appears

² A contrast is equipollent if the two terms in a binary contrast are both underlyingly specified (e.g., H vs. L). It is privative if only one of these is underlyingly specified (e.g., H vs. no tone) (e.g., van der Hulst, 1999).

to be shared in the literature, the exact definition of functional load differs greatly across studies (e.g., Oh, Coupé, Marsico, & Pellegrino, 2015; van Severen et al., 2013, and references therein). The functional load of tone has been proposed to depend on the tonal inventory (i.e., the number of tones, and, related to this, their information value), the distributional restrictions of tones or 'tonal density' (Gussenhoven, 2004a; i.e., the percentage of syllables that require a tone feature), the importance of tones for lexical disambiguation (i.e., how many tonal minimal pairs are there in the language?), and the extent to which f_0 is the only cue to the tonal distinction (i.e., do duration or voice quality play a role?) (e.g., Kristoffersen, 2000; Pierrehumbert & Beckman, 1988; Tong, Francis, & Gandour, 2008; Wang, Bene, Jongman, & Sereno, 2004; Wu, Tu, & Wang, 2012). The smaller the inventory, the greater the distributional restrictions and the smaller the number of tonal minimal pairs, the more restricted a tone system is considered to be (Voorhoeve, 1973).

The functional load of word-level pitch patterns in the native language (L1) has been assumed to influence sensitivity to word-level pitch in speakers of these languages (e.g., Goss, 2015; Schaefer & Darcy, 2014; Wang et al., 2004; Wu et al., 2012). The scale between more or less restricted tone systems based on the functional load of lexical pitch thus appears to be valid on the basis of experimental evidence. However, to date there exist hardly any cross-linguistic investigations comparing word-level pitch sensitivity in speakers of the many different kinds of tone languages to pitch processing in speakers of languages without lexical tone. The existing cross-linguistic studies on tone perception mostly involved Asian tone languages and English as the non-tonal language (for a review on tone acquisition studies, see Singh & Fu, 2016).

To extend our insights into the influences of different prosodic systems on pitch processing and its development, this thesis looks at acquisition and processing of lexical tone in infant, child and adult speakers of a group of tone languages spoken in the south of the Netherlands: Limburgian dialects of Dutch. In terms of the functional load of tone these can be considered restricted tone systems (see section 2.2.3).

Interestingly, the Limburgian dialects show complex intonation systems, which is rare for languages that also have lexical tone (Gussenhoven & van der Vliet, 1999). Lexical tones co-occur with intonation tones in syllables with main stress, causing the lexical tone contrast to show up in very different shapes (Gussenhoven, 1999; see section 2.2.4). Unlike for example Mandarin Chinese, any legitimate monosyllabic pitch contour simultaneously represents discursual and lexical meanings. Due to this intricate tone-intonation interaction, the Limburgian prosodic system can be argued to be of greater complexity than that of Mandarin Chinese, but also than that of closely related Standard Dutch, a system with intonation contrasts only. This difference in complexity between Limburgian and Standard Dutch has been shown to determine how pitch processing is differently distributed over the brain in Limburgian and Dutch listeners, supporting the view of function-driven, language-dependent pitch processing (Fournier, Gussenhoven, Jensen, & Hagoort, 2010; see also Gandour, 2007).

Given the differences in complexity of the prosodic systems, we expected not only behavioural differences in adult Limburgian and Dutch speakers in tone perception experiments, we also wanted to explore the possibility that Limburgian children's sensitivity to prosody develops differently from that of Dutch children.

2.2 Word-level pitch in Limburgian

In this section, we offer a description of the language under investigation. First, we will present some important facts about Limburgian in general, which are crucial for understanding some of the choices we have made in the course of our research. Secondly, we will discuss the specific feature of Limburgian that we are interested in: the word prosodic contrast, its different realizations due to the interaction with utterance-level prosody, and its lexical representation.

2.2.1 The linguistic context in the province of Limburg

The Limburgian dialects of Dutch belong to the Central Franconian dialect-continuum, which covers the provinces of Limburg in the Netherlands and Belgium as well as the north of the German Rhineland-Palatinate and the southwest of North-Rhine Westphalia (see Figure 2).

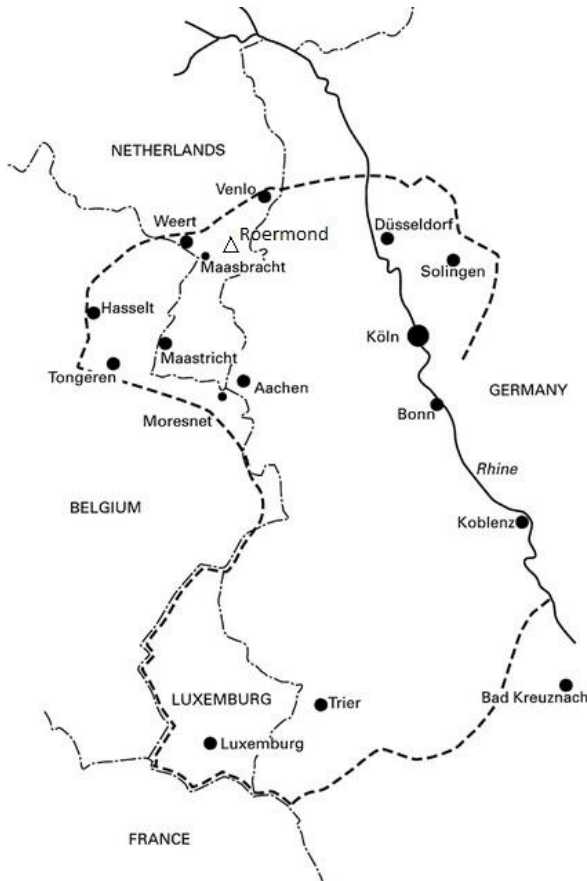


Figure 2. Geographical distribution of the lexical tone contrast in the Cologne-Trier area. Adapted from Gussenhoven & Bruce (1999).

The Dutch province of Limburg has about 1.1 million inhabitants (www.cbs.nl), 75% of which speak a Limburgian dialect (Driessen, 2006). We

use the term dialect to refer to a regional linguistic variety that differs from the standard language, in this case Standard Dutch, at the phonological, morphosyntactic and lexical level (O'Grady, Archibald, Aronoff, & Rees-Miller, 2001). The probably most striking difference between Limburgian and Dutch is the fact that many Limburgian dialects use word-level pitch to distinguish between words, whereas pitch is not lexically contrastive in Dutch. Pitch is used in both languages as a cue to word stress and in post-lexical intonation (e.g. Gussenhoven, 1988; Gussenhoven & van der Vliet, 1999).

Still, mutual intelligibility between Standard Dutch and Limburgian is fairly high (van Bezooijen & van den Berg, 1999a, 1999b). This is probably due to the existence of many cognates, leading to a high degree of phonological overlap. Inhabitants from the provinces in the western and central part of the Netherlands (i.e., Zuid-Holland, Noord-Holland, and Utrecht), who have experience with accented Dutch but are assumed to have little experience with dialects, showed good understanding in a series of translation tasks (van Bezooijen & van den Berg, 1999a, 1999b). Their task was to listen to Limburgian fragments containing target nouns referring to concrete, common objects and translate the target nouns into Dutch. Depending on the amount of context provided with the target noun, participants translated 80% of the target nouns correctly in a low context condition (one target noun, no other content words) and 86% of the targets correctly in a condition with more context, allowing top-down information to be combined with bottom-up information.

Speaking a Limburgian dialect is not evaluated as a characteristic of lower social status. The dialects are spoken by virtually all native speakers, from manual workers to university graduates, in both formal and informal contexts (Cornips, de Rooij, Stengs, & Thissen, 2016; Driessen, 2006). Speaking a Limburgian dialect is an expression of regional or local loyalty (Cornips, 2013).

42 Setting the Tone

Most Limburgians also speak Standard Dutch, the official language used in many formal and institutional settings. According to research conducted by the newspaper *De Limburger/Limburgs Dagblad* of a representative sample of 1.078 respondents in spring 2016, Limburgian is spoken most with one's own partner (64%) or children (62%) at home, with parents (66%), and with friends (71%). Dutch is used most at work or at school (53%), in civil services (65%), and in the hospital (75%). The majority of the Limburgians can thus be considered bidialectal (Cornips, 2014). Bidialectalism can be understood as bilingualism involving closely related linguistic varieties, where an indigenous variety operates alongside more widespread norms in a community of speakers. According to Cornips and Hulk (2006, p. 356) "bidialectalism has increased so much that monolingual speakers of nonstandard dialects have become the exception." That is, children often acquire a local dialect in addition to the standard language. Limburgian children may be immersed in one particular dialect from birth and encounter Standard Dutch upon entering the educational system as well as through media and peers, or alternatively one parent speaks a dialect and the other parent either speaks Dutch (possibly with a Limburgian accent) or a different Limburgian dialect.

Much like bilinguals, bidialectal speakers must be able to monitor who can and who cannot be addressed in their dialect, choose appropriate articulatory settings and inhibit competing phonetic and lexical variants of the language variety not currently in use (Kirk, Fiala, Scott-Brown, & Kempe, 2014). However, it has also been suggested that bidialectals are very different from bilinguals. Basing themselves on their work on Shetland (i.e., northern Scotland) bidialectals, Smith and Durham (2012) propose that, whereas a bilingual speaker might have two distinct grammars for, say French and English, a bidialectal speaker has one grammar, and within this, two dialects, resulting in a mixing of variants, blending into each other in everyday use.

Research on bilingual infant speech perception, not to mention bidialectal infant speech perception, is a relatively young area of

psycholinguistic research and the available studies have yielded mixed results (for a review, see Fennell, Tsui, & Hudon, 2016). Although the main goal of this thesis is to provide insight into the acquisition of a lexical tone system and not to investigate the particularities of bidialectal infant speech perception, we must acknowledge the special input conditions that our Limburgian population is exposed to. As a result of the highly variable input, the higher probability of hearing accented speech (e.g., Bosch & Ramon-Casas, 2011), and the large amount of lexical overlap due to many cognates (e.g., Sebastián-Gallés & Bosch, 2009), we cannot exclude the possibility that Limburgians may exhibit a greater acceptance of acoustic variation in phonetic categories. Due to these factors, it has been proposed that bi-/multidialectals or close-language-bilinguals might be more willing to accept small acoustic-phonetic changes and show greater flexibility of what they will allow as representative of a target label (Fennell et al., 2016). These input-related factors due to bidialectalism could additionally influence our Limburgian participants' sensitivity to pitch in comparison to our Dutch control groups.

2.2.2 The dialect of Roermond and its relation to other varieties

In this thesis, the focus is on the Limburgian dialect of Roermond, a city of approximately 57,000 inhabitants (status quo January 1, 2017, www.roermond.nl) in the centre of the Dutch province of Limburg. The choice to focus on one particular dialect instead of on Limburgian as a whole stems from the fact that Limburgian is not a homogeneous linguistic variety. We cannot speak of 'the Limburgian dialect'. The term Limburgian is to be understood as an umbrella term for many different dialects. Comparable to the pitch-accents in different varieties of Japanese, Norwegian, and Swedish (Tamaoka, Saito, Kiyama, Timmer, & Verdonschot, 2014; Wetterlin, 2007), the Limburgian word prosodic contrast may have different phonetic realizations across dialects, be embedded in different intonation systems or may be absent altogether (e.g., Gussenhoven, 2000a; Gussenhoven & Peters, 2008). Different

'dialect regions' have been classified among others on the basis of shared phonetic-phonological features (e.g., van de Wijngaard & Keulen, 2007). Dialects that belong to the same region show a greater degree of phonological (also prosodic) overlap than dialects from different regions. The Roermond dialect belongs to the dialect region called *East-Limburgian* (e.g., Bakker & van Hout, 2012; see area B in Figure 3).

Indeling van de Limburgse dialecten

- A Ripuarisch
- B Oost-Limburgs
- C Centraal-Limburgs
- D West-Limburgs
- E West-Limburgse overgangsdialecten
- F Brabants-Limburgs
- G Kleverlands

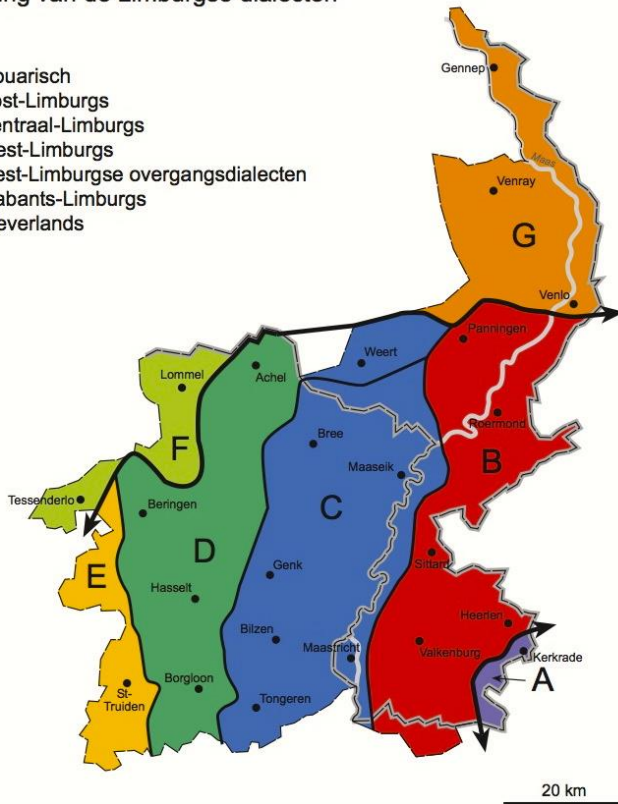


Figure 3. The different dialect regions in Dutch and Belgian Limburg (B = East-Limburgian). Retrieved from <http://willydolsstichting.nl/wp-content/uploads/2013/07/Taalkaart-Limburg.jpg>.

It can be assumed that the realization of the word prosodic contrast within the East-Limburgian dialect region does not show much variation (e.g., Heijmans, 2003). The word prosodic contrast is realized differently in the North-

Limburgian dialect of Venlo than in the dialect of Roermond, due to the fact that Venlo has a larger inventory of intonation contours that interact with the word-level contrast (e.g., Fournier, 2008; Gussenhoven & van der Vliet, 1999). The dialect of Venlo belongs to a different dialect region, called *Kleverlands* (see area G in Figure 3; Bakker & van Hout, 2012).

In the course of participant recruitment for our study in Roermond, it proved not to be feasible to limit ourselves to people with a purely Roermond background. People in Roermond (and in Limburg in general) are often exposed to various dialect combinations, among other things depending on the linguistic background of the parents. To be able to recruit a large enough sample within the time available, we decided to widen our criterion for participation to any infant or child exposed to an East-Limburgian dialect spoken by at least one parent or caregiver. Adult participants had to report to speak an East-Limburgian dialect themselves. The inclusion of participants from the entire East-Limburgian dialect region is based on the above-mentioned assumption that the Roermond tonal grammar is very similar to that of other dialects in this region (see Figure 3). Any inclusion of speakers from different dialect regions would have complicated our research to an extent that was outside the scope of this thesis.

The choice of the Roermond dialect in particular is partly motivated by the existence of a series of behavioural and neurophysiological studies investigating both the perception and production of the word prosodic contrast with adult native speakers of Roermond Dutch (Fournier, 2008; Fournier, Verhoeven, Swerts, & Gussenhoven, 2006; Fournier et al., 2010). Fournier and colleagues found that the contrast in Roermond is rather stable in comparison to the contrast in Venlo, as indicated by better discrimination of the word prosodic contrast in a perception experiment with adult speakers of the Roermond dialect as opposed to speakers of the Venlo dialect. The less stable status of the contrast in the dialect of Venlo can partly be attributed to the fact that Venlo is located in the periphery of the tonal area (see Figure 2). The

contrast might therefore be recessive under the influence of neighboring linguistic varieties that do not have the tone contrast (Fournier, 2008; Fournier et al., 2006). Another reason to focus on the dialect of Roermond is the fact that both its vocabulary and (tonal) grammar are well documented (e.g., Gussenhoven 2000a, 2000b; Kats, 1939, 1985).³ The fact that adult speakers of Roermond Dutch show sensitivity to the word prosodic contrast was a good starting point for our research and made it all the more reasonable to ask how children growing up with this dialect acquire this feature.

2.2.3 Word-level pitch in Roermond Dutch

Like in many varieties of North Germanic (Norwegian, Swedish, and some variants of Danish), in many Limburgian dialects a word prosodic contrast can signal lexical and morphological differences. This phenomenon has intrigued linguists at least since 1881 (Schmidt, 1986). According to Gussenhoven (2000a), the contrast emerged as a reaction to an analogical change in a neighbouring dialect. Short vowels in singular noun forms whose plural forms had undergone vowel lengthening as a result of *Open Syllable Lengthening* were analogically lengthened. A neighbouring generation of speakers decided not to adopt the lengthened vowels as phonologically long, as this would have obliterated the contrast between the singular and plural forms within their dialect. Their intriguing solution to this problem was to interpret the different pitch characteristics of the lengthened vowel, high level pitch, as tonal. The origin of the Central Franconian word prosodic contrast is thus assumed to be related to the desire to keep morphological minimal pairs distinct. The contrast was generalized after having reached morphological status as an indicator for grammatical number (Gussenhoven & Peters, 2008; but cf. Roos, 2009; Schmidt, 2002).

³ Yet another reason to focus on the East-Limburgian dialect region is the fact that the author [SR] is a native speaker of an East-Limburgian dialect.

48 Setting the Tone

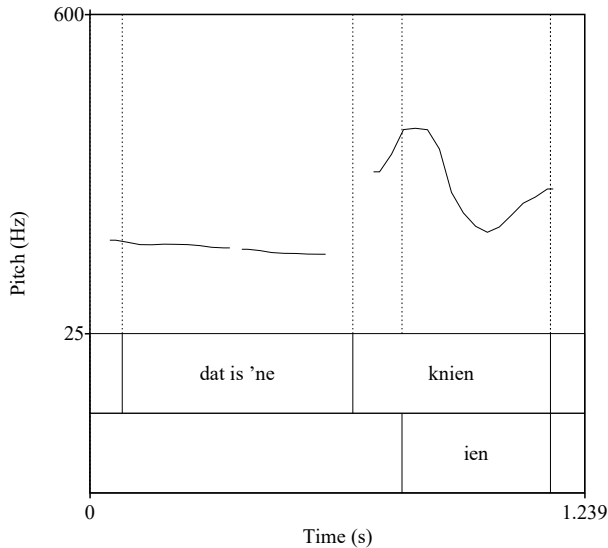


Figure 4B. F0 contour of the Limburgian sentence *dat is 'ne KNIEN* 'that is a rabbit'. The rhyme of the target word carries accent 2.

Whereas duration has repeatedly been put forward as an enhancing feature, the primary acoustic cue to the word prosodic contrast in Roermond Dutch is f_0 . This is based on perception data and acoustic inspection of production data of adult speakers of the Roermond dialect (e.g., Fournier et al., 2006). Yet, syllables bearing accent 2 tend to be slightly longer than their accent 1 counterparts in Roermond as well as in many other Limburgian areas (e.g., Fournier et al., 2006, and references therein; Köhnlein, 2016), which is in line with findings from other Franconian dialects (e.g., Werth, 2011). This can be understood if we assume that the origin of accent 2 was a lengthened vowel (Gussenhoven, 2000a). However, the word prosodic contrast in Roermond Dutch is clearly tonal, whereas Weert Dutch (belonging to the *Central Limburgian* dialect region) is clearly durational (Fournier, 2008; Fournier et al., 2006).

Tone in Limburgian has a lower functional load than tone in many Chinese dialects. Relative to Mandarin Chinese, there are few tonal minimal pairs. Fournier (2008) reported to have counted around 80 minimal pairs in

Kats' (1985) dictionary of Roermond Dutch. Note that these were mostly of a morphologically complex nature (e.g., verb inflections). There is only a two-way contrast, which has been described as privative. Gussenhoven and Peters (2008, p. 88) state that "the word accent contrast (...) amounts to a contrast between the absence of lexical tone (accent 1) and its presence (accent 2)". Moreover, in Dutch Limburgian dialects and in the dialect of Cologne, the contrast can only be realized on minimally bimoraic syllables with primary stress, meaning that an unbound multisyllabic morpheme can only carry one tone contrast. With respect to the functional load of tone, we assume Limburgian to be comparable to languages like Swedish (Bruce, 1977; Gussenhoven, 2004b; Riad, 2013), Norwegian (Kristoffersen, 2000; Steien & van Dommelen, 2016; Wetterlin, 2007) and Japanese (Kubozono, 1993). There is, however, one difference between these latter languages and the Limburgian dialect of Roermond (Gussenhoven, 2004b). In Swedish, and in most cases also in Norwegian, word-level pitch patterns extend over a larger temporal (i.e., multisyllabic) domain, that is, the tone contrast cannot appear on monosyllables.⁴ However, the tone contrast in Roermond Dutch is always realized within a single syllable. In this respect, Roermond Dutch is more akin to tone languages such as Mandarin Chinese.

Despite the relatively small number of minimal pairs, the lexical tone system is a crucial element of the Roermond Dutch grammar, since any primary stressed bimoraic syllable (i.e., a syllable consisting of a long vowel, a diphthong, or a short vowel plus a sonorant consonant) carries either accent 1 or accent 2 (Gussenhoven, 2000b). A substantial number of syllables thus cannot be pronounced correctly if a speaker does not know their tonal specification. For example, in Roermond Dutch, *boum* [bɔum] ('tree'; [bo:m] 'boom' in Dutch) is pronounced with accent 2, whereas *sjaop* [ʃɔ:p] ('sheep'; [sxa:p] 'schaap' in Dutch) is pronounced with accent 1 (for diachronic

⁴ Some dialects in Northern Norway and Sweden also have a surface accent distinction in monosyllabic words (Wetterlin, 2007).

descriptions of the allocation of the respective accents, see Gussenhoven, 2000a; Roos, 2009). Pronouncing any of these words with the wrong accent would turn them into a non-existing word. Pitch is thus part of a word's mental representation and as such connected to the speaker's knowledge of how a word is to be pronounced.

2.2.4 The tone-intonation interaction in Roermond Dutch

As in any other language, pitch in Limburgian also serves post-lexical and paralinguistic (i.e. pragmatic) functions. As has already been put forward in the first section of this chapter, of particular interest is the fact that the Limburgian dialects show complex intonation systems. As in most other Franconian dialects (Köhnlein, 2016), and even more so than in Swedish (Bruce, 1977), Limburgian lexical tones co-occur with intonation tones in syllables with main stress. The occurrence of different intonation contours causes the lexical tone contrast to have drastically different pitch contours, depending on whether the tone-bearing syllable is in focus, whether it appears in a question or a statement, and whether the syllable appears in phrase-final or non-final position (see Figure 5). This is different from Standard Chinese, where different intonations cause changes of register rather than of contours (Wu, 2000). In Limburgian, different prosodic contexts thus give rise to different phonetic realizations of lexical tones. In this way, every monosyllabic pitch contour represents a unique combination of intonational and lexical meaning. From the f_0 information on a monosyllabic word, a speaker of Limburgian thus has to decide (1) which lexical tone category it is, (2) which sentence type it is, and (3) whether the word is in focus. The Roermond tone categories do not share phonological forms with the other tone category, as is typically the case in Chinese tone sandhi.

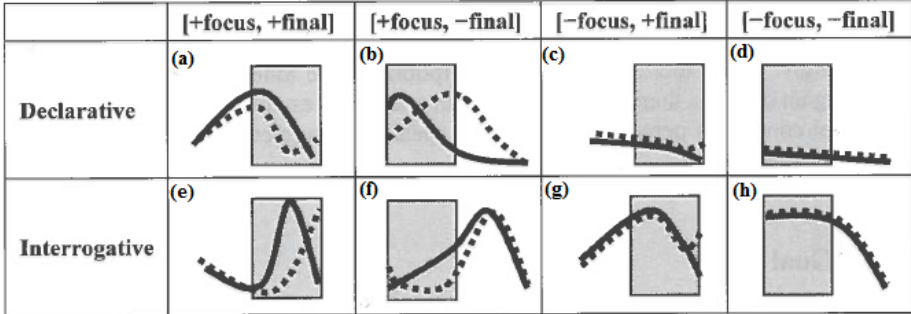


Figure 5. The different phonetic realizations of accent 1 (solid lines) and accent 2 (dashed lines) in Roermond Dutch depending on information status (+- focus), position (+- final), and sentence type (declarative vs. interrogative). Shaded portions indicate the stressed syllable. Adapted from Fournier (2008).

Fournier et al. (2006) showed that identification of the Limburgian tones by adult speakers of Roermond Dutch varied drastically depending on the prosodic context. Specifically, listeners recognized accents 1 and 2 in focused and/or phrase-final contexts, but failed to do so in phrase-internal, unaccented contexts. These perception data are on a par with Gussenhoven’s (2000b; see section 2.3) description of the Roermond tonal grammar, confirming a categorical neutralization of the Roermond tone contrast in non-final, non-focal contexts (see Figure 5, panels d and h, and section 2.3).

Due to this intricate tone-intonation interaction, the Limburgian prosodic system is of greater complexity than that of, for example, closely related Standard Dutch, a system with intonation contrasts only. This difference in complexity has been shown to determine how pitch processing is distributed over different brain cortices in Limburgian and Dutch listeners, in support of the view that pitch processing is function-driven and language-dependent (e.g., Gandour, 2007). In a mismatch paradigm using MEG, Fournier et al. (2010) showed that speakers of Roermond Dutch processed lexical tone contrasts predominantly in the left hemisphere, but intonation contrasts predominantly in the right hemisphere. Standard Dutch controls processed all contrasts in the left temporal cortex. The authors conclude that the greater functional

complexity of pitch in the Roermond linguistic system led to a different topography for the processing of pitch contrasts.

The fact that the Limburgian accents have many realizations could make it difficult for a language learner to abstract away from all these variations and recover the underlying tone system. Studies on lexical tone acquisition in Japanese and Swedish (Ota, 2003) and Sesotho (Demuth, 1995) have shown that the reliability of the mapping between underlying tones and their surface realizations has an impact on the speed of acquisition of tone. The tone-intonation interaction in Limburgian is more complex than in Swedish due to the larger number of intonation patterns leading to more surface variation (Gussenhoven, 2004b). It is thus reasonable to expect that learners of Limburgian are confronted with a difficult task. By looking at acquisition and processing of tone in Limburgian, we can add to the discussion of the influence of functional load and complexity on prosodic development.

The fact that pitch is the acoustic parameter that distinguishes accent 1 and accent 2 is not under discussion, but the underlying, phonological nature of the Limburgian word prosodic contrast is. Opinions differ on what the lexical representations of accent 1 and accent 2 words might look like. These different views will be the subject of the next section, where our aim is to point out that the tonal perspective taken in this thesis is to some extent controversial.

2.3 The phonological nature of word-level pitch in Roermond Dutch

In languages that make use of word-level pitch contrasts, pitch must be specified in the mental representation of words in one way or another. According to the tonal approach following the generative tradition (e.g., Gussenhoven, 2000b; Gussenhoven & Peters, 2004; Fournier, 2008), the Limburgian word prosodic surface contrast derives from the presence of tones in the mental lexicon. This could be achieved by including different tones (e.g., H and L) or the presence (e.g., H) versus absence of a tone. In the former case

the contrast is called equipollent, in the latter case it is called privative (van der Hulst, 1999). Gussenhoven (2000a) and Gussenhoven and Peters (2008) argue that the Limburgian word accent contrast amounts to the absence of lexical tone (accent 1) and its presence (accent 2), i.e., a privative contrast. However, no behavioural experimental data exist in support of or against this claim. By studying the acquisition and processing of the Limburgian tone contrast, we hope to contribute to the discussion on its lexical representation.

For describing the tonal grammar of the language under investigation, the framework of Autosegmental Metrical Theory is used (Goldsmith, 1976; Ladd, 2008; Pierrehumbert, 1980). The tonal contours of Roermond Dutch can be represented as sequences of H (high) and L (low) tones on a tier separate from consonants and vowels on the segmental tier. The tone bearing unit in Roermond Dutch – the element with which the tone can be associated – is the mora. Tones are considered autosegments because they can behave independently from consonants and vowels. The moras with which tones associate in Limburgian must occur in accented syllables or in the final syllable of the intonational phrase (henceforth: IP).

To describe the Roermond tonal grammar in more detail, we must consider a few properties of the IP, our constituent of interest. IPs are phonological units that carry complete intonation patterns, including minimally one stressed syllable, signalling utterance-level communicative meaning (Werth, 2011). Within this constituent, we observe intonational and lexical tones.

The intonational (post-lexical) tones are focal (or pitch) accents and boundary tones. Focal accents are aligned with the accented syllable. They are characterized as starred tones (T*) and are associated with the first mora of the primary stressed syllable of the accented word. In Dutch and Limburgian, focal accents are, among other things, used to mark focus, that is, to highlight the word that conveys new or important information. Boundary tones are characterized as T%. These tones mark the edges of intonational phrases and

54 Setting the Tone

indicate discourse meaning (e.g., question vs. statement) and phrasing (e.g., continuation).

For Roermond Dutch, the intonational contours for declarative and interrogative sentences look as given in Table 1:

Table 1. The intonation contours of Roermond Dutch (Gussenhoven, 2000b).

declarative:	H*L L%
interrogative:	H*L H%

Lexical tones are characterized as T_{LEX} . For now, we follow the assumption that the accent 1-accent 2 opposition is privative, i.e., a no-tone vs. lexical tone distinction. The second sonorant mora of accent 2 words is associated with an extra H (i.e., H_{LEX}). This extra tone represents the lexical tone that is inherent to accent 2 words and lacking in accent 1 words. Accent 2 is taken to be the lexically specified accent because the different realizations of accent 1 can be accounted for by intonation tones only. Accent 1 is considered the unmarked, default option. Any word with a bimoraic stressed syllable that has not been specified as H by the end of lexical phonology surfaces as accent 1. Monomoraic syllables cannot be specified for tone (but cf. Peters, 2008, for a different situation in the Belgian Limburgian dialect of Hasselt). For a schematic description of the proposed underlying representations of accent 1, accent 2, and monomoraic syllables in Roermond Dutch, see Table 2:

Table 2. Underlying representations of mono- and bimoraic syllables in Roermond Dutch, following Gussenhoven (2000b) and Fournier (2008). σ = syllable, μ = mora.

Monomoraic syllables	σ μ
Accent 1 syllables	σ └─┘ μ μ
Accent 2 syllables	σ └─┘ μ μ H_{lex}

In section 2.2.4, we explained that the phonetic realizations of accents 1 and 2 vary under the influence of the prosodic context (i.e., information status, position in the phrase, sentence type, see Figure 5). To explain how the phonetic realizations of accents 1 and 2 in the Roermond dialect are derived from their assumed underlying representations as depicted in Table 2, Gussenhoven (2000b) has proposed four post-lexical phonological rules (or constraints).

First, Roermond Dutch loses its lexical H tone in non-focal, non-final positions, causing the neutralization of the contrast in those contexts. The lexical H tone is only realized when it is associated with focal (containing T*) or final (containing T%) syllables. As reported in section 2.2.4, this neutralization has been shown to cause poor recognition of the contrast in non-focal, non-final contexts (see Figure 5, panels d and h), whereas recognition in focal-final contexts is excellent (at least for adult speakers; Fournier et al., 2006).

Secondly, final boundary tones are realized before the lexical H if this H occurs on the final mora of the IP. In other words: when the final syllable of an IP carries accent 2, the boundary tone occurs before the lexical tone (*align Tlex right*). This constraint is apparently ranked higher than a constraint which says

56 Setting the Tone

that boundary tones should be aligned with the edges of IPs. Consider Figure 6, panel a, illustrating the underlying representation of the word *bein₂* [bɛɪn] ‘leg’.⁵ The boundary L% occurs after the lexical tone (H_{LEX}). However, this is not what we observe in production. Focus-final accent 2 words (as well as accent 2 words uttered in isolation) are produced with a rise at the end. The constraint *align Tlex right* can account for this phenomenon, leading to the surface representation shown in Figure 6, panel b.

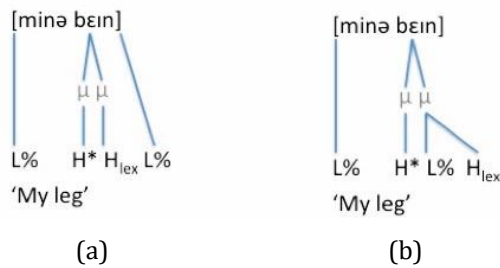


Figure 6. Underlying representation (a) and surface representation (b) of *Miene BEIN₂* (‘my leg’).

Third, the lexical H in syllables carrying accent 2 undergoes assimilation to L if it occurs after a L* in the same syllable in non-final position (*No Rise*). See for example Figure 7, where *bein₂* [bɛɪn] occurs in such a focal, non-final interrogative context. Due to assimilation of the lexical H on the second mora to the focal L on the first mora, *bein₂* [bɛɪn] ‘leg’ is pronounced with low pitch (see also the contour for accent 2 in a focal, non-final interrogative context in Figure 5, panel f).

⁵ Subscripts indicate accents 1 and 2.

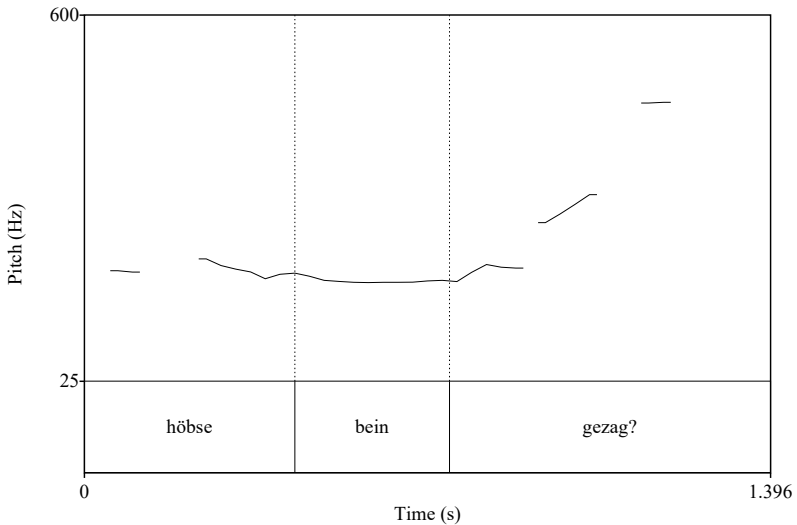


Figure 7. F0 contour of the Limburgian sentence *Höbse BEIN₂ gezag?* (‘did you say leg?’)

Fourth and finally, final boundary tones spread leftward into the phrase (*leftward tone spreading*) to a free mora in focal, non-final accent 1 syllables. In a declarative context, this results in a steep pitch fall in the tone-bearing syllable, as is shown for *book*₁ [bo:k] ‘book’ in Figure 8 (see also the contour of accent 1 in a focal, non-final declarative context in Figure 5, panel b). In an interrogative, this constraint causes a steep rise in the tone-bearing syllable (see the contour of accent 1 in a focal, non-final interrogative context in Figure 5, panel f).

58 Setting the Tone

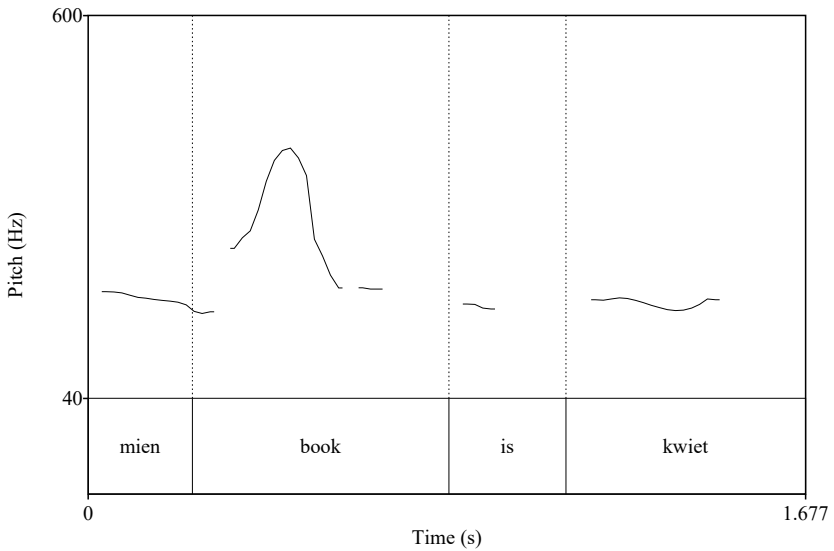


Figure 8. F0 contour of the Limburgian sentence *Mien BOOK₁ is kwiet* ('My book is lost').

Note that the boundary tone can only spread to the second mora if it does not already have a lexical tone. If we replaced *book*₁ [bo:k] 'book' with an accent 2 word, e.g. *sjoon*₂ [ʃo:n] 'shoe', where the second mora of the accent 2 word is already occupied by H_{LEX}, pitch would fall gradually until the end of the IP (see the contour of accent 2 in a focal, non-final declarative context in Figure 5, panel b).

Roos (2009) offers an alternative proposal from a diachronic perspective, according to which accent 1 is underlyingly specified as *Glottal Tension* (realized as a low tone) on the second mora in words carrying accent 1. Accent 2 words lack an underlying specification: their pitch contour results from spreading of the pitch accent on the first mora to the second (see also van Oostendorp, 2005). Unfortunately, there is no experimental evidence for either one of the above-mentioned proposals.

Apart from the fact that there is no consensus in the 'tonal camp' upon the question which tone is underlyingly specified (e.g., Boersma, unpublished; Gussenhoven, 2000a, 2000b; Roos, 2009), some scholars have questioned

whether the Limburgian accents come from lexical tones. They argue that there is no lexical tone in Limburgian, but that the contrast emerges from different foot structures (e.g., Hermans, 2012; Köhnlein, 2016). Proponents of this *metrical approach* argue that the only tones needed are the intonational focal and boundary tones. The tonal surface contrast can be attributed to the association of post-lexical intonation tones to tone bearing units (i.e., moras). The association of tones varies depending on the foot type.

So far there has been no experimental evidence that could solve the discussion on the phonological nature of the Limburgian accent contrast. As stated before, we will assume that the contrast derives from the presence of lexical tone. Although the primary goal of this thesis is to investigate acquisition and processing of the Limburgian tone contrast, some of our studies might contribute to the discussion on the lexical nature of the contrast by testing for perceptual asymmetries.

In the next chapter, we will first look at how (and whether) the perception of the Limburgian tone contrast develops in Limburgian and Dutch 6- to 12-month-old babies.

References

- Bakker, F., & van Hout, R. (2012). De afgrenzing tussen de noordelijke en zuidelijke dialecten in Noord-Limburg. *Taal en Tongval*, 64(2), 159-193.
- van Bezooijen, R., & van den Berg, R. (1999a): Taalvariateiten in Nederland en Vlaanderen: Hoe staat het met hun verstaanbaarheid? *Taal en Tongval*, 51(1), 15-33.
- van Bezooijen, R., & van den Berg, R. (1999b). Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions. In R. van Bezooijen & R. Kager (Eds.), *Linguistics in the Netherlands 16*, (pp. 1-12), Amsterdam: John Benjamins.
- Boersma, P. The history of the Franconian tone contrast. Unpublished manuscript, University of Amsterdam. Retrieved from <http://www.fon.hum.uva.nl/paul/papers/FranconianToneHistory68.pdf>.
- Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics*, 39, 514-525.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Cornips, L. (2013). Recent developments in the Limburg dialect region. In F. Hinskens & J. Taeldeman (Eds.), *Language and Space: An International Handbook of Linguistic Variation* (pp. 378-399). Berlin/Boston: De Gruyter Mouton.
- Cornips, L. (2014). Socio-syntax and variation in acquisition: problematizing monolingual and bidialectal acquisition. In K.K. Grohmann (ed.): Three factors and beyond: Socio-syntax and language acquisition. Special issue of *Linguistic Variation* 14(1), 1-25.

- Cornips, L., & Hulk, A. C. J. (2006). External and Internal Factors in Bilingual and Bidialectal Language Development: Grammatical Gender of the Dutch Definite Determiner. In C. Lefebvre, L. White, & C. Jourdan (Eds.), *L2 Acquisition and Creole Genesis. Dialogues* (pp. 355-378). Amsterdam/Philadelphia: Benjamins.
- Cornips, L., de Rooij, V., Stengs, I. L., & Thissen, L. (2016). Dialect and local media. Reproducing the multi-dialectal space in Limburg (the Netherlands). In J. Thøgersen, N. Coupland, & J. Mortensen (Eds.), *Style, Media, and Language Ideologies* (pp. 189-216). St. Charles, MO: Novus Press.
- Demuth, K. (1995). The acquisition of tonal systems. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 111-134). Hillsdale: Lawrence Erlbaum.
- Driessen, G. (2006). Ontwikkelingen in het gebruik van streektalen en dialecten in de periode 1995-2003. *Toegepaste Taalwetenschap in Artikelen*, 75, 103-113.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford University Press.
- Fennell C.T., Tsui, A.S., & Hudon, T.M. (2016). Speech perception in simultaneously bilingual infants. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the Lifespan: Factors Moderating Language Proficiency* (pp. 43-62). Berlin/New York: De Gruyter Mouton.
- Fournier, R. (2008). *Perception of the tone contrast in East-Limburgian dialects*. LOT Dissertation series.
- Fournier, R., Gussenhoven, C., Jensen, O., & Hagoort, P. (2010). Lateralization of tonal and intonational pitch processing: An MEG study. *Brain Research*, 1328, 79-88.

62 Setting the Tone

- Fournier, R., Verhoeven, J., Swerts, M., & Gussenhoven, C. (2006). Perceiving word prosodic contrasts as a function of sentence prosody in two Dutch Limburgian dialects. *Journal of Phonetics*, 34(1), 29-48.
- Fromkin, V. (Ed.) (1978). *Tone. A linguistic survey*. New York: Academic Press.
- Gandour, J. (2007). Neural substrates underlying the perception of linguistic prosody. In C. Gussenhoven, & T. Riad (Eds.), *Tones and tunes: Experimental studies in word and sentence prosody*, vol. 2 (pp. 3-25), Berlin/New York: Mouton de Gruyter.
- Goldsmith, J. (1976). *Autosegmental phonology*. PhD MIT.
- Goss, S. (2015). *The effects of internal and experience-based factors on the perception of lexical pitch accent by native and nonnative Japanese listeners*. PhD dissertation, Ohio State University.
- Gussenhoven, C. (1999). Discreteness and gradience in intonational contrasts. *Language and Speech*, 42(2-3), 283-305
- Gussenhoven, C. (2000a). On the origin and development of the Central Franconian tone contrast. In A. Lahiri (Ed.), *Analogy, leveling, markedness. Principles of change in phonology and morphology* (pp. 215-260).
- Gussenhoven, C. (2000b). The lexical tone contrast of Roermond Dutch in Optimality Theory. In M. Horne (Ed.), *Prosody: Theory and Experiment* (pp. 129-167). Amsterdam: Kluwer.
- Gussenhoven, C. (2004a). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2004b). Tone in Germanic: Comparing Limburgian with Swedish. In G. Fant, H. Fujisaki, J. Cao, & Y. Xu (Eds.), *From Traditional*

Phonology to Modern Speech Processing (pp. 129-136). Beijing: Foreign Language Teaching and Research Press.

Gussenhoven, C., & Bruce, G. (1999). Word prosody and intonation. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 233-266). Berlin/New York: Mouton de Gruyter.

Gussenhoven, C. & Peters, J. (2004). A tonal analysis of Cologne Schärfung. *Phonology*, 21(2), 251-285.

Gussenhoven, C., & Peters, J. (2008): De tonen van het Limburgs. *Nederlandse Taalkunde*, 13(1), 87-114.

Gussenhoven, C. & van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.

Hermans, B. (2012). The phonological representation of the Limburgian tonal accents. In B. Botma, & B. Noske (Ed.), *Phonological explorations. Empirical, Theoretical and Diachronical issues* (pp. 223-239). Berlin/Boston: De Gruyter.

Heijmans, L. (2003). The relationship between tone and vowel length in two neighboring Dutch Limburgian dialects. In P. Fikkert & H. Jakobs (Eds.), *Development in Prosodic Systems* (pp. 7-46). Berlin/New York: Mouton de Gruyter.

van der Hulst, H. (1999). Word Accent. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe*. Berlin: De Gruyter.

Hyman, L.M. (1985). *A theory of phonological weight*. Dordrecht: Foris.

Hyman, L.M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, 31, 213-238.

64 Setting the Tone

- Kats, J.C. (1939). *Het phonologisch en morphologisch systeem van het Roermondsch dialect*. Roermond: J.J. Romen & Zonen.
- Kats, J.C. (1985). *Remunjs Woardebook*. Roermond: H. van der Marck en Zonen.
- Kirk, N., Fiala, L., Scott-Brown, K., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, 26(6), 640-648.
- Köhnlein, B. (2016). Contrastive foot structure in Franconian tone-accent dialects. *Phonology*, 33, 87-123.
- Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford university press.
- Kubozono, H. (1993). *The organization of Japanese prosody*. Tokyo: Kurosio.
- Ladd, D.R. (2008). *Intonational Phonology*. 2nd edition. Cambridge: Cambridge University Press.
- O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (Eds.). (2001). *Contemporary Linguistics*. Boston: Bedford/St. Martin's.
- Oh, Y.M., Coupé, C., Marsico, E., & Pellegrino, F. (2015). Bridging phonological system and lexicon: insights from a corpus study on functional load. *Journal of Phonetics*, 53, 153-176.
- Oostendorp, M. van. (2005). Expressing inflection tonally. *Catalan Journal of Linguistics*, 4, 107-126.
- Ota, M. (2003). The development of lexical pitch accent systems: an autosegmental analysis. *Canadian Journal of Linguistics*, 48(3/4), 357-383.
- Peters, J. (2008). Tone and intonation in the dialect of Hasselt. *Linguistics*, 46, 983-1018.

- Pierrehumbert, J. (1980). *The phonetics and phonology of English intonation*. PhD dissertation, MIT.
- Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Linguistic inquiry monograph 15. Cambridge: MIT Press.
- Riad, T. (2013). *The phonology of Swedish*. Oxford: Oxford University Press.
- Roos, N. (2009). *The weak past tense in Dutch and low German*. PhD dissertation, Radboud University, Nijmegen.
- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, 5(4), 489-522.
- Schmidt, J. (1986). *Die mittelfränkischen Tonakzente (Rheinische Akzentuierung)*. Stuttgart: Franz Steiner.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874-887.
- van Severen, L., Gillis, J.J.M., Molemans, I., van den Berg, R., de Mayer, S., & Gillis, S. (2013). The relation between order of acquisition, segmental frequency and function: the case of word-initial consonants in Dutch. *Journal of Child Language*, 40, 703-740.
- Singh, S., & Fu, C. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, 87(3), 834-854. doi:10.1111/cdev.12512.
- Smith, J. & Durham, M. (2012). Bidialectalism or dialect death? Explaining generational change in the Shetland Islands, Scotland. *American Speech*, 87(1), 57-88.

- Steien, G.B., & van Dommelen, W.A. (2016). The production of Norwegian tones by multilingual non-native speakers. *International Journal of Bilingualism*. doi: 10.1177/1367006916673218.
- Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R. (2014). Is pitch accent necessary for comprehension by Japanese speakers? – An ERP investigation. *Journal of Neurolinguistics*, 27, 31-40.
- Tong, Y., Francis, A., & Gandour, J. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, 23(5), 689-708.
- Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*, 4(1), 1-22.
- Wang, Y., Behne, D., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, 25, 449-466.
- Werth, A. (2011). *Perzeptionsphonologische Grundlagen der Prosodie. Eine Analyse der mittelfränkischen Tonakzentdistinktion*. Stuttgart: Franz Steiner Verlag.
- Wetterlin, A. (2007). *The Lexical Specification of Norwegian Tonal Word Accents*. Doctoral dissertation, University of Konstanz.
- Wijngaard, T. van de, & Keulen, R. (2007). De indeling van de Limburgse dialecten. In R. Keulen, T. van de Wijngaard, H. Crompvoets, & F. Walraven (Eds.), *Riek van klank. Inleiding in de Limburgse dialecten* (pp. 15-23). Sittard: Veldeke Limburg.
- Wu, Z. (2000). From traditional Chinese phonology to modern speech technology: Realization of tone and intonation in spoken Chinese.

Proceedings of the International Conference on Spoken Language Processing (ICSLP 2000), Vol. 1, B1-B12.

Wu, X., Tu, J.-Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623-641.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

3. No perceptual reorganization for Limburgian tones? A cross-linguistic investigation with 6- to 12-month-old infants

Based on: Ramachers, S., Brouwer, S., and Fikkert, P. (2017). 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.

Abstract

Despite the fact that many of the world's languages use lexical tone, the majority of language acquisition studies has focused on non-tone languages. Research on tone languages has typically investigated well-known tone languages such as Mandarin and Cantonese Chinese. The current study looked at a Limburgian dialect of Dutch that uses lexical pitch differences, albeit in a rather restricted way. Using a visual habituation paradigm, 6- to 12-month-old Limburgian and Dutch infants were tested for their ability to discriminate Limburgian tones. The results showed that both Limburgian and Dutch infants discriminate the Limburgian tones throughout their first year of life. The role of linguistic experience, acoustic salience and the degree of similarity to the native prosodic system are discussed.

3.1 Introduction

Infants learn their native language (L1) at a surprisingly rapid pace. Before producing their first words, they have already learned many of the phonetic contrasts that are relevant for distinguishing meaning in their L1. From birth onwards, infants are able to perceive almost any phonetic difference, irrespective of its existence in the L1 (for a review, see Werker & Tees, 1992). However, between 6 and 12 months infants typically begin to perceive phonetic contrasts in accordance with their native phonology. This change or relocation of attention is often called PERCEPTUAL TUNING or PERCEPTUAL REORGANIZATION (henceforth PR). It is a process of perceptual decline, maintenance, or facilitation of phonetic discrimination under the influence of ambient language input. Native vowel perception seems to be in place somewhere between 6 and 8 months of age (e.g., Polka & Werker, 1994; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), whereas attenuation of non-native consonant contrasts appears later, at around 10 to 12 months of age (e.g., Kuhl et al., 2006; Tsushima et al., 1994).

One of the factors guiding PR is the experience with the ambient language and, at the same time, a lack of experience with non-native speech sounds. However, presenting PR as a process of maintenance and loss under the influence of L1 experience is too limited. Some native contrasts have proven to be rather difficult and are acquired relatively late (e.g., for nasal place discrimination: Narayan, Werker, & Beddor, 2010; for discrimination of vowel duration: Sato, Sogabe, & Mazuka, 2010a), whereas certain non-native contrasts remain perceivable into adulthood (e.g., Best, McRoberts, & Sithole, 1988). Additional factors that are held responsible for this more nuanced pattern of speech sound perception are acoustic salience, general perceptual maturation and/or the extent to which non-native sounds are similar to or different from native sounds (Best, 1995). The interaction between psychoacoustic and experiential factors gives way to multiple routes to adult perception of different speech sounds. Best et al. (1988) for example showed

that American English-learning infants between 6 and 14 months of age as well as English-learning adults are able to discriminate non-native Zulu click contrasts. They argue that these contrasts are sufficiently different from English contrasts and are therefore easy to perceive. Polka and Bohn (1996) show a similar result for the German front-back high vowel contrast /y/-/u/ with English-learning infants. Narayan et al. (2010) show that even for native contrasts it depends on their relative acoustic salience how much input is required, or in other words how long it takes for infants to successfully discriminate the respective native contrast.

Apart from the mere experience with certain (non-)native segmental contrasts, also their psychoacoustic salience and the extent to which they are similar to native contrasts might thus determine the pathway to adult perception of these speech sound contrasts.

3.1.1 The case of lexical tone

PR studies on suprasegmental, or prosodic, contrasts are relatively scarce compared to studies investigating segmental contrasts. Examples of suprasegmental contrasts are phoneme duration (i.e., contrasts between long versus short vowels or single versus geminate obstruents), lexical stress (e.g., English *INsight* vs. *inCITE*, capitals indicate the stressed syllable) and lexical tone.

In tone languages, pitch patterns are used to distinguish words, similar to what phonemes do at the segmental level. By some estimates the number of languages using lexical tone is as high as 60-70% (Yip, 2002). Some of these languages make extensive use of tone. In a tone language like Mandarin Chinese, for example, almost every syllable has to be pronounced with one of four tones (Duanmu, 2000; Gussenhoven, 2001). Other tone languages are more restricted in their use of lexically contrastive pitch. These languages, for example Tokyo Japanese, have been referred to as RESTRICTED TONE LANGUAGES

(Voorhoeve, 1973). From a typological point of view, the phonetic and functional distinction between tone languages on the one hand and restricted tone languages on the other hand is hard to draw (e.g., Gussenhoven, 2004a). Nevertheless, many phonologists still recognize an intermediate type of language, also referred to as a PITCH-ACCENT LANGUAGE. According to Yip (2002, p. 257), pitch-accent languages are a subtype of tone languages. Hyman (2009) too argues against pitch-accent languages as a third prosodic type of language. He states that all languages that appear to be intermediate between tone languages and stress languages (e.g., Tokyo Japanese, Somali, Swedish) do not form a coherent class. All languages that have been called pitch-accent languages can be analyzed as tonal, because what all these languages have in common is that pitch, to a greater or lesser extent, is necessary for determining the meaning of a word (Hyman, 2009; Werth, 2011). We will use the term tone language for languages that use pitch to signal word-level meaning differences, unless when referring to studies that used a different term.

With this study, we want to contribute to the field of lexical tone acquisition by investigating a relatively unknown and understudied population: infants growing up with tonal Limburgian dialects of Dutch. From a typological point of view, these dialects are particularly interesting for two reasons. First, tone in Limburgian has a relatively low functional load compared to many Chinese dialects. Secondly, there is a complex interaction between lexical and intonational tones, which is unusual for tone languages, leading to many different surface realizations of the underlying tone categories. By looking at the developmental perception of lexical tone in Limburgian infants, we can add to the discussion of the influence of functional load and complexity on the acquisition of a lexical tone system. We compared the Limburgian infants to a control group of infants acquiring Standard Dutch, a non-tonal language. Before we discuss Limburgian in more detail, we will first review the extant work on lexical tone acquisition.

3.1.2 Perceptual reorganization for lexical tones

With respect to the development of native tone and pitch-accent perception, studies have shown good discrimination performance throughout the first year of life (e.g., Sato, Sogabe, & Mazuka, 2010b; Tsao, 2017; Yeung, Chen, & Werker, 2013). Sato et al. (2010b) have looked at native perception of lexical pitch-accent in Japanese, and showed that 4- and 10-month-old Japanese infants could readily discriminate Japanese pitch-accents. However, as previously documented for native segment perception (e.g., Narayan et al., 2010), native perception of lexical pitch also appears to be subject to effects of psychoacoustic salience. Tsao (2008) tested Mandarin 10- to 12-month-olds on Mandarin tone contrasts differing in acoustic salience and indeed attested contrast-dependent differences in performance. For example, tones 1 (high level) and 3 (low dipping) were discriminated with higher accuracy than tones 2 (low rising) and 3 (see also Tsao, 2017).

The existing studies on non-native lexical tone perception in the first year of life have yielded mixed results (for a critical review, see Singh & Fu, 2016). Some studies indicate that non-tone-learning infants show a decline in lexical tone discrimination around 9 months (English and Yorùbá infants, Yorùbá tones: Harrison, 2000; English and Chinese infants, Thai tones: Mattock & Burnham, 2006; English and French infants, Thai tones: Mattock, Molnar, Polka, & Burnham, 2008; English, Cantonese and Mandarin infants, Cantonese tones: Yeung et al., 2013). Other more recent studies by Liu and Kager (2014) and Chen and Kager (2016) investigated discrimination of Mandarin tones in Dutch infants and found a maintained sensitivity to non-native pitch patterns. Liu and Kager showed that 5- to 18-month-old Dutch infants can discriminate Mandarin tone 1 (high level) from tone 4 (high falling). Apparently, the Dutch infants in this study did not go through a process of PR for these tones, or the process was overridden by other factors like psychoacoustic salience. Indeed, 8- to 15-month-olds did not succeed in discriminating a reduced variant of this contrast in which the difference in fundamental frequency between the two

pitch patterns had been made less extreme. However, five- to six-month-olds and 17-18-month-olds were successful in this condition, leading to a U-shaped developmental pattern. The authors propose that the perceptual salience of a tonal contrast – which may depend on the distance between the pitch patterns – influences the process of PR. The original, non-reduced contrast was apparently salient enough to overcome PR.

Another potential explanation put forward by Liu and Kager (2014) for the success of the 8- to 15-month-old Dutch infants on the original, non-reduced contrast is that Dutch infants interpret the Mandarin tones as instances of Dutch intonation contours. Dutch has a rich intonation system involving meaningful variation in pitch contours (Gussenhoven, 2005). Yeung et al. (2013) also consider the possible influence of native intonation when discussing their finding that Mandarin and Cantonese Chinese 4- to 9-month-old infants attend differently to one particular Cantonese tone (tone 25). Referring to a study on adult native and non-native tone perception (So & Best, 2010), Yeung et al. attribute this difference in perception to the way in which Cantonese tone 25 relates to the native tone inventories of Cantonese and Mandarin. That tone perception is partly governed by the precise relation between the tones used in an experiment and the learner's native tone or intonation inventory is among others documented by Reid et al. (2014) and by Schaefer and Darcy (2014).

In line with Liu and Kager (2014), Chen and Kager (2016) also found improvement rather than attenuation in sensitivity towards non-native lexical tones. Dutch 4-, 6- and 12-month-old infants were tested on the Mandarin tone 2 (low rising) – tone 3 (dipping) contrast, which is assumed to be a non-salient contrast (e.g., Tsao, 2008). Four-month-olds did not show discrimination, whereas 6- and 12-month-olds discriminated the non-native tone contrast both in a condition without token variation, encouraging phonetic listening, and in a condition with token variation that encouraged phonological listening. The authors concluded that this improvement shows that infants improve in perceiving non-native tones acoustically (i.e., phonetically) and that they also

become more capable of normalizing variable tokens when they grow older. They attribute this finding to general cognitive/perceptual maturation.

To summarize, psychoacoustic salience and general perceptual maturation are claimed to play a role in native and in non-native tone perception. In non-native tone perception, the extent to which non-native tones are similar or different to tone or intonation categories from the L1 might also influence the developmental perception of lexical tone contrasts.

However, as became apparent from this section, most studies on lexical pitch perception have focused on a small set of target languages that make a rather extensive use of lexical tone (e.g., Cantonese, Mandarin). Granting that all languages that use pitch as a distinctive feature in at least a few words can be considered tone languages, we still have to recognize that, within the family of tone languages, large differences exist. First, they differ with respect to the functional load of tone. The smaller the tonal inventory, the greater the distributional restrictions, and the smaller the number of tonal minimal pairs, the more restricted a tone system might be considered to be (Voorhoeve, 1973). The functional load of word-level pitch has been assumed to influence listeners' sensitivity to lexical pitch patterns (e.g., Goss, 2015; Schaefer & Darcy, 2014; Wang, Behne, Jongman, & Sereno, 2004; Wu, Tu, & Wang, 2012).

Secondly, tone languages differ with respect to the complexity of their intonation systems. Intonation is found in all of the world's languages (Singh & Fu, 2016; Yip, 2002). It is used to distinguish sentence types, like statements versus questions, to indicate whether a speaker has finished or intends to continue speaking, or to show which parts of an utterance present new or important information. Usually, tone languages do not have complex intonation systems (e.g., Gussenhoven & van der Vliet, 1999). In Standard Chinese, for example, different intonations only cause changes of pitch register rather than of pitch contours (e.g., Wu, 2000). In Swedish (Bruce, 1977), and even more so in Limburgian (Gussenhoven, 2004b), lexical and intonational pitch patterns interact, leading to various phonetic realizations of an underlying tone

category. It has been shown that the reliability of the mapping between underlying tones and their surface realizations has a large impact on the acquisition of a lexical tone system (e.g., Demuth, 1995; Ota, 2003).

This study aims at extending our knowledge about the influence of the functional load of tone and the complexity of the prosodic system on developmental pitch perception. It looks at infants acquiring an understudied language, i.e., a Limburgian dialect of Dutch, and compares them to a control group of infants growing up with non-tonal Standard Dutch. The language under investigation will be discussed in more detail in the next section.

3.1.3 Limburgian dialects of Dutch

The Limburgian dialects of Dutch belong to the Central Franconian dialect-continuum, which covers the provinces of Limburg in the Netherlands and Belgium as well as the north of the German Rhineland-Palatinate and the southwest of North-Rhine Westphalia (Fournier, 2008; Gussenhoven, 2000a; see Figure 2 in Chapter 2 of this thesis).

Approximately 75% of the inhabitants of the Dutch province of Limburg speak a Limburgian dialect (Driessen, 2006), which is a regional variety of Standard Dutch. Differences between Limburgian and Dutch exist at the phonological, morphosyntactic and lexical level. Still, mutual intelligibility between Dutch and Limburgian is fairly high (van Bezooijen & van den Berg, 1999), probably due to the existence of many cognates, leading to a high degree of phonological overlap. Most Limburgians also speak Dutch, the official language used in formal and institutional settings, and are considered bidialectal (Cornips, 2014). We understand bidialectalism as bilingualism involving closely related linguistic varieties. Children in Limburg thus often acquire a local dialect in addition to the standard language. They may be immersed in one particular dialect from birth and encounter Standard Dutch upon entering the educational system as well as through media and peers, or alternatively one parent speaks a dialect and the other parent either speaks Dutch (possibly with a Limburgian accent) or a different Limburgian dialect.

Research on the influence of bilingualism on early speech perception is a young area of psycholinguistic research. Extant studies involving both distant- and close-language-bilinguals have yielded mixed results (for a review, see Fennell, Tsui, & Hudon, 2016). To our knowledge, no studies looked at early speech perception in bidialectal infants (but see Durrant, Delle Luche, Cattani, & Floccia, 2015, for a study on the specificity of early lexical representations in multidialectal children). Nevertheless, we cannot rule out the possibility that, next to the differences in the native prosodic system that are the topic of this study, additional input-related factors could influence the development of lexical tone perception in Limburgian participants in comparison to their Dutch peers. As a consequence of the highly variable input Limburgians are exposed to, the higher probability of hearing accented speech (e.g., Bosch & Ramon-Casas, 2011) and the large amount of lexical overlap due to many cognates (e.g. Sebastián-Gallés & Bosch, 2009), Limburgian infants might exhibit a greater acceptance of acoustic variation in phonetic categories and show (temporary) discrimination difficulties (Fennell et al., 2016).

Recently, Limburgian bidialectals have been the subject of studies looking at linguistic development in older children and on the cognitive benefits of bilingualism. Cornips (2014) looked at vocabulary acquisition and grammatical development in Limburgian 4- to 6-year-old bidialectal children. They were compared to monolingual Dutch and bilingual peers. The bilingual children had Dutch as their L1, and the other language varied across participants. The bidialectal children did not differ significantly from monolingual Dutch peers, but they did differ from their bilingual peers in that the bidialectals scored significantly higher on vocabulary and were faster in acquiring grammatical gender in Dutch. From this we might conclude that Limburgian bidialectals, at least with respect to certain vocabulary and grammar development measures, cannot be directly compared to bilinguals. Recent work by Blom, Boerma, Bosma, Cornips, and Everaert (2017) also shows that six- to seven-year-old bidialectal Limburgian children differ from

Polish-Dutch bilingual peers but not from monolingual Dutch peers on a flanker task tapping into executive functioning. More research is needed to provide insights into the potentially unique acquisition pathways of bidialectal children and into the mechanisms that are responsible for these potential differences. For now, we cannot exclude the possibility that Limburgian infants might differ from Dutch infants with respect to their early phonetic development as a result of their special input properties.

3.1.4 Lexical tone in Roermond Dutch

In many Limburgian dialects, word-level pitch patterns signal lexical and grammatical differences. For example, in the dialect of Roermond, *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 also serves a grammatical function with accent 1 systematically indicating plurality: *knien* [kni:n] with accent 2 means 'rabbit', but pronounced with accent 1 it means 'rabbits' (see Figures 1A and B for the pitch contours of accents 1 and 2, realized on a focal-final target word in a declarative sentence spoken by a female native speaker). Based on perception and production data with adult speakers of the Roermond dialect, Fournier, Verhoeven, Swerts, and Gussenhoven (2006) concluded that pitch is the primary acoustic cue to the Limburgian tone contrast.

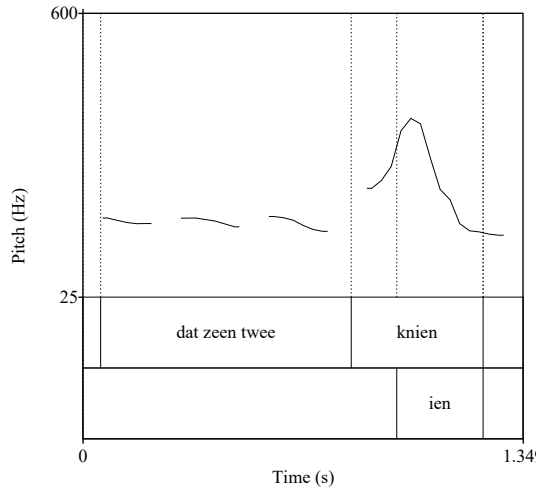


Figure 1A. F0 contour of the Limburgian sentence *dat zeen twee KNIEN* 'those are two rabbits'. The rhyme of the target word carries accent 1.

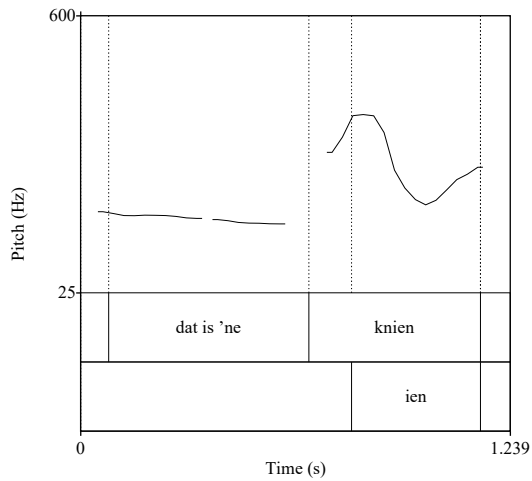


Figure 1B. F0 contour of the Limburgian sentence *dat is 'ne KNIEN* 'that is a rabbit'. The rhyme of the target word carries accent 2.

Tone⁶ in Limburgian has a lower functional load than tone in many Chinese dialects. Relative to Mandarin Chinese, there are few tonal minimal pairs. Fournier (2008) counted around 80 in Kats' (1985) dictionary of the Roermond dialect. There is only a two-way contrast distinguishing between a fall (H*L, accent 1) and a fall-rise (H*LH, accent 2)⁷, and the contrast can only be realized on bimoraic syllables with primary stress. Any unbound multisyllabic morpheme can only carry one accent. With respect to the functional load of tone, we assume Limburgian to be comparable to Swedish (Gussenhoven, 2004b; Riad, 2013) and Norwegian (Kristoffersen, 2000; Wetterlin, 2007). However, these languages differ from Limburgian with respect to the temporal domain of tone. In Swedish, and in most cases also in Norwegian, the tone contrast can only appear on minimally bisyllabic words (Wetterlin, 2007). In Roermond Dutch, the tone contrast is realized within a single syllable.

As in any other language, pitch in Limburgian also serves post-lexical and paralinguistic (i.e. pragmatic) functions. As has been mentioned earlier, of particular interest is the fact that the Limburgian dialects show complex intonation systems, which is rare for a language that also has lexical tone (Gussenhoven & van der Vliet, 1999). The Roermond word accents can have drastically different pitch contours depending on whether the tone-bearing syllable is in focus, whether it appears in a question or a statement and whether the syllable appears in phrase-final or non-final position. In this way, every monosyllabic pitch contour represents a unique combination of intonational and lexical meaning. This is different from Standard Chinese, where different intonations cause changes of register rather than of contours (Wu, 2000). The

⁶ Note that some scholars have questioned whether the Limburgian accents come from lexical tones. They argue that there is no lexical tone in Limburgian, but that the contrast emerges from different foot structures (e.g., Köhnlein, 2016).

⁷ Accent 1 as a fall and accent 2 as a fall-rise are the citation tones as spoken in isolation or in the final focal syllable of a declarative utterance. The realization of the tone contrast depends on the prosodic context, as will be explained later.

fact that the Limburgian accents have many realizations could make it difficult for a language learner to uncover the underlying tone system (Ota, 2003). Despite the fact that, on the one hand, these different surface realizations can be regarded meaningless allophonic variants of the underlying lexical tone system, on the other hand they are meaningful at the post-lexical level in that they all signal unique information with respect to sentence type and information status. In other words, speakers of Limburgian need to attend to every single aspect of these variants to uncover the differences in intonational meaning. This might (eventually) lead to a greater sensitivity to pitch differences in general.

With respect to the number of surface realizations, the Limburgian lexical tone system is of greater complexity than that of Swedish (Gussenhoven, 2004b). Due to the use of lexical tone, Limburgian is also more complex than closely related Dutch, a system with intonation contrasts only. This difference in complexity has been shown to determine how pitch processing is distributed over the different brain cortices in Limburgian and Dutch listeners, in support of the view that pitch processing is function-driven and language-dependent (Fournier, Gussenhoven, Jensen, & Hagoort, 2010).

3.1.5 The present study

The aim of the present study is twofold. First, we want to provide insights into the development of lexical pitch perception in Limburgian infants growing up with a restricted tone language (Limburgian) and a non-tone language (Dutch) simultaneously. In accordance with previous results on native tone and pitch-accent perception in the first year of life, we expect that 6- to 12-month-old native Limburgian infants successfully discriminate the Limburgian tones throughout their first year of life. Because of their daily experience with both lexical tones and intonational tones, we expect their perception of native pitch patterns to be maintained or perhaps even enhanced between 6 and 12 months.

Secondly, we compare Limburgian infants to Dutch infants learning an intonation language that does not use lexical pitch. Given the differences in complexity of the prosodic systems, we wanted to explore the possibility that Limburgian children's sensitivity to lexical pitch develops differently from that of Dutch children. For the Dutch infants, two outcomes are possible. On the one hand, they could show an age-related performance decline (as the non-native listeners in Mattock & Burnham, 2006; Mattock et al., 2008; and Yeung et al., 2013), which would result in an interaction involving both the language background and the age of the participants. On the other hand, Dutch infants could maintain their discrimination of the Limburgian tones, which would result in the absence of an interaction involving language and age. Maintenance of discrimination could be explained by psycho-acoustic salience, general perceptual maturation or comparability to the native intonation system overriding PR, as was the case for the Dutch infants in Liu and Kager (2014) and Chen and Kager (2016).

3.2 Method

Discrimination performance of 6-, 9- and 12-month-old Limburgian and Dutch infants was assessed using the HYBRID VISUAL HABITUATION PROCEDURE (HVHP, Houston, Horn, Qi, Ting, & Gao, 2007). Visual habituation experiments traditionally consist of two phases: a habituation phase and a test phase. During habituation, the infant is habituated to one particular type of stimulus. In the subsequent test phase, the infant is presented with instances of the habituated stimulus (SAME trials) and with instances of a different type of stimulus (SWITCH or CHANGE trials). If the infant notices the difference between the two types of stimuli, it is expected to look longer in the switch trials than in the same trials. The HVHP is a variant of the traditional VHP (Stager & Werker, 1997) and combines aspects of the Stimulus Alternating Preference Procedure (SAPP; Best & Jones, 1998) and of the Oddity Paradigm (Picton, Alain, Otten, Ritter, & Achim, 2000). Houston et al. (2007) have shown that the HVHP elicits greater

mean looking time differences between same and switch trials than versions of the VHP that either incorporate only the SAPP or the Oddity Paradigm. More details follow in the Procedure section.

3.2.1 Participants

A total number of 54 Limburgian infants were recruited via health care institutions and daycare centers in the city of Roermond, Limburg, the Netherlands. Thirty-nine full-term infants were included in the analysis: 11 six-month-olds (range = 6 months, 8 days to 7 months, 11 days; 8 boys), 10 nine-month-olds (range = 8 months, 15 days to 10 months, 10 days; 3 boys) and 18 twelve-month-olds (range = 10 months, 28 days to 13 months, 7 days; 8 boys). An additional 15 infants were tested but excluded from analysis due to failing to reach the habituation criterion ($N = 2$; both 12-month-olds) or not completing enough test trials ($N = 13$; 6-month-olds: $N = 1$, 9-month-olds: $N = 4$, 12-month-olds: $N = 8$).

It was impossible to recruit infants with exclusive exposure to Roermond Dutch. Since the tonal grammar of Roermond Dutch is very similar to the tonal grammar of other East-Limburgian dialects (e.g., Heijmans, 2003), infants who were exposed to any East-Limburgian dialect spoken by at least one caregiver were allowed to participate (see section 2.2.2 in Chapter 2 for more details). In the case of 25 of the 39 children included in the analysis, both parents spoke East-Limburgian, in six cases exactly the same dialect and in 19 cases a different dialect; in a further five cases, one parent spoke an East-Limburgian dialect, while the other spoke a Limburgian dialect from a different region, and in nine cases, one parent spoke East-Limburgian and the other Standard Dutch. Overall, 23 of the participants had at least one parent speaking the dialect of Roermond.

A total number of 126 Dutch infants were recruited from the subject pool of the Baby Research Center of Radboud University in Nijmegen, the Netherlands. All infants were full-term and grew up in monolingual Dutch-

speaking families. Eighty-three infants were included in the analysis, of which 28 six-month-olds (range = 5 months, 5 days to 6 months, 30 days; 12 boys), 29 nine-month-olds (range = 8 months, 0 days to 9 months, 23 days; 16 boys), and 26 twelve-month-olds (range = 11 months, 3 days to 12 months, 20 days; 11 boys). An additional 43 infants were tested but excluded from the analysis due to failing to reach the habituation criterion ($N = 4$; all 12-month-olds), equipment or experimenter error ($N = 7$), failing to contribute at least 50% of same and/or switch trials ($N = 31$; 6-month-olds: $N = 10$, 9-month-olds: $N = 13$, 12-month-olds: $N = 8$) or an inappropriate linguistic background ($N = 1$).

To make sure that none of the Dutch infants had substantial experience with a Limburgian dialect, their parents were asked questions related to the linguistic input of their child during an intake phone call. An infant was regarded to have substantial experience with Limburgian and thus not suitable for participation if (1) one of the parents or primary caregivers was a native speaker of a Limburgian dialect, (2) the infant had weekly contact with a native speaker of Limburgian.

3.2.2 Stimuli

Stimuli were two pseudo-words that only differed in tone: *taag*₁ [ta:ç] and *taag*₂ [ta:ç].⁸ This minimal pair was created from an existing tonal minimal pair in East-Limburgian, indicating a grammatical difference: *daag*_{1/2} [da:ç] ('days', 'day'). A female native speaker of the Roermond dialect recorded multiple tokens in declarative sentences in sentence final, focal position. Stimuli were recorded in context to avoid listing intonation, and only in one prosodic context to avoid differences in the phonetic realization of the tones (recall that the phonetic realization of the Limburgian tones depends on the prosodic context). The speaker was selected for her ability to speak both the dialect and accentless Dutch. She was asked to produce the words in a child-directed

⁸ Subscripts indicate accent 1 and accent 2.

manner. Recordings took place in a sound attenuated booth at the Max Planck Institute for Psycholinguistics in Nijmegen. The stimuli were digitized onto a PC using PRAAT (44,1kHz) (Boersma & Weenink, 2012). We followed the guidelines presented in Turk, Nakai, and Sugahare (2006) for excising the pseudoword stimuli from their carrier sentences.

Five tokens per tone type were selected. The selection was based on native intuition of a native speaker of an East-Limburgian dialect [SR] and careful listening of a trained phonetician [Carlos Gussenhoven]. Some variation with respect to absolute pitch height was allowed to introduce minimal variation between tokens of one tone category (see Figures 2A and 2B).

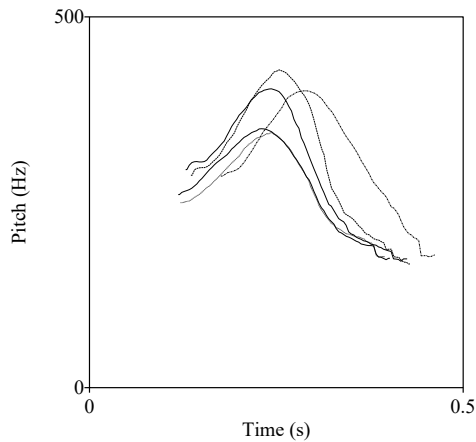


Figure 2A. F0 contours of the syllable nuclei of the five different tokens of *taag* with accent 1.

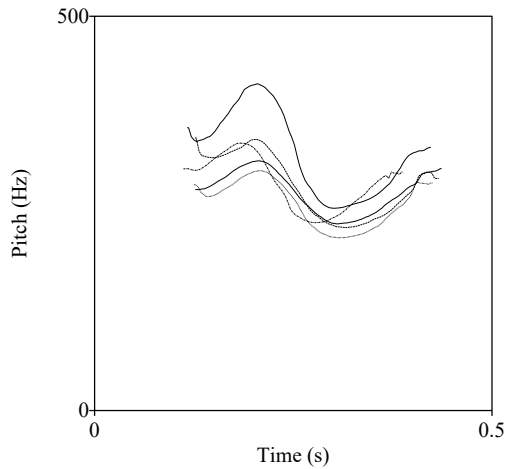


Figure 2B. F0 contours of the syllable nuclei of the five different tokens of *taag* with accent 2.

As for between-category variation, we wanted participants to discriminate the contrast based on pitch differences rather than duration differences. 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. See Table 1 in the Appendix for the acoustic measurements per token. Tokens with deviant voice quality (e.g., creaky voice) were excluded. We controlled for intensity by equalizing all selected tokens to 65dB. Discriminability of the tokens was piloted in an ABX-discrimination task with twelve adult speakers of the Roermond dialect. Their mean discrimination accuracy was 87,5%, indicating that they could readily distinguish between accent 1 and accent 2.

3.2.3 Apparatus

Limburgian infants were tested in an office in a daycare center in Roermond. They sat in front of a 24-inch LCD screen (Philips 249C4QHSB) and were recorded via a digital video camera (Sony HC40) mounted on a tripod below the table. Behind the monitor were two speakers (Logitech Z130). The video

camera broadcast the recording to a 13-inch Apple MacBook Air. Recordings were made with the video software Vidi (version 0.4.7). The experiment was presented using the LOOK software (Meints & Woodford, 2008), run on a laptop (HP EliteBook Folio 9470m). To mimic the set-up in Nijmegen, the monitor was placed inside a wooden frame. During testing, the experimenter and the parent listened to masking music through noise-cancelling headphones (Sennheiser HME 110).

Dutch infants were tested in a dimly lit room in the Baby Research Center in Nijmegen. Infants sat on their caregiver's lap in a baby car seat to reduce external distraction. The experiment was run in a test booth (size: 128 x 177cm), which is partly closed by black wooden partitions, left and right from the 47-inch television screen (LG 47LK530 ZC). A digital video camera (Sony Handycam DCR_HC85E PAL) was placed 30cm below the screen, hidden by a black curtain with an opening for the lens. The video camera provided a broadcast of the infant's behavior to a monitor behind the TV. Recordings of the infant for offline coding were made using Virtual Dub (Version 1.9.11). The experiment was controlled using the same software as in Roermond. Both the experimenter and the parent wore noise-cancelling headphones (Sennheiser HMEC 300) that played masking music.

3.2.4 Procedure

Infants' discrimination was assessed using the HVHP (Houston et al., 2007). The experiment was divided into a habituation phase and a test phase. The habituation phase consisted of maximally 12 trials. Habituation trials consisted of four of the five selected tokens, repeated blockwise in pseudo-random order with a maximum of 32 stimuli per trial (eight repetitions of four tokens). As soon as the mean looking time in three consecutive trials had reached less than 65% of the looking time on the first three trials, the habituation criterion was met and the infant would proceed to the test phase. If a child did not habituate in twelve trials, the habituation criterion was not met and the infant was

excluded from the analysis. Half of the infants were habituated on *taag*₁ and half of the infants on *taag*₂ to test for a possible perceptual asymmetry (e.g., Tsao, 2008).

The test phase consisted of four switch trials and eight same trials. Same trials consisted of alternations between the fifth and thus novel token of the habituated type and the four tokens the infant was habituated on (Houston et al., 2007 call these trials NON-ALTERNATING). Switch trials were made up of alternations between one token of the novel type (thus carrying the tone the infant was *not* habituated on) and the four tokens from the habituation phase (Houston et al., 2007 call these ALTERNATING). Thus, in same trials there was token alternation whereas in switch trials there was type alternation. Which type of trial was presented first was counterbalanced between infants. As a result, there were two pseudo-random trial orders, with switch trials never occurring twice in a row.

The visual stimulus in both phases was a dynamic checkerboard pattern (e.g., Horowitz, 1975; Stager & Werker, 1997). Trials were separated with a visual attention getter consisting of a purple flashing light. The inter-stimulus interval in all trials was one second.

Online looking time was recorded by pressing a designated key whenever the child was looking at the image on the TV. A trial was initiated as soon as the infant fixated the attention getter. If the infant looked away for more than two seconds, the trial was aborted. If looking time within a trial did not exceed two seconds, the trial was restarted.

3.2.5 Data pre-processing

Infants' video recordings were coded offline using ELAN (version 4.5.0; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) with a resolution of 40 frames per second. A random 20% of the videos were recoded by a second experienced coder. Intercoder reliability was very high for both same and switch trials (Pearson's $r = .995$, $p = .01$).

Next to participant exclusion criteria, which were already mentioned in the participants section, we adopted a series of trial exclusion criteria. A trial was excluded if (1) offline coded looking time within the trial was less than two seconds and (2) a trial was restarted more than once. Moreover, (3) all trials after three restarted trials were excluded as well, as we thought three restarts signal a serious decrease in attention and engagement. Trials could also be excluded as a consequence of (4) technical error, (5) external interference, (6) crying/fussing of the child or (7) difficulty in coding looking behavior. The eighth and last trial exclusion criterion was maximum looking time within a trial (54 seconds), which could signal that the infant is staring and not really paying attention to the experiment. This criterion did not apply to any participant. For a detailed overview of trial and participant exclusion per language and age group, see Table 2 in the Appendix.

Infants' mean looking time during the switch trials was compared to their mean looking time in same trials. In this procedure, a significant main effect of trial type indicates that the child discriminates between same and switch trials. If this is the case, children are expected to show a novelty effect and thus attend longer to switch trials.

3.3 Results

To see whether Limburgian and Dutch infants differed in the time they needed to habituate, we ran an independent *t*-test comparing the number of habituation trials between Limburgian and Dutch infants. There was no significant difference in the length of the habituation phase between Limburgian infants ($M = 5.6, SD = 2.1$) and Dutch infants ($M = 5.4, SD = 1.9$), $t(120) = .55, p > .1$.

Discrimination of Limburgian accent 1 and accent 2 was analyzed in a three-way mixed ANOVA with trial type (switch vs. same) as the within subjects factor and language (Limburgian vs. Dutch) and age (6-month-olds vs. 9-

month-olds vs. 12-month-olds) as between-subjects factors. The dependent variable was mean looking time.

Because the assumption of normality was not met, the raw mean looking times towards switch and same trials were logarithmically transformed to fit a normal distribution. After transformation, Levene’s test indicated equal variances ($p > .05$). The analysis was run on the transformed data. To enable a more straightforward interpretation of the data, however, we will report the raw means.

Figures 3A and 3B show infants’ looking times to same and switch trials in the test phase, split by age group and language background.

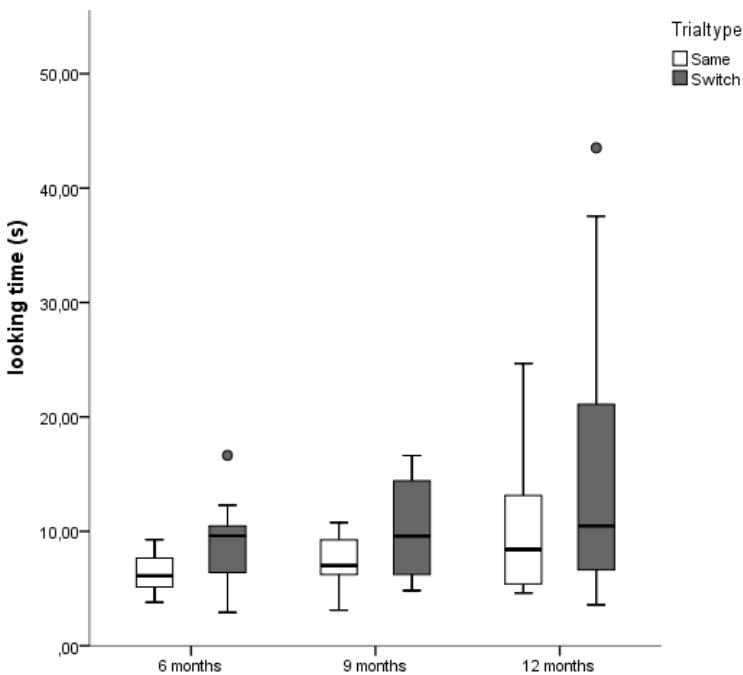


Figure 3A. Boxplot showing the median and interquartile range of looking time in same and switch trials for *Limburgian* infants at age 6, 9, and 12 months. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers.

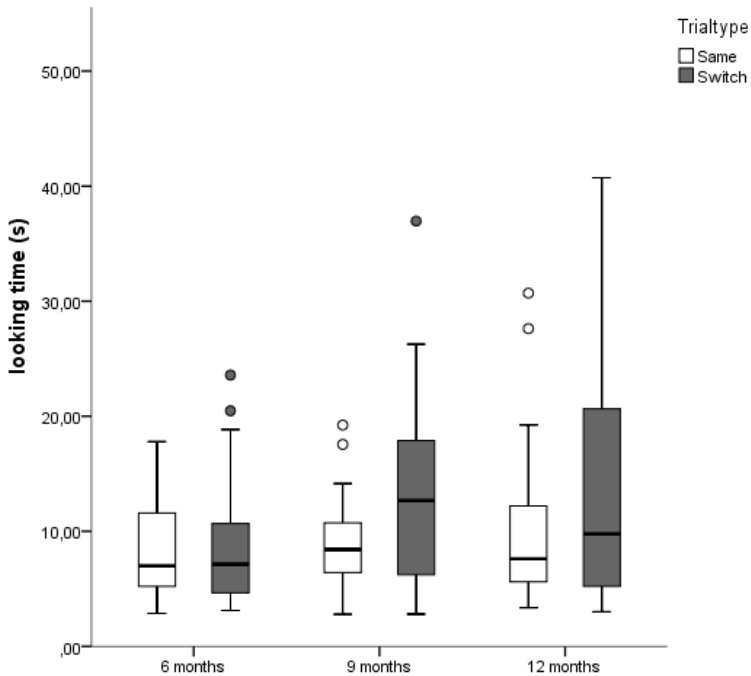


Figure 3B. Boxplot showing the median and interquartile range of looking time in same and switch trials for *Dutch* infants at age 6, 9, and 12 months. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers.

Results showed a significant main effect of trial type, indicating that infants looked significantly longer during switch ($M_{switch} = 11.9$ s, $SD = 8.2$ s) than during same trials ($M_{same} = 9.2$ s, $SD = 5.6$ s), $F(1,116) = 15.90$, $p < .001$, $\eta_p^2 = 0.12$, observed power = 0.977. We also found a marginally significant main effect of age, showing that overall looking times increased with age; $F(1,116) = 2.94$, $p = .057$, $\eta_p^2 = 0.05$, observed power = 0.563. No other main effects nor interactions reached significance (all $ps > .1$).⁹

⁹ An additional analysis was run on a sample excluding participants whose looking times in same and/or switch trials were over two standard deviations from the mean. For the Limburgian sample this meant that three 12-month-olds had to be excluded, resulting in a total Limburgian sample of 36 infants. For the Dutch infants, two 6-month-olds, three 9-month-olds and three 12-month-olds had to be excluded, yielding a total sample of 75 Dutch infants. This analysis yielded the same results.

3.4 Discussion

The current study explored the developmental pattern in infants' native and non-native perception of Limburgian lexical tones between 6 and 12 months of age. We expected that native Limburgian infants would discriminate the Limburgian tones throughout their first year of life. Non-native Dutch babies could either show a performance decline or they could maintain their sensitivity to the Limburgian contrast.

The main effect of trial type and the absence of any significant interactions indicate that both Limburgian and Dutch babies are able to discriminate the Limburgian tonal contrast throughout their first year of life, meaning that our hypothesis with respect to the development of the Limburgian infants is met. Because of their daily exposure to lexical tones, we expected them to discriminate the native tonal contrast at all the ages at test.

This finding is in line with earlier research on native lexical pitch-accent and tone acquisition (e.g., Sato et al., 2010b; Tsao, 2017; Yeung et al., 2013). Note, however, that the sample sizes per age group for the Limburgian sample were relatively small. We cannot exclude the possibility that larger sample sizes could have yielded a more nuanced developmental pattern, like an improvement in discrimination (e.g., Tsao, 2017).

The present data, however, provide no support for PR in non-native tone perception, since there was no interaction involving age and/or language background. Although the interaction with age was not significant, visual inspection of the data suggests that Dutch infants' discrimination at 6 months is weaker than their discrimination at 9 and 12 months (see Figure 3B). Our failure to find an interaction with age and/or language background needs to be interpreted with caution because of the relatively small Limburgian sample size. However, it is in line with recent studies by Liu and Kager (2014) and Chen and Kager (2016) in showing persistent sensitivity to non-native pitch patterns in Dutch infants.

These findings complement results from a vast amount of studies on adult non-native tone perception, demonstrating persistent tone sensitivity in speakers of non-tone languages (e.g., Hallé, Chang, & Best, 2004; So & Best, 2010, 2014). Singh and Fu (2016) propose that the developmental pathway of tone perception should therefore not be compared to the development of phonetic segment perception, but should rather be studied and described as a phenomenon in its own right, because tone appears to stand out among other phonological categories.

Several explanations can be put forward for our finding that Dutch infants maintain sensitivity to Limburgian lexical tones: the influence of acoustic salience, general perceptual maturation, and the influence of native intonation. Regarding acoustic salience, because of the fact that there is a wide variety of phonetic contrasts available in the world's languages, it stands to reason that some of these contrasts will be perceptually easier to differentiate than others. A number of studies on the development of (non-)native infant speech perception have shown that factors like psychoacoustic salience can indeed affect the developmental perception of segmental and suprasegmental contrasts in the first year of life (e.g., Best et al., 1988; Narayan et al., 2010; Polka & Bohn, 1996; Tsao, 2008, 2017).

Liu and Kager (2014) have shown that 5- to 18-month-old Dutch infants can readily discriminate Mandarin tones 1 (high level) and 4 (high falling). However, 8- to 15-month-olds did not succeed on a reduced variant of this contrast in which the difference between the contours had been made less extreme. The authors suggest that the perceptual salience of a tonal contrast influences the process of PR. The original, non-reduced contrast was apparently salient enough to be resistant to perceptual tuning. It is possible that the perceptual difference between accents 1 and 2 in Limburgian – at least in their citation form – is also salient enough to override experience-based PR.

Recall that our data suggest weaker performance of the Dutch infants at 6 months than at 9 and 12 months. If perceptual salience were the

explanation for maintenance of discrimination by non-natives in the second half of the first year of life, we would also expect salience to affect discrimination in a positive way at a younger age. Since this seems not to be the case, we would like to put forward other explanations that might be more consistent with the Dutch infants' results.

One alternative explanation for the perceptual pattern found in our Dutch infants is general perceptual maturation. Chen and Kager (2016), reporting similar results, propose that general auditory development and the acquisition of language-specific acoustics might complement each other. When infants grow older, they get better at perceiving acoustical information in general (see also Tsao, 2017).

Another alternative explanation is the influence of native post-lexical prosody. Dutch has a rich intonation system and has intonation contours that are similar to the Limburgian tones in our stimuli. Limburgian accent 1 is similar to the highly frequent falling H*L L% nuclear contour of Dutch, while accent 2 is quite like the Dutch H*L H% as used on phrase-final monosyllables. A fall-rise can signal a question, a reminder, a suggestion, or non-finality (Gussenhoven, 2004). However, as accent 2 in Limburgian or Limburgian-accented Dutch can also occur in prosodic contexts that are unexpected for Dutch listeners (e.g., in declaratives), it is usually the tone that is perceived as 'foreign' by speakers of Dutch.

An infant exposed to an intonation language like Standard Dutch thus receives input that contains melodies applied to both one-word or multiword utterances. Importantly, infant directed speech (IDS) tends to consist of short utterances (Fernald & Hurtado, 2006; Snow, 1977). Frota, Butler, and Vigário (2014) examined the input directed to four European Portuguese infants ranging from 11 to 16 months of age and found that 43% of the intonational phrases in IDS consisted of single prosodic word utterances. Nearly half of these utterances were of the declarative or question sentence type (27.3 and 19.4%, respectively). Van de Weijer (2001) investigated single-word utterances in IDS directed to a Dutch infant between six and nine months old and showed

that almost 40% of all utterances were single-word utterances. Most of these utterances were either fillers (e.g., *Yes, Hm*), social expressions (e.g., greetings) or vocatives (e.g., name of the addressee). The percentage of declaratives, interrogatives and imperatives was only 8.6%. However, van de Weijer's findings apply to only one child. For now, we assume that Dutch infants are probably also regularly exposed to various intonation contours realized on single-word utterances, which is exactly the kind of stimuli that have been used in the current experiment.

The Dutch infants in our experiment could thus perceive the Limburgian tones, which were instantiated on isolated words, as instances of native sentence-level intonation, which they probably also frequently encounter on single-word utterances in their daily lives (Frota et al., 2014). This possibility is also put forward by Liu and Kager (2014) in explaining the maintenance of Mandarin tone discrimination in Dutch infants. Best et al. (1988) had already suggested that, at least for consonants, around 10-12 months of age “a phonemic process appears that assimilates speech sounds to native categories whenever possible” (p. 345). In her PERCEPTUAL ASSIMILATION MODEL (PAM), Best (1995) offers an explanation for the process that phonetic and phonological properties of the L1 influence perception in a foreign language. The model proposes that non-native sound contrasts tend to be perceived according to their degree of similarity to native sounds. The perceived distance between the non-native contrast and the closest native sound(s) (if any) leads to differences in non-native contrast discriminability. The first applications of PAM concerned auditory perception of consonants and vowels, but it has also successfully been extended to the perception of suprasegmentals (PAM-S; Hallé, Chang, & Best, 2004; Reid et al., 2015; So & Best, 2010, 2014). The PAM-framework has also been applied to infant speech perception development (Best, 1994; Yeung et al., 2013).

In the first year of life, infant perception of speech contrasts shifts from phonetic (or universal) to phonological (or language-specific) due to growing

experience with the native phonological system. Despite infants' sensitivity to the prosodic properties of language at birth (e.g., Nazzi, Floccia, & Bertoncini, 1998), a number of studies have shown that infants' perception of certain aspects of prosody reflect the nature of the native prosodic system only at around 9 months of age (e.g. Jusczyk, Cutler, & Redanz, 1993; Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009; Skoruppa et al. 2009; but see Frota et al., 2014). Our data suggest that Dutch infants' discrimination of the Limburgian tones was less strong at 6 months than at 9 and 12 months. This pattern might reflect their accumulating knowledge of native sentence-level prosody. Frota et al. (2014) have shown that English infants can discriminate segmentally varied, single-prosodic word utterances presented with statement or yes-no question intonation at 6 and 9 months of age. At 9 months, Dutch infants might also have accumulated enough knowledge of native sentence-level intonation to enable them to distinguish questions from statements. Other studies on infant discrimination of sentence-level intonation did not offer direct evidence of intonation discrimination because their (whole-sentence) stimuli differed in both intonation and word order (Geffen & Mintz, 2011) or because they found an overall preference for stimuli with question intonation (Soderstrom, Ko, & Nevzorova, 2011).

It has already been stated that the Limburgian tones are quite similar to frequent Dutch intonation contours. The PERCEPTUAL ASSIMILATION MODEL FOR SUPRASEGMENTALS (PAM-S; So & Best, 2010, 2014) outlines a number of different patterns of perceptual assimilation. These patterns go with different expectations regarding discriminability. If two non-native sounds are similar to two native sounds, TWO-CATEGORY (TC) ASSIMILATION applies. Following PAM-S, Limburgian accent 1 and accent 2 could both be classified as instances of TC assimilation: two non-native categories, accent 1 and accent 2, assimilate to two native categories, the Dutch sentence-level intonation contours H*L L% and H*L H%. TC assimilation typically leads to high discrimination performance (e.g., Best, McRoberts, & Goodell, 2001). A discrimination paradigm with Limburgian and Dutch adults similar to Braun and Johnson (2011) could

provide insights into the relative contribution of psychoacoustic salience and assimilation to native prosody.

Another interesting question for future research is whether the Limburgian tones are interpreted as lexically relevant by native speakers of Limburgian. We expect tones to be part of the phonological representation of Limburgian words that meet the segmental requirements for the tone contrast to appear (see section 3.1.4). In Dutch, however, pitch differences are not contrastive at the word level, so we would expect at least Dutch listeners not to attribute lexical relevance to tones. The interpretation of Limburgian tones in children and adults acquiring Limburgian or Dutch is currently under study in our laboratory. Previous word learning studies with English children (e.g., Hay, Graf Estes, Wang, & Saffran, 2015; Quam & Swingley, 2010; Singh, Hui, Chan, & Golinkoff, 2014) have indeed shown that around 20 months of age, when word learning is well underway, non-tone-language children treat pitch differences as not lexically contrastive.

In conclusion, the present study contributes to the field of lexical tone acquisition by studying a relatively unknown and understudied population: infants growing up with restricted tonal Limburgian dialects of Dutch. We compared them to non-tonal Dutch peers growing up with an intonation language. Both Limburgian and Dutch infants perceive Limburgian accents 1 and 2 throughout their first year of life. For future research it is important to disentangle the relative contribution of acoustic salience and linguistic experience to these results, as well as provide insights into the function or importance that both of these populations assign to pitch contour differences at the word level.

References

- Best, C.T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: the transition from speech to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.
- 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. 171-204). York Timonium, MD: York Press.
- Best, C.T., & Jones, C. (1998). Stimulus-alternating preference procedure to test infant speech discrimination. *Infant Behavior and Development*, *21*, 295.
- Best, C.T., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, *109*(2), 775-794.
- Best, C.T., McRoberts, G.W., & Sithole, N.M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345-360.
- van Bezooijen, R., & van den Berg, R. (1999). Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions. In R. van Bezooijen & R. Kager (Ed.), *Linguistics in the Netherlands 16* (pp. 1-12). Amsterdam: John Benjamins.
- Blom, E., Boerma, T., Bosma, E., Cornips, L., & Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Frontiers in Psychology*, *8*(552). doi:10.3389/fpsyg.2017.00552.

- Boersma, P., & Weenink, D. (2012). *PRAAT: doing phonetics by computer* [Computer program]. Version 5.3.22, <http://www.praat.org/>.
- Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics*, 39, 514-525.
- Braun, B., & Johnson, E. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, 39, 585-594.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, 25(5), 426-439.
- Cornips, L. (2014). Socio-syntax and variation in acquisition: problematizing monolingual and bidialectal acquisition. In K.K. Grohmann (Ed.), *Three factors and beyond: Socio-syntax and language acquisition*. Special issue of *Linguistic Variation*, 14(1), 1-25.
- Demuth, K. (1995). The acquisition of tonal systems. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 111-134). Hillsdale: Lawrence Erlbaum.
- Driessen, G. (2006). Ontwikkelingen in het gebruik van streektalen en dialecten in de periode 1995-2003. *Toegepaste Taalwetenschap in Artikelen*, 75, 103-113.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford University Press.
- Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, 42, 447-465.

100 Setting the Tone

- Fennell C.T., Tsui, A.S., & Hudon, T.M. (2016). Speech perception in simultaneously bilingual infants. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the Lifespan* (pp. 43-62). De Gruyter Mouton.
- Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9(3), F33-F40.
- Fournier, R. (2008). *Perception of the tone contrast in East-Limburgian dialects*. Doctoral dissertation, Radboud University, Nijmegen. LOT Dissertation series.
- Fournier, R., Gussenhoven, C., Jensen, O., & Hagoort, P. (2010). Lateralization of tonal and intonational pitch processing: An MEG study. *Brain Research*, 1328, 79-88.
- Fournier, R., Verhoeven, J., Swerts, M., & Gussenhoven, C. (2006). Perceiving word prosodic contrasts as a function of sentence prosody in two Dutch Limburgian dialects. *Journal of Phonetics*, 34(1), 29-48.
- Frota, S., Butler, J., & Vigário, M. (2014). Infant's perception of intonation: Is it a statement or a question? *Infancy*, 19(2), 194-213.
- Geffen, S., & Mintz, T. (2011). Seven-month-olds discrimination of statements and questions. In A.K. Biller, E.Y. Chung, & A.E. Kimball (Eds.), *Online Supplement of the Proceedings of the 36th Boston University conference on language development*. Somerville, MA: Cascadilla Press.
- Goss, S. (2015). *The effects of internal and experience-based factors on the perception of lexical pitch accent by native and nonnative Japanese listeners*. Doctoral dissertation, Ohio State University.
- Gussenhoven, C. (2000a). On the origin and development of the Central Franconian tone contrast. In A. Lahiri (Ed.), *Analogy, leveling, markedness*.

Principles of change in phonology and morphology (pp. 215-260). Berlin, Boston: de Gruyter.

Gussenhoven, C. (2000b). The lexical tone contrast of Roermond Dutch in Optimality Theory. In M. Horne (Ed.), *Prosody: Theory and Experiment* (pp. 129-167). Amsterdam: Kluwer.

Gussenhoven, C. (2001). Suprasegmentals. In N. J. Smelser and P. B. Baltes (Eds.), *International Encyclopedia of the Social and the Behavioural Sciences* (pp. 15294-15298). Oxford: Pergamon.

Gussenhoven, C. (2004a). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

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

Gussenhoven, C. (2005). Transcription of Dutch intonation. In S.-A. Jun (Ed.), *Prosodic typology: The Phonology of Intonation and Phrasing* (pp. 118-145). Oxford: Oxford University Press.

Gussenhoven, C., & Bruce, G. (1999). Word prosody and intonation. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 233-266). Berlin/New York: Mouton de Gruyter.

Gussenhoven, C., & Peters, J. (2008). De tonen van het Limburgs. *Nederlandse Taalkunde*, 13(1), 87-114.

Gussenhoven, C., & van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.

102 Setting the Tone

- Hallé, P., Chang, Y.-C., & Best, C. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581-616.
- 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(1), 10-22.
- Höhle, B., Bijeljic-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.
- Horowitz, F.D. (1975). Visual attention, auditory stimulation, and language discrimination in infants. *Monographs of the Society for Research in Child Development*, 39(5/6), Serial No. 158.
- Houston, D., Horn, D., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, 12(2), 119-145.
- Hyman, L.M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, 31, 213-238.
- Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English. *Child Development*, 64, 675-687.
- Köhnlein, B. (2016). Contrastive foot structure in Franconian tone-accent dialects. *Phonology*, 33, 87-123.
- Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.

- Kuhl, P., Stevens, E., Hayashi, A., Degushi, 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.
- Kuhl, P., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by six months of age. *Science*, 255, 606-608.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133, 385-394.
- Mattock, K. & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization. *Infancy*, 10, 241-265.
- 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.
- Meints, K., & Woodford, A. (2008). *Lincoln Infant Lab Package 1.0: a new programme package for IPL, Preferential Listening, Habituation and Eyetracking*. <http://www.lincoln.ac.uk/psychology/babylab.htm>.
- 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(3), 407-420.
- Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, 21, 779-784.
- O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (Eds.). (2001). *Contemporary Linguistics*. Boston: Bedford/St. Martin's

104 Setting the Tone

- Ota, M. (2003). The development of lexical pitch accent systems. *Canadian Journal of Linguistics*, 48(3/4), 357-383.
- Picton, T.W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology and Neuro-Otology*, 5, 111-139.
- Polka, L. & Bohn, O. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, 100(1), 577-592.
- 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.
- 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.
- Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Rattanasone, N.X., & Best, C.T. (2015). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Attention, Perception, & Psychophysics*, 77, 571-591.
- Riad, T. (2013). *The phonology of Swedish*. Oxford: Oxford University Press.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010a). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, 46(1), 106-119.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010b). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, 22(11), 2503-2513.

- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, 5(4), 489-522.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874-887.
- Singh, S., & Fu, C. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, 87(3), 834–854. doi:10.1111/cdev.12512.
- Singh, L., Hui, T.J., Chan, C., & Golinkoff, R. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science*, 17(1), 94-109.
- Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés N., Alves Limussuri, R., & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, 12(6), 914-919.
- Snow, C.E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4(1), 1-22.
- So, C. & Best, C.T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, 53(2), 273-293.
- So, C., & Best, C.T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36, 195-221.
- Soderstrom, M., Ko, E.-S., & Nevzorova, U. (2011). It's a question? Infants attend differently to yes/no questions and declaratives. *Infant Behavior and Development*, 34, 107-110.

- Stager, C.L. & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381-382.
- 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*(2), 111-124.
- Tsao, F.-M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in Psychology*, doi:10.3389/fpsyg.2017.00558.
- Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C.T. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: language-specific developmental changes in speech perception abilities. *1994 International Conference on Spoken Language Processing* (pp. 1695-1698). Yokohama, Japan.
- Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic Segment Durations in Prosodic Research: A Practical Guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Methods in Empirical Prosody Research* (pp. 1-27). Berlin: de Gruyter.
- Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*, *4*(1), 1-22.
- Wang, Y., Behne, D., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, *25*, 449-466.
- van de Weijer, J. (2001). The importance of single-word utterances for early word recognition. *Proceedings of Early Language Acquisition (ELA2001)*. Lyon: France.
- Werker, J.F., & Tees, R.C. (1992). The organization and re-organization of human speech perception. *Annual Review of Neuroscience*, *15*, 377-402.

- Werth, A. (2011). *Perzeptionsphonologische Grundlagen der Prosodie. Eine Analyse der mittelfränkischen Tonakzentdistinktion*. Stuttgart: Franz Steiner Verlag.
- Wetterlin, A. (2007). *The Lexical Specification of Norwegian Tonal Word Accents*. Doctoral dissertation, University of Konstanz.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. *Proceedings of the Fifth International Conference on Language Resources and Evaluation LREC*. Genoa: Italy.
- Wu, X., Tu, J.-Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623-641.
- Wu, Z. (2000). From traditional Chinese phonology to modern speech technology: Realization of tone and intonation in spoken Chinese. *Proceedings of the International Conference on Spoken Language Processing (ICSLP 2000)*, Vol. 1, B1-B12.
- Yeung, H.H., Chen, K.H., & Werker, J. (2013.). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Appendix

Table 1. Acoustic measurements of the accent 1 and accent 2 tokens of *taag* used during habituation and test phase. Duration in seconds, timing of the pitch peak measured from word onset in seconds, pitch range in Hz.

Tone	Token	Duration	Timing pitch peak	Pitch range
1	1	0,63	0,24	231,29
1	2	0,61	0,24	164,78
1	3	0,61	0,25	258,69
1	4	0,71	0,28	222,83
1	5	0,62	0,23	174,33
2	1	0,62	0,2	156,48
2	2	0,62	0,2	84,41
2	3	0,61	0,18	100,2
2	4	0,67	0,2	79,1
2	5	0,64	0,2	110,82

Table 2: Number and percentage of excluded trials and participants per language and age group.

Trial exclusion n (%)	Dutch				Limburgian			
	6 mo.	9 mo.	12 mo.	Total	6mo.	9 mo.	12 mo.	Total
	1: LT < 2 s.	19 (5,7)	26 (7,5)	20 (6,4)	65 (6,5)	5 (3,8)	4 (3,3)	14 (6,5)
2: max. LT	-	-	-	-	-	-	-	-
3: > 1 restart	-	5 (1,4)	4 (1,3)	9 (0,9)	2 (1,5)	2 (1,7)	4 (1,9)	8 (1,7)
4: 3 restarts	2 (0,6)	10 (2,9)	-	12 (1,2)	1 (0,8)	-	3 (1,4)	4 (0,9)
5: ext. interference	1 (0,3)	-	-	1 (0,1)	-	1 (0,8)	2 (0,9)	3 (0,6)
6: crying/fussing	21 (6,3)	12 (3,4)	8 (2,6)	41 (4,1)	1 (0,8)	8 (6,7)	1 (0,5)	10 (2,1)
7: difficulty coding	-	2 (0,6)	2 (0,6)	4 (0,4)	-	-	4 (1,9)	4 (0,9)
8: equip./exp. error	-	-	-	-	-	-	1 (0,5)	1 (0,2)
Total excluded trials n (%)	43 (12,8)	55 (15,8)	34 (10,9)	132 (13,3)	9 (6,8)	15 (12,5)	29 (13,4)	53 (11,3)
Total n trials	336	348	312	996	132	120	216	468
Participant exclusion n (%)								
1: habituation	4 (8,9)	-	-	4 (3,2)	-	-	2 (7,1)	2 (3,7)
2: equip./exp. error	4 (8,9)	2 (4,5)	1 (2,7)	7 (5,6)	-	-	-	-
3: < 50% test trials	8 (17,8)	13 (29,5)	10 (27)	31 (24,6)	1 (8,3)	4 (28,6)	8 (28,6)	13 (24,1)
4: linguistic background	1 (2,2)	-	-	1 (0,8)	-	-	-	-
Total excluded participants n (%)	17 (37,8)	15 (34,1)	11 (29,7)	43 (34,1)	1 (8,3)	4 (28,6)	10 (35,7)	15 (27,8)
Total n participants	45	44	37	126	12	14	28	54

110 Setting the Tone

4. Distinctive function, distinctive processing? Native and non-native perception of Limburgian lexical tones by Limburgian and Dutch adults

Based on: Ramachers, S., Brouwer, S., & Fikkert, P. (in prep.). Distinct functions lead to distinct processing. Native and non-native perception of Limburgian lexical tones.

Abstract

The present study examined the perception of lexical tone in adult speakers of Limburgian, a restricted tone language. Their performance was compared to a control group of speakers of non-tonal Standard Dutch. In this way, we could investigate whether and how (1) the functional load of a lexical pitch contrast and (2) the different functions of pitch across languages influence word-level pitch processing. Hardly any cross-linguistic investigations exist that investigated pitch sensitivity in speakers of more restricted tone languages. In this study, tone perception was tested by means of a categorical AXB discrimination task. By manipulating the position of the tone contrast within our stimuli, we wanted to investigate whether Dutch adults' perception of non-native tones is influenced by native intonation categories. By including between-category and within-category tone variation, we examined whether Limburgian adults' perception of word-level pitch is more categorical than that of Dutch speaking adults. Results show that, overall, adult speakers of Limburgian outperform Dutch-speaking adults by showing better discrimination of Limburgian lexical tones. However, Dutch participants performed above chance, indicating that they still show some sensitivity to word-level pitch differences. We also show that Limburgian adults perform equally well in discriminating between- and within-category tone variation. The results are discussed against the background of the prosodic systems of Limburgian and Dutch.

4.1 Introduction

The languages of the world differ with respect to the functions they attribute to pitch (i.e., fundamental frequency or f_0 in Hz). Pitch can signal meaningful information at different levels of linguistic structure (i.e., at the word level or at the utterance level). These cross-linguistic differences in the functionality of pitch can influence the way in which pitch information is processed in speakers of different languages. This study tests the discrimination of a lexical tone contrast in adult speakers of Limburgian, which can be considered a restricted tone language, compared to a control group of adult speakers of a non-tone language (Dutch). The main aim is to see whether the differences in the functionality of pitch cause differences in pitch processing in these groups.

Linguists distinguish between two of the major uses of linguistic pitch as TONE and INTONATION. By some estimates, approximately 60-70% of the worlds' languages (Yip, 2002) have lexical tone. These languages use pitch patterns to distinguish words or grammatical morphemes the way vowels and consonants do in all languages. Some of these tone languages make very extensive use of lexical pitch differences. In Mandarin Chinese, for example, almost every syllable has to be pronounced with one of four tones (Duanmu, 2000). Other tone languages are more restricted in their use of lexically contrastive tones. These languages, for example Tokyo Japanese or Limburgian, have been referred to as RESTRICTED TONE LANGUAGES or PITCH-ACCENT LANGUAGES (cf., Hyman, 2009; Voorhoeve, 1973). We follow Yip (2002) and Hyman (2009) in their assumption that, be it to a greater or lesser extent, in all these languages pitch is necessary for determining the meaning of a word. Depending on the functional load of tone, which is partly based on the number of tonal minimal pairs in a specific language, a tonal system can be considered a more or less restricted tone language (Voorhoeve, 1973).

Intonation is used in all languages of the world (Gussenhoven, 2004; Singh & Fu, 2016; Yip, 2002). Speakers of non-tone languages such as English and Dutch have to learn that pitch is used at the utterance-level to distinguish between sentence types (e.g., questions and statements) or to signal emotions

(i.e., paralinguistic meaning). Furthermore, it can highlight aspects of grammatical structure, for example by marking focus (e.g., Gussenhoven, 2004), and it functions as one of the cues to word stress (e.g., Cooper, Cutler, & Wales, 2002; Cutler & van Donselaar, 2001).

Importantly, in languages like English and Dutch, pitch is not used to distinguish words – except inasmuch as pitch is involved in signaling word stress (e.g., Dutch *VOORKomen* ‘appear’ vs. *voorKOMEN* ‘prevent’),¹⁰ where pitch is only one of several correlated cues. The fact that pitch is not used to signal lexical distinctiveness in non-tone languages might prevent speakers of these languages from distinguishing monosyllables that differ in pitch only (Schaefer & Darcy, 2014).

4.1.1 How different functions of pitch affect pitch processing

Distinctive processing of functionally different pitch patterns has its roots in early infancy. Infants growing up with a tone language are well able to discriminate native lexical tones throughout their first year of life (e.g., Ramachers, Brouwer, & Fikkert, 2017; Chapter 3 of this thesis; Sato, Sogabe, & Mazuka, 2010; Tsao, 2017; Yeung, Chen, & Werker, 2013). Studies with infants acquiring a non-tone language have yielded mixed results (for a critical review, see Singh & Fu, 2016). Some seminal studies indicate that non-tone-learning infants show a decline in lexical tone discrimination around nine months (English and Yorùbá infants, Yorùbá tones: Harrison, 2000; English and Chinese infants, Thai tones: Mattock & Burnham, 2006; English and French infants, Thai tones: Mattock, Molnar, Polka, & Burnham, 2008; English, Cantonese and Mandarin infants, Cantonese tones: Yeung et al., 2013). Other more recent studies by Liu and Kager (2014), Chen and Kager (2016), and Ramachers et al. (2017; Chapter 3 of this thesis) investigated discrimination of Mandarin and

¹⁰ According to Cutler and van Donselaar (2001), Dutch has approximately 13 minimal stress pairs.

Limburgian tones in Dutch infants and found a maintained sensitivity to non-native pitch patterns, partly depending on the contrast tested. How can we explain this maintained sensitivity?

From previous studies on the perception of both segmental and suprasegmental contrasts in both infants and adults, we have learned that the perception of phonetic contrasts does not always follow the 'all-or-nothing'-principle (e.g., Best, McRoberts, & Sithole, 1988; Tong, Lee, Lee, & Burnham, 2015). Certain non-native contrasts remain perceivable into adulthood. This can be explained by their acoustic salience and/or the extent to which they are similar to or different from native sounds (Best, 1995). It is thus the combined influence of native language (henceforth L1) experience as well as more general factors like salience that determine speech sound discrimination. The same factors have also been shown to have an influence on lexical tone perception. Earlier studies on native and non-native tone perception in both infants and adults support the assumption that cross-linguistic differences in the functions attributed to pitch lead to differences in the processing of pitch (e.g., Mattock et al., 2008; Braun & Johnson, 2011; Braun, Galts, & Kabak, 2014; Hallé, Chang, & Best, 2004; Schaefer & Darcy, 2014). However, it has also led to the insight that a simple binary distinction between tone languages on the one hand and non-tone languages on the other is too simplistic to account for the attested patterns, in that speaking a tone-language apparently is not a prerequisite for being able to perceive tones (e.g., Braun & Johnson, 2011; Braun et al., 2014; Schaefer & Darcy, 2014). Despite the advantage in lexical tone discrimination enjoyed by tone language speakers, a vast number of studies have demonstrated sensitivity to tones in speakers of non-tone languages, albeit to a lesser extent (Braun & Johnson, 2011; Braun et al., 2014; Broselow, Hurtig, & Ringen, 1987; Burnham et al., 2015; Gandour, 1983; Hallé et al., 2004; Lee, Vakoch, & Wurm, 1996; Qin & Mok, 2011; Schaefer & Darcy, 2014; So & Best, 2008, 2010, 2011, 2014; Tong et al., 2015; Wayland & Guion, 2003; Xu, Gandour, & Francis, 2006).

For example, Hallé et al. (2004) tested perception of Mandarin tones in Mandarin and French speakers. In a non-categorical AXB discrimination task¹¹, French listeners showed significantly lower mean discrimination performance ($M = 74\%$) than Mandarin participants ($M = 88\%$). Still, the French listeners show a non-negligible sensitivity to tone contour variations. The authors propose that they do not process them in a contrastive, but rather in a more psychophysical way because the French lack of native tonal referent categories. Hallé et al. conclude that their listeners' performance is best explained by the different functions of pitch in their L1 phonologies. They state that Mandarin tones are phonemic for Mandarin listeners but not for French listeners.

In a similar vein, So and Best (2014) tested how native speakers of Australian English and French, non-tone languages that differ in the complexity of their prosodic system, perceived six pairs of Mandarin tones in a categorical AXB discrimination task featuring target words in a sentence context. The French listeners' mean discrimination performance ($M = 70\%$) was slightly yet significantly higher than that of the English listeners ($M = 66\%$). Moreover, accuracy on specific tone pairs differed across groups. So and Best attribute the different behavior of the two groups to the different prosodic systems of French and English.

Tong et al. (2015) compared perception of Cantonese tones in Cantonese-English bilingual adults to English monolingual adults. Despite the

¹¹ In an AXB discrimination task, listeners are presented with triads of stimuli. For each triad, they have to decide whether the first or the third stimulus (the standards, belonging to two different types, for example two distinct tones) resembled the second stimulus (the target, also belonging to one of these types) most. The paradigm can be categorical or non-categorical. In a categorical task, the matching standard and target stimulus are different tokens of the same type, i.e., the listener has to abstract away from phonetic variation in order to uncover the underlying type. In a non-categorical task, the matching standard and target stimulus are phonetically identical, i.e., the listener does not have to abstract away from phonetic variation. Categorical tasks are considered higher-level tasks and are more suitable to tap into language-specific processing. Non-categorical tasks can elicit lower-level acoustic-phonetic processing.

significant difference in accuracy between the tone ($M = 95\%$) and non-tone language listeners ($M = 90\%$), the monolingual English speakers still performed remarkably well. Correlating their discrimination results with findings from an extensive acoustic analysis of Cantonese lexical tones and English lexical stress, Tong and colleagues conclude that English listeners might have called upon an acoustic correlate of lexical stress (in this case spectral tilt, i.e. the overall distribution of spectral energy) as a cue to perceive Cantonese tones. Cantonese-English adults gave much less weight to this cue. This is in line with earlier findings by among others Gandour (1983) and Gandour and Harshman (1978) that native prosody determines the weighting of acoustic cues in the perception of non-native prosodic categories.

Braun and Johnson (2011) also showed that the linguistic function of native pitch contrasts guides listeners' pitch processing. Speakers of Dutch participated in a non-categorical ABX discrimination paradigm with Mandarin tones 2 (rising) and 4 (falling). By manipulating the position of tone 2 in their target disyllabic pseudo-words, the authors created a condition in which pitch carried no linguistic information for the Dutch listeners (tone 2 in pre-final position) and a condition where pitch signaled post-lexical information (tone 2 in final position). A rise on an utterance-final syllable signals an interrogative contour in Dutch and thus signals meaningful linguistic information for Dutch listeners. In congruent trials, the target X (the third stimulus) matched the standard A or B both at the segmental and suprasegmental level (e.g., *mo₂va₂-no₄ba₄-mo₂va₂*; subscripts indicating the respective tones). Tone 4 was always realized on the final syllable. In incongruent trials, one standard matched the target X along the suprasegmental but not the segmental dimension, and the other matched X along the segmental but not the suprasegmental dimension (e.g., *mo₂va₂-no₄ba₄-mo₂va₄*). Dutch listeners' reaction times to incongruent trials relative to congruent trials were significantly slower in the condition where tone 2 was realized on the final syllable compared to when it was realized on the pre-final syllable, implying that the Dutch had paid significantly more attention to pitch when it appeared on the final syllable. A replication with

native Mandarin speakers showed that they experienced even more difficulty (i.e., slower reaction times) in incongruent trials. This is attributable to the fact that the pitch patterns, in both conditions, signal lexical differences for speakers of Mandarin. The authors conclude that the function of pitch in the L1 determines the extent to which listeners attend to word-level pitch: If the listeners' L1 has lexical tone, they attend more closely to word-level pitch patterns than listeners whose L1 does not have lexical tone.

In sum, previous studies on adults' perception of non-native lexical tones have shown that prosodic experience from listeners' L1 guides their perception of pitch patterns to a considerable extent. A pitch movement that signals a lexical distinction in one language can signal a post-linguistic or paralinguistic distinction in another. As a result, non-tone language speakers may still attend to pitch, though perhaps less closely than tone language speakers.

The studies reported above mainly focused on well-known tone languages like Mandarin Chinese that make very extensive use of tone. To date, hardly any cross-linguistic investigations exist that investigated pitch sensitivity in adult speakers of more restricted tone languages. To extend our insights into the influences of the functional load of a lexical pitch contrast and of the different functions of pitch on word-level pitch processing, the present study looks at tone perception in adult speakers of a group of restricted tone languages, Limburgian dialects of Dutch. We compare lexical pitch perception in adult speakers of Limburgian to adult speakers of non-tonal Dutch, to see whether the differences in the functionality of pitch in these languages brings about differences in pitch processing in speakers of these languages. For a detailed description of the use of pitch in Limburgian, the reader is referred to Chapter 2 of this thesis.

In line with previous work on non-native tone perception, we assume that Dutch listeners will perceive non-native Limburgian tones according to the properties of the prosodic system of Dutch. Several theoretical models have

been developed to account for listeners' reliance on L1 prosodic categories when perceiving non-native prosodic contrasts, for example the PERCEPTUAL ASSIMILATION MODEL FOR SUPRASEGMENTALS (PAM-S; So & Best, 2014 and references therein) and the FUNCTIONAL PITCH HYPOTHESIS (Schaefer & Darcy, 2014). These models will be explained in the following section.

4.1.2 Theoretical models for the naïve perception of non-native pitch contrasts

PAM-S is an extension of the original PERCEPTUAL ASSIMILATION MODEL (PAM; Best, 1995). The model suggests that listeners make use of their native prosodic categories when perceiving pitch movements (e.g., lexical tones) from a foreign language. In PAM-S, a non-native prosodic category may be perceptually assimilated as either a CATEGORIZED or an UNCATEGORIZED prosodic category. A non-native tone is categorized if a listener perceives it as an instance of a native prosodic category. For example, English listeners assimilate Mandarin tone 4 (falling) as their statement intonation category (e.g., Broselow et al., 1987; So & Best, 2008). Recall that Braun and Johnson (2011) suggested that their Dutch listeners probably perceived utterance-final Mandarin tone 2 (rising) as question intonation. With respect to the present study, Dutch listeners could perceive Limburgian accent 1 (H*L) as statement intonation (H*L L%). A non-native tone remains uncategorized if its phonetic features fall within native phonological space but in between two or more native prosodic categories (So & Best, 2014).

The degree of similarity between non-native and native categories determines what type of assimilation takes place. PAM-S applies the same assimilation types as the original PAM framework. TWO-CATEGORY (TC) assimilation occurs when two non-native prosodic categories are mapped onto two separate native prosodic categories. Discrimination in these cases is expected to be excellent. SINGLE-CATEGORY (SC) assimilation occurs when two non-native prosodic categories assimilate equally well or poorly to a single

native prosodic category. In this case, poor discrimination is predicted. However, if one of the two native categories assimilates better than the other, discrimination is moderate to very good. This is called a CATEGORY-GOODNESS (CG) assimilation. UNCATEGORIZED-CATEGORIZED (UC) assimilation arises when one non-native category is assimilated to a native prosodic category and the other is left uncategorized. Discrimination may range from poor to good depending on the perceived similarity between the two tones. The same holds for UNCATEGORIZED-UNCATEGORIZED (UU) assimilations (So & Best, 2014). The only difference with the original PAM framework is that non-native prosodic categories, in contrast to certain non-native segmental categories, cannot be considered NON-ASSIMILABLE (NA). Since pitch variation is found in all natural languages at the sentential intonation level, pitch patterns in speech are thus not completely foreign to any non-tone language speaker. They cannot be perceived as non-speech, except when they are freed from their linguistic context and for example presented as pure tones or musical pitch (Burnham & Mattock, 2007; Hallé, et al., 2004; Singh & Fu, 2016). The Limburgian tone contrast could be a case of TC assimilation for Dutch listeners if they perceive accent 1 (H*L) as a statement (H*L L%) and accent 2 (H*LH) as a question (H*L H%). If so, they are expected to show excellent discrimination.

Another model that tries to explain differences in pitch processing with reference to native prosody is the FUNCTIONAL PITCH HYPOTHESIS (Schaefer & Darcy, 2014). This model tries to explain naïve pitch perception with reference to (1) the functionality of pitch in the L1, defined as the degree to which linguistic pitch differentiates lexical items in the L1, and (2) the size of the linguistic domains in which pitch is realized. In Mandarin Chinese, pitch contrasts manifest themselves within the syllable. For this reason, Mandarin listeners are expected to show high accuracy in discriminating syllable-level tonal contrasts. In Japanese and English (a pitch-accent language and a lexical stress language, respectively), pitch contrasts spread over multiple syllables. Therefore, Japanese and English listeners are expected to be less accurate in

perceiving pitch at the syllable level than speakers of a tone language. Using an AXB discrimination task with monosyllables differing only in Thai tones, Schaefer and Darcy showed that Mandarin listeners indeed discriminated tonal contrasts with higher accuracy than the other groups. They significantly outperformed English participants, but were only marginally ($p = 0.087$) more accurate than the Japanese group. The authors conclude from these findings that the specific prosodic domain in which pitch differentiates lexical items in the L1 constrains performance in a discrimination task. Unfortunately, the authors do not mention whether the difference between Japanese and English participants was significant. Together with the marginally significant (in case of adult data it is perhaps better to speak of a non-significant) difference in performance between Mandarin and Japanese participants, we doubt that only the size of the domain in which pitch is realized in the L1 is decisive for non-native tone perception. Rather, it appears to be a combined effect of whether pitch is lexically contrastive in the L1 (which is the case in Mandarin and Japanese, but not in English) and the temporal domain over which native pitch patterns are realized (within a single syllable in Mandarin, over multiple syllables in Japanese, at the utterance level in English).

4.1.3 The present study

To investigate whether differences in the functionality of pitch in Limburgian and Dutch cause differences in pitch processing, and to see whether adults show the same patterns as found in the discrimination study with Limburgian and Dutch infants (Ramachers et al., 2017; Chapter 3 of this thesis), we ran a forced-choice AXB categorical discrimination task. We chose a categorical AXB task instead of an AX discrimination task for three reasons. First, an AXB task is not as strongly influenced by low-level acoustic comparisons and thus more likely to elicit language-specific listening. Secondly, it has a lower response bias than an AX task (e.g., McGuire, 2010; Strange & Shafer, 2008). Thirdly, the use of this paradigm is consistent with prior PAM investigations (e.g., Best,

McRoberts, & Goodell, 2001; Hallé & Best, 2007; Harnsberger, 2001; So & Best, 2014; Tyler, Best, Faber, & Levitt, 2014). In this AXB task, we presented pseudo-words, which do not require lexical access and are therefore suited to compare the perception of speech stimuli by listeners from different languages (Braun & Johnson, 2011).

We included trials that featured tonal *between*-category variation and trials with *within*-category variation. Trials with between-category variation ought to test lexical tone discrimination. Within-category variation 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 native lexical tone categories, Limburgian listeners are expected to be worse in perceiving within-category than between-category pitch differences. To efficiently categorize f_0 contours into tones, they must ignore irrelevant tonal variations (i.e., within-category variations) (Hallé et al., 2004; Xu et al., 2006). As such, we expect our Limburgian listeners to be more accurate and faster in between-category variation trials than within-category variation trials. However, unlike speakers of typically studied dense tone languages, Limburgians are exposed to a considerable amount of surface variability in pitch contours because of the intricate interaction between lexical and intonational tones in Limburgian. From any pitch contour within a stressed syllable, Limburgians have to abstract both its intonational and lexical meaning. As a consequence of this lifelong experience, they might just as well yield a greater sensitivity to pitch differences in general and perform well in within-category pitch discrimination in contrast to tone language speakers in earlier studies. Our Dutch listeners, not being constrained by lexical tone categories, are not expected to perform differently in between- and within-category variation trials and treat any pitch pattern as equally (un)meaningful.

To see whether Dutch adults' perception of Limburgian tones is influenced by native prosodic categories, we manipulated the position of the tones in our pseudo-word stimuli comparable to the Braun and Johnson (2011)

study. Half of the trials featured monosyllabic pseudo-words where the tonal contrast is realized in utterance final position. In this position, the Dutch listeners were expected to interpret the tones as instances of native post-lexical intonation categories. As mentioned above, the citation forms of the Limburgian tones might be interpreted as Dutch statement and question intonation patterns. The tones in monosyllables were thus expected to elicit TC assimilation, leading to excellent discrimination performance. In a study with 6- to 12-month-old Limburgian and Dutch infants, Dutch infants were successful in discriminating Limburgian tones in monosyllables throughout their first year of life (Ramachers et al., 2017; Chapter 3 of this thesis). By testing adult speakers of Dutch, we can find out whether this discrimination ability declines.

The other half of the trials consisted of disyllabic pseudo-words with trochaic stress. In these pseudo-words, the tones were realized on the penultimate syllable and thus occurred in non-final position. Crucially, in this position, the tones supposedly do not signal a linguistically meaningful contrast for speakers of Dutch (e.g., Braun & Johnson, 2011). As a consequence, we expected Dutch listeners to show poorer discrimination performance in these trials as compared to monosyllabic trials. Moreover, it is important to note that the tones' surface realizations in the disyllabic stimuli differed from their realizations in the monosyllabic stimuli. Due to differences in the prosodic contexts, the tones surface as a falling tone (accent 1) and a level tone (accent 2) in disyllables with trochaic stress, whereas they surface as a falling tone (accent 1) and a fall-rise (accent 2) in monosyllables (see Figure 2 below). For details on the tone-intonation interaction in Limburgian we refer the reader to Chapter 2, section 2.2.4.

Dependent measures were accuracy (proportion of correct responses) and reaction times in ms. Reaction times are assumed to reflect difficulties in decision making (Braun & Johnson, 2011; Gili-Fivela, 2012; Schneider, Dogil, & Möbius, 2011) and could offer a more nuanced interpretation of the data.

4.2 Method

4.2.1 Participants

Limburgian and Dutch adults were recruited and tested in public libraries in Roermond (Limburg, the Netherlands) and Nijmegen (Gelderland, the Netherlands). The Limburgian listeners ($N = 20$, 10 males) ranged in age from 23 to 67 years ($M = 49$ years). All Limburgian participants were born and raised in the East-Limburgian dialect region and lived there at the time of testing. People that reported to speak a dialect other than one from the East-Limburgian dialect region were excluded from participation. The Dutch listeners ($N = 19$, 8 males) ranged in age from 19 to 66 years ($M = 45$ years). None of our Dutch participants had weekly contact with people speaking a Limburgian dialect in their presence. Moreover, none of them grew up or lived in the province of Limburg.

None of the participants reported to have command of a tonal L2 like Mandarin Chinese or Swedish. They all reported normal hearing and no speech, language, or attention deficits. Because of the fact that musical experience can have an influence on pitch processing (e.g., Burnham & Brooker, 2002; Burnham, Brooker, & Reid, 2015), we tried to keep the relative amount of musically trained individuals comparable across groups. In the end, six of the Limburgian participants (30%) and seven of the Dutch participants (37%) reported to have had over 3 years of musical training. Four additional participants (1 Dutch, 3 Limburgian) were excluded from the analysis because they turned out not to meet our linguistic criteria after completion of the background questionnaire that was administered after the experiment. Participants took part either voluntarily or for a small amount of money.

4.2.2 Stimuli

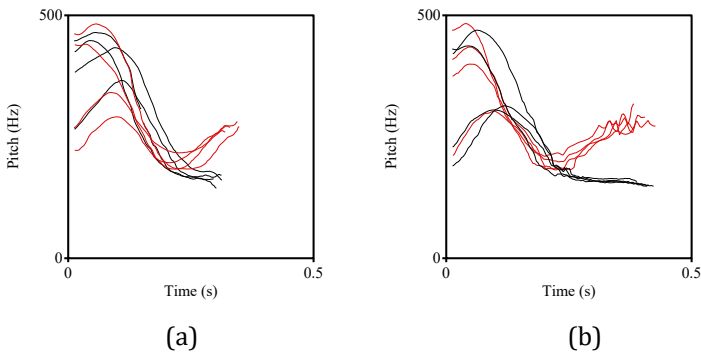
Test stimuli were four pseudo-words that could be pronounced with accent 1 and accent 2: Two monosyllables, *taag*_{1/2} [ta:ç]¹² and *moon*_{1/2} [mo:n], and two disyllables, *zeeve*_{1/2} [ke:və] and *perger*_{1/2} [pɛɾɣɔɾ]. These pseudo-words were created by substituting the onsets of existing tonal minimal pairs in East-Limburgian: *daag*₁ [da:ç] ‘days’, *daag*₂ [da:ç] ‘day’, *sjoon*₁ [ʃo:n] ‘beautiful’, *sjoon*₂ [ʃo:n] ‘shoe’, *zeeve*₁ [ze:və] ‘to sift’, *zeeve*₂ [ze:və] ‘seven’, *erger*₁ [ɛɾɣɔɾ] ‘worse’, and *erger*₂ [ɛɾɣɔɾ] ‘dispute/argument’.¹³ 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 surface as a falling tone and a level tone, respectively, in the disyllabic stimuli (see Figure 2 below). The pseudo-words complied with the phonotactic rules of both Limburgian and Dutch, but none of them carried any lexical meanings.

Multiple tokens of the pseudo-words were recorded by a female native speaker from the East-Limburgian dialect region [SR]. Stimuli were recorded in a sentence context to avoid listing intonation. To avoid any more prosody-dependent differences in the phonetic realization of the tones than those resulting from the position-manipulation, tokens were only recorded in a focal-declarative context. Recordings took place in a sound-attenuated studio at Radboud University, Nijmegen. Stimuli were digitized onto a PC using Adobe Audition (version CS6, 44.1 kHz). For each pseudo-word, four acoustically different tokens were selected based on careful listening by a native speaker [SR] to introduce minor phonetic variability. We followed the guidelines presented in Turk, Nakai, and Sugahare (2006) for excising the pseudo-word stimuli from their carrier sentences. We controlled for perceptual loudness by equalizing all selected tokens to 65 dB(A). Stimulus excision and acoustic

¹² Subscripts indicate accents 1 and 2.

¹³ The item *taag* was also included because this study is a continuation of the infant discrimination study reported in Ramachers et al. (2017 / Chapter 3 of this thesis) where *taag* also served as a stimulus.

measurements were done in PRAAT (version 5.3.22, Boersma & Weenink, 2012). For the tone-bearing portion (TBP) of our stimuli we measured maximum and minimum f_0 , f_0 excursion, average f_0 , endpoint f_0 and duration. We also measured the full duration of each token (see Table 1 in the Appendix). Measurements were done manually, taking auditory as well as spectral properties into account. A series of independent t -tests revealed that accents 1 and 2 in monosyllables differed significantly with respect to minimum pitch within the TBP ($t(14) = -7.43, p < .001$) and with respect to the f_0 offset value measured at the end of the TBP ($t(14) = -14.75, p < .001$). Accents 1 and 2 in disyllables could be differentiated based on minimum pitch in the TBP ($t(14) = -8.02, p < .001$), pitch range in the TBP ($t(7) = 6.02, p < .001$) and f_0 offset ($t(8) = -8.50, p < .001$). No other significant differences were found (all $ps > .05$). See Figure 2 for the f_0 contours of accents 1 and 2 on the TBPs of the four pseudo-words (panels a-d).



126 Setting the Tone

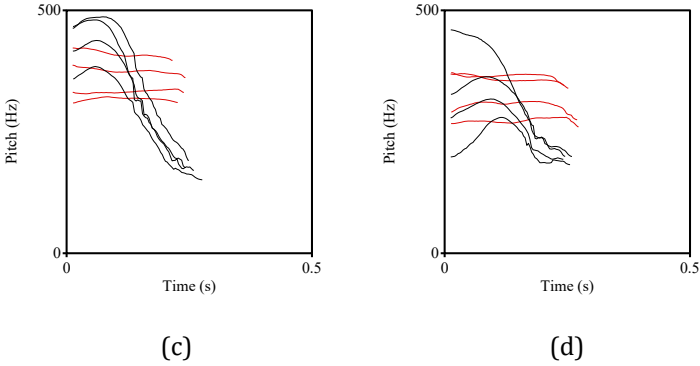


Figure 2: F₀ contours on the TBPs of our four accent 1 (solid black line) and accent 2 (solid red line) tokens of *taag* (a), *moon* (b), *keeve* (c) and *perger* (d).

During the test phase of the experiment, participants proceeded through a total number of 192 trials. The test phase consisted of 96 test trials and 96 filler trials. Among the 96 test trials, we introduced two conditions. Table 1 provides an overview of these conditions including example stimuli.

Table 1. Overview of the different trial types in the AXB discrimination task.

Trial type	A	X	B	Correct
Between-category pitch variation	taag _{1a} *	taag _{1b}	taag _{2a}	A
	perger _{2b}	perger _{1a}	perger _{1b}	B
Within-category pitch variation	moon _{2a}	moon _{2a}	moon _{2b}	A
	keeve _{1a}	keeve _{1b}	keeve _{1b}	B

*The subscripts in columns A, X and B refer to the tone (1 or 2) and the token (a or b).

The first condition constituted 64 between-category variation trials (BETWEEN). In this condition, tonal minimal pairs (which share the same segmental structure) served as stimuli to test lexical tone discrimination.

Participants had to compare the pitch pattern of token X (e.g., accent 1) to the pitch patterns of token A (e.g., accent 2) and token B (e.g., accent 1) and had to evaluate which of them was closer to the target X. Four triplets were constructed for each of the four possible triad orders (AAB, ABB, BAA, BBA), resulting in 16 different trials for each of the four pseudo-words. This made up a total number of 64 between-category variation trials (4 triad orders x 4 triplets x 4 pseudo-words). Crucially, half of the between-category trials featured the monosyllabic pseudo-words and half of the trials featured the disyllabic pseudo-words. We counterbalanced the order of presentation of both tones within a trial to control for potential perceptual asymmetries. Half of the trials included a change from accent 1 to accent 2 (accent 1 – accent 1 – accent 2 or accent 1 – accent 2 – accent 2) and the other half included a change from accent 2 to accent 1 (221 or 211). 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 (e.g., Best et al., 1988; Best et al., 2001; Polka, 1995; Polka & Bohn, 1996; Strange & Shafer, 2008; Tyler et al., 2014).

The second condition constituted 32 within-category variation trials (WITHIN). The three stimuli in these triads showed the *same* tonal pattern with only subtle differences in their actual instantiation of the pitch contour. In these trials, the X token was physically identical to the A or B token.

For the 96 filler trials, another eight pseudo-words were created that differed either in one vowel or one consonant from the aforementioned test items: *meen* [me:n], *moop* [mo:p], *tag* [taç], *kaag* [ka:ç], *kaave* [ka:və], *peeve* [pe:və], *pirger* [pɪrɣər] and *pertter* [pɛrtər]. All vowels and consonants are phonemes in both Dutch and Limburgian and were thus expected to be easily perceivable by both listener groups. The pseudo-words were consistently pronounced with accent 1, which resembles the intonation of words in isolation in Dutch.

The 96 filler trials were subdivided into three different types of fillers. Type 1 fillers introduced 32 trials where the target stimulus differed from A or B with respect to a vowel (e.g., *perger-pirger-pirger*), thus introducing vowel-minimal-pairs. Type 2 fillers introduced 32 consonant-minimal pairs (e.g., *moon-moop-moop*), whereas type 3 fillers consisted of 32 non-minimal pair triplets (e.g., *moon-taag-taag*), where X would be a completely different word than, in this specific example, A. Fillers were added to distract participants from the exact purpose of the experiment (lexical tone perception) and to give participants a sense of success and remain motivated in case they experienced difficulties with the pitch stimuli.

In accordance with earlier studies on non-native speech perception using the AXB task (e.g., Best et al., 2001; Hallé & Best, 2007; Harnsberger, 2001, Tyler et al., 2014), we chose for a relatively long inter stimulus interval (ISI; > 1000 ms) to increase the memory load of the task. Previous work has shown that, as the cognitive demands of the task increase, native-language perceptual patterns are more likely to be demonstrated (Harnsberger, 2001; Polka, 1995, and references therein; Strange & Shafer, 2008; Wayland & Guion, 2003; Werker & Tees, 1984; but cf. Dupoux, Peperkamp, & Sebastián-Gallés, 2001). At shorter ISI's, judgments can be based on auditory memory (i.e., psychoacoustic information) instead of on language specific phonetic or phonological information. At longer ISI's (> 500 ms), participants have more time to categorize the stimuli (e.g., Gerrits & Schouten, 2004; van Hessen & Schouten, 1992). We opted for an ISI of 1000 ms plus the duration of the preceding stimulus, leading to ISI's ranging from 1460 ms to 1700 ms ($M = 1600$ ms, $SD = 64$ ms). We implemented these minimally variable ISI's to allow our participants a processing time that compensates for the relative duration of the stimuli. The inter-trial interval was set to 2000 ms.

Trials were presented in pseudo-random order, with the restrictions that trials would not have the same correct answer more than three times in a row, and that test trials featuring the same pseudo-word would not appear more than three times in a row to avoid perceptual confusion.

4.2.3 Apparatus and procedure

The experiment was run on a laptop (HP Probook 6750B) using Presentation® software (Version 18.2 02.18.16, www.neurobs.com). Stimuli were presented through noise-cancelling headphones (Sennheiser HME 110). Participants were instructed that they would hear sequences of three stimuli and had to judge whether the second token (X) most resembled the first (A) or third (B) token. During presentation of the auditory stimuli, a fixation cross was presented on the screen. Time-locked to the offset of the third stimulus, the cross disappeared and the answer options '1' and '3' were presented on the screen. After each trial, participants had to press a designated button labeled '1' or '3' on a button box¹⁴ as fast as possible, but not until they had heard all three stimuli in a trial and the answer options had appeared on the screen. After the instructions, participants went through six practice trials that featured different trial types. After the practice phase, participants were allowed to ask questions before they proceeded to the actual experiment. They were told that their task during the main experiment was the same as in the practice phase, only that it would be more difficult.

Participants were allowed to pause three times during the experiment, i.e. after every 48 trials. The experiment lasted approximately 25 minutes. After the experiment, participants completed a language background questionnaire.

4.3 Results

4.3.1 Introduction to the data analysis

Two different analyses were conducted. The first analysis pertained to accuracy (proportion of correct responses) and reaction times (RTs) in between-category variation trials (BETWEEN) to test discrimination of the Limburgian

¹⁴ BITS1: developed by the technical support group of the Faculty of Social Sciences, Radboud University (see <http://tsgdoc.socsci.ru.nl/index.php?title=ButtonBoxes>).

lexical tone contrast and to see whether Dutch listeners' sensitivity to word level pitch differs as a function of the position of the pitch pattern (final or non-final). A mixed ANOVA with CONDITION (monosyllabic, disyllabic) and presentation ORDER (112/122 - 221/211) as within-subjects variables and LANGUAGE (Limburgian, Dutch) as a between-subjects variable was conducted. The second analysis compared accuracy and RT measures in trials featuring between-category variation (BETWEEN) to performance in trials featuring within-category variation (WITHIN) in a mixed ANOVA with TRIALTYPE (between-category variation, within-category variation) as a within-subjects variable and LANGUAGE as a between-subjects variable. The second analysis seeks to explore whether Limburgian participants' perception of lexical tone is more categorical than that of Dutch participants or whether Limburgians show high sensitivity to pitch in general because of their lifelong experience with a complex prosodic system.

RTs for correct responses only (= 78% overall) were included and measured from the offset of the third token (i.e., B). Participants with accuracy scores or RTs that were more than 2 SDs above or below the mean per language group per condition were excluded from the analysis (e.g., Schaefer & Darcy, 2014; Schneider et al., 2011; Tong et al., 2015). For a detailed overview of participant exclusion, see Table 2 in the Appendix.

4.3.2 Between-category variation analysis

Figure 3 shows the accuracy results from Limburgian and Dutch participants in mono- and disyllabic trials featuring between-category variation.

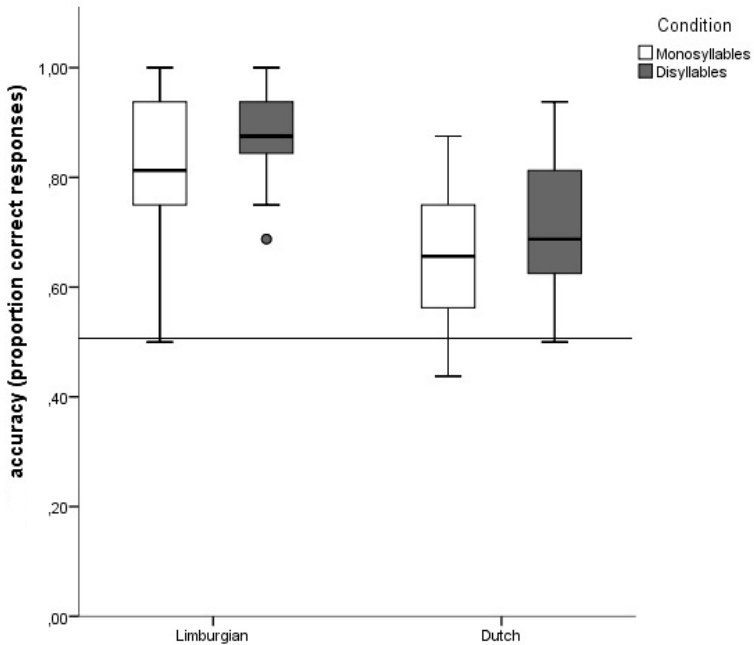


Figure 3. Boxplot showing the median and interquartile range of accuracy scores for Limburgian and Dutch participants in mono- and disyllabic trials featuring between-category variation. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers. The solid horizontal line indicates chance level (.50).

Performance in all conditions of interest was first compared to chance (.50) by means of one-sample *t*-tests. Both Limburgian ($N = 18$) and Dutch participants ($N = 18$) performed significantly above chance in all conditions (all p 's < .001, Cohen's d ranging from 2.38 to 4.95 for Limburgian participants and from 1.31 to 2.26 for Dutch participants). This shows that both Limburgian and Dutch adults were able to discriminate the East-Limburgian lexical tones in all the conditions tested.

The mixed ANOVA showed a significant main effect of LANGUAGE, $F(1,34) = 43.70$, $p < .001$, $\eta_p^2 = .56$, indicating that Limburgian participants ($M = .86$, $SE = .02$) performed significantly better than Dutch participants ($M = .69$, $SE = .02$). 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 have 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.

Data of the Limburgian participants revealed a main effect of CONDITION, $F(1,17) = 7.76, p = .01, \eta_p^2 = .31$. Limburgians perform better in disyllabic trials ($M = .89, SE = .02$) than in monosyllabic trials ($M = .84, SE = .03$). This effect has to be interpreted against the background of a significant CONDITION x ORDER interaction, $F(1,17) = 6.88, p = .02, \eta_p^2 = .29$. To be able to interpret this interaction we carried out paired-samples *t*-tests for each level of the ORDER variable on performance in monosyllabic and disyllabic trials. If first presented with an accent 2 token (221/211), 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 $d = -0.77$. However, if Limburgians first heard an accent 1 token (112/122), their performance in disyllabic and monosyllabic trials did not differ significantly ($M_{di} = .89, SD_{di} = .10; M_{mono} = .86, SD_{mono} = .12$), $t(17) = -1.03, p > .05$, Cohen's $d = -0.24$. Thus, the effect of CONDITION only arises in trials where Limburgians are first presented with an accent 2 token, as shown in Figure 4.

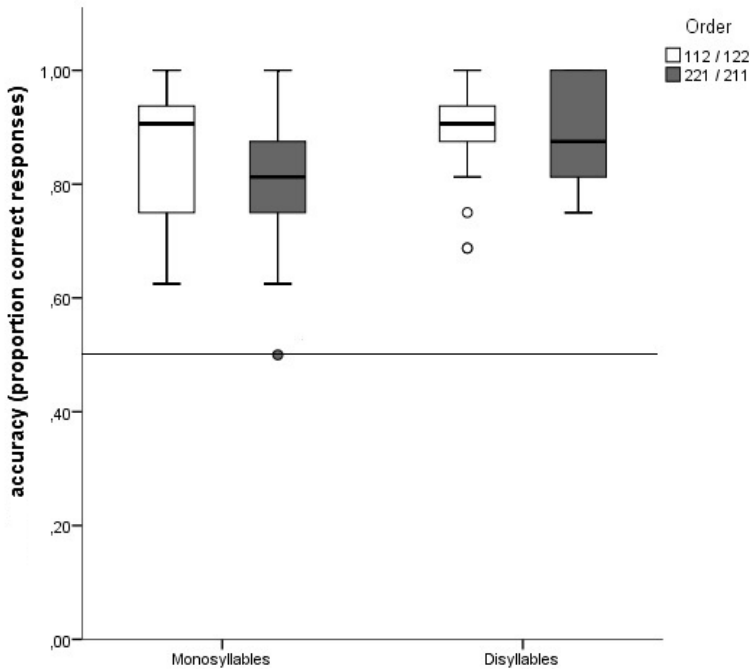


Figure 4. Boxplot showing the median and interquartile range of accuracy scores for Limburgian participants in mono- and disyllabic trials under the influence of order of tone presentation. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers. The solid horizontal line indicates chance level (.50).

The repeated-measures ANOVA on the Dutch participants' data revealed a main effect of CONDITION, $F(1,17) = 5.43$, $p = .03$, $\eta_p^2 = .24$. Like the Limburgian subjects, the Dutch subjects also 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 $> .05$).

The same analyses were run on the RT data. The RT data partly paralleled the accuracy data by showing a main effect of CONDITION ($F(1,31) = 8.04$, $p = .008$, $\eta_p^2 = .21$). Participants were faster on disyllabic ($M = 519.16$, $SE = 40.74$) than monosyllabic trials ($M = 600.48$, $SE = 51.14$). However, in contrast with the accuracy data, this difference was only significant in trials that started with an accent 1 stimulus as revealed by a significant CONDITION by ORDER

interaction ($M_{di} = 486.77$, $SD_{di} = 221.42$; $M_{mono} = 617.64$, $SD_{mono} = 296.87$; $t(32) = 4.12$, $p < .001$, Cohen's $d = 0.72$). There were no main effects or interactions with LANGUAGE (all p 's $> .05$).

4.3.3 Between-category variation vs. within-category variation analysis

Dutch and Limburgian participants' performance in between- vs. within-category variation trials is shown in Figure 5.

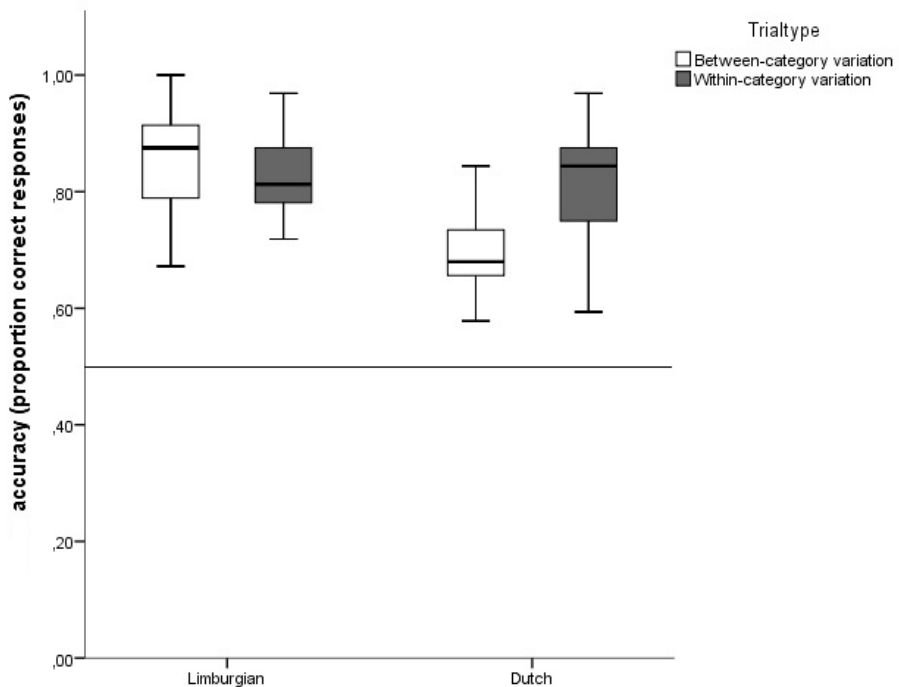


Figure 5. Boxplot showing the median and interquartile range of Limburgian and Dutch participants' accuracy scores in between- vs. within-category variation trials. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers. The solid horizontal line indicates chance level (.50).

Accuracy in both trial types (between-category variation, within-category variation) was compared to chance (.50) by means of one-sample t -

tests for both language groups separately.¹⁵ Limburgian ($N = 19$) and Dutch participants ($N = 18$) performed significantly above chance in both conditions (all p 's $< .001$, Cohen's d ranging from 3.60 (between-category variation) to 4.85 (within-category variation) for Limburgian participants and ranging from 2.65 (between-category variation) to 2.75 (within-category variation) for Dutch participants). These outcomes show that both Limburgian and Dutch adults were able to discriminate a variety of pitch differences.

Next, a mixed ANOVA with TRIALTYPE as a within-subjects variable and LANGUAGE as a between-subjects variable was conducted on the accuracy data. A significant main effect of LANGUAGE was detected, $F(1,35) = 15.31$, $p < .001$, $\eta_p^2 = .30$, indicating that Limburgians performed better ($M = .84$, $SE = .02$) than Dutch participants ($M = .75$, $SE = .02$) overall. The analysis also yielded a main effect of TRIALTYPE, $F(1,35) = 8.90$, $p = .005$, $\eta_p^2 = .20$, demonstrating that participants score better in within-category variation trials ($M = .83$, $SE = .02$) than in between-category variation trials ($M = .77$, $SE = .01$). Moreover, we found a significant TRIALTYPE x LANGUAGE interaction, $F(1,37) = 17.74$, $p < .001$, $\eta_p^2 = .30$. To break down this interaction, we performed paired-samples t -tests for each language group on their performance in between-category variation trials as compared to within-category variation trials. Dutch participants 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 $d = 0.11$). The main effect of TRIALTYPE is thus carried by the Dutch participants.

To take a closer look at the effect of LANGUAGE, we performed independent t -tests on both trial types (between-category and within-category variation). For between-category items, Limburgian participants performed

¹⁵ Note that this analysis was run on a slightly different sample ($N = 37$) due to the 2 SD exclusion criterion that was applied to each analysis and condition separately.

significantly better ($M = .85$, $SD = .10$) than Dutch participants ($M = .69$, $SD = .07$), $t(35) = -5.57$, $p < .001$, Cohen's $d = -1.84$. In within-category trials, both participant groups achieved similar results ($M_{Limburgian} = .84$, $SD_{Limburgian} = .07$; $M_{Dutch} = .81$, $SD_{Dutch} = .11$), $t(35) = -.80$, $p > .05$, Cohen's $d = -0.26$. Limburgian and Dutch participants thus only differ in trials featuring between-category variation, but not in the trials featuring within-category variation.

The same analysis was run on the RT data. The outcomes support the accuracy data. A significant TRIALTYPE \times LANGUAGE interaction ($F(1,32) = 5.66$, $p = .02$, $\eta_p^2 = .15$) showed that Dutch participants were significantly faster in within-category trials ($M = 490.77$, $SD = 193.86$) than in between-category trials ($M = 589.05$, $SD = 259.82$), $t(15) = 3.37$, $p = .004$, Cohen's $d = 0.84$), whereas Limburgian participants were equally fast in both trial types ($M_{between} = 568.25$, $SD_{between} = 277.77$; $M_{within} = 557.18$, $SD_{within} = 297.82$, $t(17) = .45$, $p > .05$, Cohen's $d = 0.11$). There was no main effect of LANGUAGE ($p > .05$).

4.4 Discussion

This study set out to answer two main questions. First, whether Limburgian and Dutch adults differ in their perception of the East-Limburgian lexical tones, and secondly, whether Dutch adults' sensitivity to the non-native Limburgian tones might be influenced by their native prosodic system. These two questions will be addressed below in two separate sections. Thirdly, we will discuss two findings with respect to our Limburgian listeners. Fourth and finally, we present suggestions for future research.

4.4.1 Do Limburgian and Dutch adults differ in their perception of Limburgian lexical tones?

Our results demonstrated that Limburgian and Dutch adults perform differently in discriminating Limburgian lexical tones. They differ in two respects. Limburgians were significantly more accurate ($M = .86$) than Dutch listeners ($M = .69$) in between-category variation trials, i.e., in trials in which

they had to discriminate between accents 1 and 2. Moreover, Dutch listeners performed better in within-category variation trials ($M = .81$) than between-category variation trials ($M = .69$), whereas Limburgians performed equally well in both trial types ($M_{within} = .84$, $M_{between} = .85$). We will elaborate on these results below.

Our findings on the between-category trials replicate previous research suggesting that tone language speakers are at an advantage in discriminating between lexical tone categories compared to naïve non-tone language speakers (Hallé et al., 2004; Huang & Johnson, 2010; Lee et al., 1996; Qin & Mok, 2011; Schaefer & Darcy, 2014; Tong et al., 2015; Wayland & Guion, 2003; Xu et al., 2006). This finding supports the assumption that cross-linguistic differences in the functions attributed to pitch patterns lead to differences in the pitch processing. Tone language speakers, who are acquainted with f_0 as a cue for lexical distinctiveness, pay more attention to pitch than speakers of non-tone languages, who have no or less experience with storing pitch as part of their lexical representations. Moreover, considering the fact that there is practically no cross-linguistic experimental research on Limburgian lexical tone perception to draw upon, we are the first to have shown that speakers of this language appear to have an advantage over speakers of a non-tone language when it comes to pitch processing. This advantage is probably due to the fact that pitch serves more functions in Limburgian than in Dutch and complements the finding by Fournier, Gussenhoven, Jensen, and Hagoort (2010) that East-Limburgian and Dutch listeners show different patterns of cerebral lateralization when perceiving word-level pitch contrasts.

Besides group differences on the between-category variation trials, performance also diverged on the within- versus between-category variation trials for the Dutch and Limburgian listeners. For the Limburgian listeners, two possible outcomes had been postulated. First, in accordance with prior studies, perception of within-category differences could have been more difficult than

perception of between-category differences because tone language listeners must usually ignore irrelevant tonal variations (i.e., within-category variations) to efficiently categorize f_0 contours into tones (Hallé et al., 2004; Hoffmann, Sadakata, Chen, Desain, & McQueen, 2014; Qin & Mok, 2011; Xu et al., 2006). Secondly, because of their exposure to a large amount of linguistically meaningful pitch information within syllables due to the tone-intonation interaction, Limburgians might exhibit a strong sensitivity to pitch in general, leading to equal performance in between- and within-category discrimination. Our results showed that Limburgian listeners performed equally well in both between- and within-category trials. We would like to propose two explanations for this outcome.

One explanation for finding equal performance by Limburgian participants in both trial types might be due to the properties of the input they are exposed to on a daily basis, causing them to exhibit a greater sensitivity to (linguistic) pitch in general. Limburgian listeners are confronted with a rather intriguing amount of linguistically meaningful pitch variation due to the interaction of tonal and intonational tones, which is typically absent in most tone languages. In fact, the variation that Limburgian listeners are exposed to in the realization of the lexical tones cannot be considered irrelevant phonetic variation, since it signals meaningful information at the post-lexical level. Moreover, being bidialectal, Limburgians have two rich prosodic systems at their disposal: Limburgian and Dutch. The Limburgians' experience with linguistic systems that make such an extensive use of pitch might positively affect their perception of pitch in general. The fact that the Limburgians are not better than the Dutch in within-category variation trials might be a consequence of the fact that they do have lexical pitch categories that still somehow guide their perception of word level pitch patterns, but obviously to a lesser extent than the tone language speakers in earlier studies.

A second explanation for the Limburgian listeners' equal performance in between- and within-category discrimination refers to the categorical nature of tones. Hallé et al. (2004) showed that their Mandarin-Chinese listeners

performed differently in between- and within-category trials¹⁶, but they also yielded generally high within-category discrimination performance. This raises questions about the categorical nature of tones. There is strong evidence for the categorical perception of consonants (Liberman, Harris, Hoffmann, & Griffith, 1957). Vowels on the other hand appear to be perceived much more continuously (Eimas 1963; Fry, Abramson, Eimas, & Liberman, 1962; Repp, 1984). Polka (1995) points out that the typical acoustic properties of vowels in natural speech (e.g., long duration) favor auditory, lower-level processing. Burnham and Mattock (2007) argue that tones are more akin to vowels than to consonants, and that they are perceived more continuously than categorically. They put forward two reasons for this assumption. First, tones are carried on vowels, and both tones and vowels extend over time. Secondly, f_0 production depends largely on quantitative changes in the rate of laryngeal vibration, rather than on qualitatively distinct speech gestures, as is the case for the production of consonants (place, manner of articulation, etc.). Hallé et al. point out that within-category discrimination is more likely if categories are loosely defined and overlap in terms of physical properties. By their nature, tones might invoke more auditory, psychophysical listening, enabling native listeners of a tone language to pick up even subtle within-category pitch differences.

As opposed to the Limburgians' performance, the performance of our Dutch listeners was *not* constant across within- ($M = .81$) and between-category trials ($M = .69$). This finding is not in line with the performance of the non-tone language speakers in previous studies (Hallé et al., 2004; Hoffmann et al., 2014; Xu et al., 2006). In those previous studies, non-native speakers showed equal performance in within- and between-category variation trials, which is explained in terms of L1 phonology. If no lexical tone categories are distinguished in the native phonological system, there are no categories that could influence perception. All tones are thus perceived as equally

¹⁶ Note that our within-category variation trials are comparable to trials featuring the stimulus pairs from the ends of the continua (pairs 1-3 and 6-8) in Hallé et al. (2004).

(un)important melodic variations, leading to similar performance in within- and between-category trials.

Yet, Dutch participants showed differential behavior in between- and within-category trials, which suggests that they may have been using *different* listening strategies in the two types of trial. For within-category trials, it is possible that the phonetic identity between the standard (A or B) and the target (X) stimulus has favored a psychophysical listening strategy. As we introduced token-variability in the between-category trials, however, it is unlikely that the Dutch listeners made use of pure acoustic matching in these trials. Whether their perceptual strategy in between-category trials was guided by native intonation categories, by acoustic salience, by their knowledge of different regional accents, or a combination of these factors, will be discussed in the next section addressing our second research question.

4.4.2 Is Dutch adults' perception of Limburgian tones influenced by their native prosodic system?

Dutch adults – albeit to a lesser extent than Limburgian adults – were able to discriminate the Limburgian tones, as did the Dutch infants in Ramachers et al. (2017; Chapter 3 of this thesis). It is possible that there is an influence of the L1, as the citation forms of the Limburgian accents 1 (H*L) and 2 (H*LH) are quite similar to Dutch intonation patterns signaling statements (H*L L%) and questions (H*L H%), respectively. Following PAM-S (e.g., So, 2012; So & Best, 2008, 2010, 2014), it is predicted that native speakers of non-tone languages assimilate foreign prosodic categories (in this case, Limburgian lexical tones) into their native prosodic categories (in this case, Dutch intonational categories). Those assimilation patterns in turn can predict discrimination performance for specific prosodic contrasts. Prior studies have indeed shown that naïve listeners' assimilations reflect the properties of the L1 prosodic system (e.g., So & Best, 2014, and references therein). If the Dutch indeed perceive and thus categorize the Limburgian tones as native intonation

categories, this could explain their above-chance performance in between-category variation trials.

To investigate this further, we compared Dutch listeners' performance on monosyllabic and disyllabic between-category variation trials. On monosyllables, we expected to find that the Dutch would be accurate in discriminating tones because in final, focal position the Limburgian tones can be perceived as Dutch post-lexical intonation categories. Recall that 12-month-old Dutch infants were also successful in discriminating Limburgian tones in monosyllables (Ramachers et al., 2017; Chapter 3 of this thesis). In final position, the tones could thus be recognized as something a Dutch listener is accustomed to hearing (Broselow et al., 1987). They could interpret the fall at the end of an utterance as a declarative, and the rise at the end of an utterance as interrogative intonation. On disyllables, however, the Dutch might be less sensitive to the tone contrast because in this position it does not correspond to an intonational contrast. If a contrast appears in an unusual position where it does not signal a meaningful linguistic contrast, listeners pay less attention to it (e.g., Braun & Johnson, 2011), they do not perceive it in a categorical manner (Hoffmann et al., 2014), and experience more difficulties mapping it onto native intonation categories (for English: Broselow et al., 1987; but see So and Best, 2014).

Strikingly, our Dutch listeners showed the opposite pattern. They performed significantly above chance in both mono- and disyllables, and they performed significantly better on disyllables ($M = .72$) than on monosyllables ($M = .66$). Discrimination of pre-final pitch movements by Dutch listeners has been found in previous work (Hoffman et al., 2014; but see Braun & Johnson, 2011). However, Hoffman et al.'s experiment, in contrast to ours, was a low memory demanding AX discrimination task that favors psychoacoustic processing. The performance of the Dutch listeners in their task is thus unlikely dependent on language-specific, phonological processing, and is therefore not directly comparable to our results.

Two reasons make it unlikely that our Dutch subjects perceived the tones in a disyllabic, non-final context as two native intonation patterns. First, it has been shown previously that Dutch listeners do not perceive pitch variation on pre-final syllables as a post-lexical contrast. Nonwords with a rise on the first syllable were equally likely to be perceived as a statement as nonwords with a fall on the first syllable (Braun & Johnson, 2011). Secondly, their discrimination performance ($M = .72$) was presumably too low to signal TC assimilation, since TC assimilation normally leads to excellent discrimination ($M > .90$). We therefore would like to suggest that the higher discrimination accuracy in disyllables could be attributed to acoustic salience. In disyllables, the difference between accents 1 and 2 may have been more apparent for two reasons. By their nature, disyllables enable listeners to compare the first and the second syllable of a word, possibly enhancing listeners' sensitivity to the pitch pattern on the first syllable relative to the second syllable. Note that word stress in Dutch is also a relative property (see also Cooper et al., 2002). Moreover, accents 1 and 2 surface as a contour and level tone in disyllables, whereas in monosyllables they both surface as contour tones.

Evidence that acoustic salience plays a role in discrimination of lexical tones comes, for example, from a study by Wayland and Guion (2003). They showed that native and non-native listeners of Thai obtained higher discrimination scores on Thai tones for closed syllables than for open syllables. They attributed this pattern to the fact that the middle and low tones in open syllables differed only in their f_0 onset and were thus less salient than their counterparts in closed syllables, where they differed both in f_0 onset and offset.

Recall that we conducted a series of independent t -tests to investigate which acoustic correlates contributed to the difference between accents 1 and 2 in mono- and disyllables (see Table 1 in the Appendix for an overview of our acoustic measurements and their accompanying statistics). With respect to pitch characteristics, the results revealed that accents 1 and 2 in monosyllables differed significantly with respect to minimum pitch within the tone bearing portion (TBP; $t(14) = -7.43$, $p < .001$) and with respect to the f_0 offset value

measured at the end of the TBP ($t(14) = -14.75, p < .001$). No other significant differences were found (all $ps > .05$).

Accents 1 and 2 in disyllables could be differentiated based on minimum pitch in the TBP ($t(14) = -8.02, p < .001$), pitch range in the TBP ($t(7) = 6.02, p < .001$) and f_0 offset ($t(8) = -8.50, p < .001$). Especially the significant difference in pitch range ($M_{\text{accent1}} = 241.7, SD = 94.5$; $M_{\text{accent2}} = 39.8, SD = 8.9$) characterizes the level-contour contrast in Limburgian and probably contributes greatly to its acoustic salience. Wang (1976) states that, from a psychoacoustic point of view, it is easier to discriminate a level tone from a contour tone than to discriminate two contour tones (but cf. Burnham & Francis, 1997; Burnham, Kirkwood, Luksaneeyanawin, & Pansottee, 1992, cited in Wayland & Guion, 2003). F_0 offset is put forward as an important cue to tone discrimination for non-tone language listeners by Gandour and Harshman (1978). Regarding the significant difference in the f_0 offset values between accent 1 and accent 2 for both monosyllabic and disyllabic stimuli (see also Figure 2, panels a-d), the Dutch listeners in our study may surely have benefited from this cue.

Not only pitch characteristics such as pitch height and pitch range, but also other cues such as duration, amplitude, and voice quality may have contributed to the differences between Limburgian accent 1 and accent 2 in non-final position (e.g., Köhnlein, 2016, and references therein). We observed two instances of this at the item level. First, post-hoc phonetic inspection of our disyllabic pseudo-word *keeve* indicated that the voice quality of the TBP (that is, the vowel [e:]) differed over time to some extent in accent 1 tokens (falling pitch), but not in accent 2 tokens (level pitch). In accent 1 tokens, the energy associated with the vowel weakens in the second half of the vowel. This energy drop appears to go hand in hand with the pitch fall. In accent 2 tokens, however, energy is constant throughout the TBP. English listeners in Tong et al. (2015) also appeared to use overall spectral energy (i.e., spectral tilt, more energy at higher frequencies) as a perceptual cue to perceive non-native

Cantonese tones. Tong and colleagues argue that speakers of English use the correlates associated with native suprasegmental categories (in this case lexical stress) during perception of other, non-native suprasegmental categories. Indeed, spectral tilt is also a correlate of stress in Dutch (Sluijter & van Heuven, 1996). It is thus possible that Dutch listeners have drawn upon their knowledge of native cues to word stress during their perception of Limburgian tones.

Another observation within our disyllabic items concerned the disyllabic stimulus item *perger*. The coda consonant in the first syllable appears as a vocalic 'r' [ɐ] in accent 2 tokens, whereas it appears as a consonant [ʀ] in accent 1 tokens. Based on native speaker intuition, we have the impression that this is also a secondary acoustic correlate of the difference between accent 2 as a level tone and accent 1 as a contour tone. Note, however, that this study was not designed for the purpose of acoustic analysis and that the number of tokens used in this study is too small to draw general conclusions. These observations should therefore be interpreted with caution. An extensive analysis of all possible acoustic correlates of the Limburgian tones (see, for example, Tong et al., 2015) would be worthwhile to investigate in the future.

Following these observations, it is possible that additional cues have influenced the perceptual strategy adopted by our Dutch listeners in between-category variation trials and that they have perceived the tones in a psychoacoustic fashion, similar to what has been proposed for the French listeners in Hallé et al. (2004).

Alternatively, some higher-order categorization process might still have taken place. Participants might not have perceived the difference between stimuli as different native intonation patterns (e.g., 'this sounds like a question and this like a statement') but rather as differing in their degree of regional-accentedness (e.g. 'this sounds Limburgian and this sounds Dutch'). The accent 2 tokens might have been perceived as 'more Limburgian accented', whereas the accent 1 tokens sounded 'less Limburgian accented' and more Dutch-like. Accents are typically defined by multiple acoustic correlates at both the segmental and suprasegmental level (e.g., Kitamura, Panneton, & Best, 2013;

Mulak, Best, Tyler, Kitamura, & Irwin, 2013). Following up on the results of the restricted acoustic analysis of our stimuli, the Dutch listeners might have picked up on several acoustic correlates (among others spectral tilt). These multiple cues might have given them the impression that certain stimuli sound more foreign than others. Gili-Fivela (2012) posits that foreign sounding stimuli can trigger a less language-specific and more psychoacoustic listening strategy. Thus, perhaps our Limburgian stimuli sounded too foreign to be mapped onto Dutch intonation contours in the first place and instead were categorized (and discriminated) in terms of degree of foreign accent. This explanation could also hold for the above-chance performance in monosyllabic trials. In these trials, accent 2 (realized as a fall-rise) could also be considered the more foreign pitch pattern. It has been suggested that accent 2 is one of the most outstanding features of Limburgian dialects because it lends them their typical 'sing-song' character (e.g., van Bezooijen & Gooskens, 1999).

To explore the possibility that Dutch listeners perceive the difference between accent 1 and accent 2 as a difference in accentedness, a rating task in which listeners rate stimuli on their degree of accentedness could be conducted. Despite the fact that all of our Dutch participants indicated no regular exposure to people speaking a Limburgian *dialect* in their presence, it is highly likely that they still have exposure to people from Limburg speaking Limburgian *accented* Dutch. The linguistic knowledge of most Dutch language users covers both standard-like and more regional varieties, for example as a consequence of exposure through the media (Impe, Geeraerts, & Speelman, 2008).

In a final attempt to shed some light on the possibly different degree of assimilability of the Limburgian tones to native Dutch prosodic categories on the basis of our discrimination data, we ran a post-hoc analysis to investigate the presence of the *native similarity effect* (Best et al., 2001). The native similarity effect involves "an asymmetry favoring discrimination when the X is the more rather than the less native-like member of the non-native contrast"

(Best et al., 2001, p. 790). Perception of tokens that are more native-like is more stable, leading to better discrimination, whereas perception of less native-like stimuli is less stable. If we would find that discrimination performance is significantly worse in trials where the X stimulus carries accent 2 compared to trials where X carries accent 1, this would signal that accent 2 is perceived as less native-like than accent 1.

To investigate whether this native similarity effect is present, we looked at the effect of an additional within-subjects variable, the tone of stimulus X (ToneX: accent 1 vs. accent 2), on the discrimination accuracy of our Dutch participants ($N = 18$). We also included Condition (monosyllabic vs. disyllabic) in the analysis, because the different surface forms of accent 2 (contour vs. level) may have consequences for the degree of assimilation to a native Dutch intonation pattern. This analysis only yielded a main effect of Condition, $F(1,17) = 5.43$, $p < .03$, $\eta_p^2 = .24$ (i.e., the same effect as reported in our prior main Language x Condition x Order analysis) and no effects nor interactions with ToneX (all $ps > .05$). Thus, the present data do not provide evidence for the possibility that there are differences in native-likeness between accents 1 and 2 or between the different surface realizations of accent 2.

4.4.3 Remarkable findings pertaining to the Limburgian listeners

Although Limburgian participants performed significantly better than Dutch subjects in discriminating between the Limburgian lexical tones, interestingly, their performance was not at ceiling in between-category variation trials ($M = 85\%$). By itself, but also compared to most other studies on native tone discrimination, this is a rather unexpected finding. Mandarin Chinese listeners in Xu et al. (2006) performed similar to our Limburgian listeners, reaching an accuracy of 86,5% in an AX same-different task with Mandarin tones. However, in Hallé et al.'s (2004) non-categorical AXB two-step discrimination task, Mandarin listeners scored above 90% in trials featuring tone pairs at the

category boundaries (i.e., between-category variation trials). In a categorical odd-one-out-paradigm, in which listeners heard three stimuli and had to indicate which one of them had a different tone, native Thai listeners in Wayland and Guion (2003) discriminated native Thai tones with 91% accuracy. Cantonese-English bilinguals in Tong et al. (2015) discriminated Cantonese tones with an accuracy of 95% in a non-categorical AXB discrimination task.

The first thing to note is that all these studies used slightly different methodologies which could invoke different listening strategies (i.e., more or less acoustic-phonetic or language-specific). Moreover, the aforementioned studies were on typically studied tone languages, whereas Limburgian is considered a restricted tone language because of the relatively low functional load of lexical tone, and Limburgian tones yield a considerable amount of surface variation. This means that it is problematic to solely interpret our data in the light of previous work. Therefore, we believe it is necessary to discuss the below-ceiling performance of our Limburgian listeners as a finding in its own right.

First, the degree of sensitivity to lexical tones might depend on the extent to which a native speaker actively uses the tone language. In our language background questionnaire, all Limburgian participants had indicated that they lived in Limburg at the time of test and that the East-Limburgian dialect was (one of) their L1(s). All our participants were bidialectal, though, as they also spoke Dutch. Possibly, some of them did not use the Limburgian dialect on a regular basis and perhaps used Dutch more frequently. The discrimination scores of those participants might have been lower than for participants who use the Limburgian dialect regularly. Future research should take this issue of language use (i.e., language dominance) into account.

Secondly, the Limburgians' relatively low discrimination scores might be a task effect (e.g., Goss & Tamaoka, 2015). Unlike other AXB discrimination experiments (e.g., Hallé et al., 2004; So & Best, 2014; Tong et al., 2015; Wayland & Guion, 2003), we added a substantial number of filler trials containing vowel

and consonant differences to direct participants' attention away from the purpose of the experiment (i.e., lexical tone perception). Moreover, our experiment featured both between-category differences as well as within-category differences. This variety of phonetic differences presented to our listeners might have resulted in doubt or confusion as to what participants should pay attention to. We did not tell our participants beforehand that they should pay attention to pitch differences, whereas a number of prior studies did explicitly direct their participants' attention to the contrast in question (e.g., Harnsberger, 2001; Polka, 1995; So & Best, 2014; Wayland & Guion, 2003). All the above-mentioned decisions were taken to let our task resemble natural listening conditions as much as possible, but they could have complicated the task as well.

Thirdly, the functional load of a lexical pitch contrast may influence the way it is processed (e.g., Goss, 2015; Wu, Tu, & Wang, 2012). If we assume that the lexical tone contrast in Limburgian has a relatively low functional load compared to Mandarin or Thai, this could perhaps explain the Limburgians' relatively low discrimination performance. More studies investigating discrimination and categorization of lexical tones in different tone languages are desirable to find out whether functional load has an impact on category strength and, as a result, on discrimination and/or categorization.

Another unexpected finding on the between-category trials is that our Limburgian participants performed significantly better on disyllables than on monosyllables *only* in trials where an accent 2 token was presented first. This implies that Limburgians find it easier to discriminate a contour after a level tone than to discriminate a level after a contour tone. If the cause of this is phonological in nature, it is possible that the effect emerged from accent 2 being the (more) lexically specified category from which any deviation would be noticed. However, if this was the case, we would have expected the same order effect for our monosyllabic trials. If the cause was phonetically motivated, in that it is perceptually easier to discriminate a contour after a level than a level after a contour, we would have expected the same effect in our Dutch

listeners as well. Our results are thus inconclusive regarding the cause of this perceptual asymmetry, and more research is required to see whether it is a genuine effect.

4.4.4 Directions for future research

In this study we have shown that adult speakers of East-Limburgian perform better in a lexical tone discrimination task than speakers of Dutch, probably due to the lexical distinctiveness of tone in Limburgian. Moreover, speakers of Limburgian are highly sensitive to both between- and within-category pitch variation, probably as a result of their accumulated experience with two rich prosodic systems. Speakers of Dutch also showed sensitivity to lexical tones, albeit to a lesser extent than the Limburgians. Whether this sensitivity to a non-native prosodic contrast can be explained by assimilation to native intonation categories cannot be answered on the basis of our results. Dutch listener's perception could also be influenced by particular stimulus item characteristics (e.g., differences in spectral tilt). We would like to propose some follow-up studies that could improve the understanding of our partly unexpected results.

To clarify whether and how the Dutch assimilate the different realizations of Limburgian tones to native prosodic categories, it is necessary to carry out a categorization task followed by a goodness-rating task (e.g., Best et al., 2001; Højen & Flege, 2006; Polka, 1995; So & Best, 2014; Tyler et al., 2014). In such a two-step task, participants would be asked to categorize Limburgian tones into a set of predetermined native categories, for example 'statement', 'question', 'uncertainty', 'don't know'. After choosing an answer category, participants rate the goodness of fit by judging how well the target tone matches the native category they have just chosen. These 'goodness of fit' ratings make it possible to distinguish between single-category (SC) and category-goodness (CG) assimilations. Suppose both Limburgian tones would be categorized as 'statement', then the goodness of fit rating could make clear whether both tones are equally good or bad instances of the statement category

(an instance of SC assimilation), or that one of the tones fits the native category better than the other (CG assimilation).

A further possibility for future research would be to employ a discrimination task with the target stimuli embedded in a sentence context instead of presented in isolation. This is more naturalistic and might encourage listeners more to invoke their native prosodic categories (Gili-Fivela, 2012; So & Best, 2014). Moreover, anecdotal evidence suggests that Dutch listeners have a hard time perceiving the difference between the Limburgian tones in running speech. Another way to increase the chance of linguistic processing to be applied is to increase the task demand even further by adding stimuli from multiple speakers (e.g., Dupoux et al., 2001).

To broaden our understanding of how lexical tone processing is influenced by the functional load of tone and the extent to which tones have different surface realizations, native tone discrimination in restricted tone languages such as Limburgian and Norwegian, Swedish and Japanese should be compared to native discrimination in denser tone languages such as Mandarin and Cantonese.

4.5 Conclusion

This study is the first to have shown that adult speakers of East-Limburgian dialects have an advantage over speakers of a non-tone language (in this case: Dutch) when it comes to word-level pitch processing. From this we can conclude that speakers of Limburgian attend more closely to word level pitch than speakers of Dutch, probably as a result of the lexical distinctiveness and possibly also the more local occurrence of pitch variation in Limburgian. Moreover, Limburgian listeners performed equally well on discriminating between- and within-category pitch variation. Their high sensitivity to various pitch patterns might be a result of their ample exposure to meaningful pitch variations due to the intricate tone-intonation interaction in Limburgian and because of their command of two languages with rich prosodic systems. The fact that Dutch listeners still showed significant sensitivity to different kinds of

pitch variations could be due to their experience with pitch as a cue to post-lexical prosody (i.e., intonation) and stress.

References

- Best, C. (1995). A direct realist view on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and Methodological issues in cross-language speech research* (pp. 167-200). York: Timonium.
- Best, C., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, *109*(2), 775-794.
- Best, C., McRoberts, G., & Sithole, N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345-360.
- van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties. The contribution of different linguistic levels. *Journal of Language and Social Psychology*, *18*(1), 31-48.
- Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.22. Retrieved from <http://www.praat.org/>.
- Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Language Research*, *30*, 323-350.
- Braun, B., & Johnson, E. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, *39*, 585-594.

- Broselow, E., Hurtig, R., & Ringen, C. (1987). The perception of second language prosody. In G. Loup & S. Weinberger (Eds.), *Interlanguage phonology. The acquisition of a second language sound system* (pp. 350-364). New York: Newbury House.
- Burnham, D., & Brooker, R. (2002). Absolute pitch and lexical tones: tone perception by non-musician, musician, and absolute pitch non-tonal language speakers. *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2002)* (pp. 257-260).
- Burnham, D., Brooker, R., & Reid, A. (2015). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music, 43*(6), 881-897. DOI: 10.1177/0305735614546359.
- Burnham, D., & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. In A. S. Abramson (Ed.), *Southeast Asian linguistic studies in honour of Vichin Panupong* (pp. 29-46). Bangkok: Chulalongkorn University Press.
- Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattasanone, N., Schwarz, I.-C., & Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics, 36*(6), 1459-1491. <http://dx.doi.org/10.1017/S0142716414000496>.
- Burnham, D., Kirkwood, K., Luksaneeyanawin, S., & Pansottee, S. (1992). Perception of Central Thai tones and segments by Thai and Australian adults. In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics* (pp. 546-560). Bangkok: Chulalongkorn University Printing House.
- Burnham, D., & Mattock, K. (2007). The perception of tones and phones. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech*

learning: In honor of James Emil Flege (pp. 259-280). Amsterdam: John Benjamins.

Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, 25(5), 426-439.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, 45(3), 207-228.

Cutler, A., & van Donselaar, W. (2001). *Voor naam* is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, 44(2), 171-195.

Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford University Press.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *Journal of the Acoustical Society of America*, 110(3), 1606-1618.

Eimas, P. (1963). The relation between identification and discrimination along speech and non-speech continua. *Language and Speech*, 6(4), 206-217.

Fournier, R., Gussenhoven, C., Jensen, O., & Hagoort, P. (2010). Lateralization of tonal and intonational pitch processing: An MEG study. *Brain Research*, 1328, 79-88.

Fry, D., Abramson, A., Eimas, P., & Liberman, A. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, 5(4), 171-189.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 49-175.

- Gandour, J., & Harshman, R. (1978). Cross-language differences in tone perception: a multidimensional scaling investigation. *Language and Speech*, 21, 1-33.
- Gerrits, E., & Schouten, M. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, 66(3), 363-376.
- Gili-Fivela, B. (2012). Testing the perception of L2 intonation. In M. G. Busa & A. Stella (Eds.), *Methodological perspectives on Second Language Prosody. Papers from ML2P 2012* (pp. 17-30). Padova: CLEUP Sc.
- Goss, S. (2015). *The effects of internal and experience-based factors on the perception of lexical pitch accent by native and nonnative Japanese listeners*. Doctoral dissertation, Ohio State University.
- Goss, S., & Tamaoka, K. (2015). Predicting lexical accent perception in native Japanese speakers: An investigation of acoustic pitch sensitivity and working memory. *Japanese Psychological Research*, 57(2), 143-154.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Hallé, P., & Best, C. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America*, 121(5), 2899-2914.
- Hallé, P., Chang, Y.-C., & Best, C. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- Harnsberger, J. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, 110(1), 489-503.

156 Setting the Tone

- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581-616.
- van Hessen, A., & Schouten, M. (1992). Modeling phoneme perception. II: A model of stop consonant discrimination. *Journal of the Acoustical Society of America*, 92(4), 1856-1868.
- Hoffmann, C., Sadakata, M., Chen, A., Desain, P., & McQueen, J. (2014). Within-Category Variance and Lexical Tone Discrimination in Native and Non-Native Speakers. *Proceedings TAL 2014*, 45- 49.
- Højen, A., & Flege, J. (2006). Early learners' discrimination of second-language vowels. *Journal of the Acoustical Society of America*, 119(5), 3072-3084.
- Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and non-native listeners. *Phonetica*, 67, 243-267.
- Hyman, L.M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, 31, 213-238.
- Impe, L., Geeraerts, D., & Speelman, D. (2008). Mutual intelligibility of standard and regional Dutch language varieties. *International Journal of Humanities and Arts Computing*, 2(1-2), 101-117.
- Kitamura, C., Panneton, R., & Best, C.T. (2013). The development of language constancy: Attention to native versus non-native accents. *Child Development*, 84(5), 1686-1700.
- Köhnlein, B. (2016). Contrastive foot structure in Franconian tone-accent dialects. *Phonology*, 33, 87-123.

- Koster, M., & Cutler, A. (1997). Segmental and suprasegmental contributions to spoken-word-recognition in Dutch. *Proceedings of Eurospeech 97* (pp. 2167-2170). Rhodes.
- Lee, Y-S., Vakoch, D.A., & Wurm, L.H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*(5), 527-542.
- Liberman, A., Harris, K., Hoffmann, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 61*, 379-388.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition, 133*, 385-394.
- Mattock, K. & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*, 241-265.
- 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.
- McGuire, G. (2010). A brief primer on experimental designs for speech perception research. *Laboratory report, 77*.
- Mulak, K.E., Best, C.T., Tyler, M.D., Kitamura, C., & Irwin, J.R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development, 84*(6), 2064-2978.
- Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America, 97*(2), 1286-1296.

- Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*(1), 577-592.
- Qin, Z., & Mok, P. (2011). Discrimination of Cantonese tones by Mandarin, English and French speakers. *Proceedings of the psycholinguistic representation of tone* (pp. 50-53). Hong Kong: Causal Productions.
- Ramachers, S., Brouwer, S., & Fikkert F. (2017). 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.
- Repp, B. H. (1984). Categorical perception: Issues, methods, and findings. In N.L. Lass (Ed.) *Speech and Languages: Advances in Basic Research and Practice*, Vol. 10. New York: Academic.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, *22*(11), 2503-2513.
- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, *5*(4), 489-522.
- Schneider, K., Dogil, G., & Möbius, B. (2011). Reaction time and decision difficulty in the perception of intonation. *Proceedings of Interspeech*, 2221-2224.
- Singh, S., & Fu, C. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, *87*(3), 834–854. doi:10.1111/cdev.12512.
- Sluijter, A., & Heuven, V. van. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, *100*(4), 2471-2485.

- So, C. (2012). Cross-language categorization of monosyllabic foreign tones: Effects of phonological and phonetic properties of native languages. In T. Stolz, N. Nau & C. Stroh (Eds.), *Monosyllables: From phonology to typology* (pp. 55-69). Berlin: Akademie Verlag.
- So, C., & Best, C. (2008). Do English speakers assimilate Mandarin tones to English prosodic categories? In *Proceedings of the 9th annual conference of the International Speech Communication Association (Interspeech 2008), incorporating the 12th Australasian International Conference on Speech Science and Technology (SST 2008)* (p. 1120). Baixas, France: International Speech Communication Association.
- So, C., & Best, C. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech, 53*, 273–293.
- So, C., & Best, C. (2011). Categorising Mandarin tones into listeners' native prosodic categories: The role of phonetic properties. *Poznan Studies in Contemporary Linguistics, 47*, 133-145.
- So, C., & Best, C. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition, 36*, 195-221.
- Strange, W., & Shafer, V. (2008). Speech perception in second language learners: Re-education of selective perception. In J. Hansen Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-191). Amsterdam: John Benjamins.
- Tong, X., Lee, S.M.K., Lee, M.M.L., & Burnham, D. (2015). A Tale of Two Features: Perception of Cantonese Lexical Tone and English Lexical Stress in Cantonese-English Bilinguals. *PLoS ONE, 10*(11): e0142896. doi:10.1371/journal.pone.0142896.

- Tsao, F.-M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in Psychology*, doi:10.3389/fpsyg.2017.00558.
- Turk, A., Nakai, S., & Sugahare, M. (2006). Acoustic Segment Durations in Prosodic Research: A Practical Guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter & J. Schließer (Eds.), *Methods in Empirical Prosody Research* (pp. 1-27). Berlin: De Gruyter.
- Tyler, M., Best, C., Faber, A., & Levitt, A. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4-21.
- Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*, 4(1), 1-22.
- Wang, W. (1976). Language change. *Annals of the New York Academy of Sciences*, 280, 61-72.
- Wayland, R., & Guion, S. (2003). Perceptual discrimination of Thai tones by naive and experienced learners of Thai. *Applied Psycholinguistics*, 24, 113-129.
- Werker, J., & Tees, R. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75(6), 1866-1878.
- Wu, X., Tu, J.-Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623-641.
- Xu, Y., Gandour, J., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch contrasts. *Journal of the Acoustical Society of America*, 120(2), 1063-1074.

Yeung, H.H., Chen, K.H., & Werker, J. (2013.). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Appendix

Table 1. Acoustic measurements (Mean (SD)) and *p*-values of the independent *t*-tests on the test stimuli.

Condition	Monosyll.	Monosyll.	Disyll.	Disyll.	<i>p</i>
Tone	Accent 1 (n = 8)	Accent 2 (n = 8)	Accent 1 (n = 8)	Accent 2 (n = 8)	<i>p</i>
Min F ₀ TBP* (Hz)	153.6 (7.9)	194.6 (13.5)	159.8 (28.9)	312.2 (45.3)	<.001
Max F ₀ TBP (Hz)	414.0 (81.5)	404.8 (90.1)	401.4 (78.1)	352.1 (47.5)	>.1
Mean F ₀ TBP (Hz)	262.2 (36.8)	280.3 (26.6)	302.7 (48.7)	340.5 (44.1)	>.1
F ₀ range TBP (Hz)	260.4 (81.9)	210.2 (99.1)	241.7 (94.5)	39.8 (8.9)	<.001
F ₀ offset TBP (Hz)	150.55 (6.18)	276.05 (23.3)	168.14 (13.2)	312.86 (46.3)	<.001
Duration TBP (s)	.37 (.06)	.38 (.04)	.27 (.01)	.26 (.02)	>.1
Duration token (s)	.57 (.05)	.56 (.06)	.66 (.04)	.62 (.06)	>.1

*TBP stands for Tone Bearing Portion. Note that the TBP for *moon* stimuli consisted of the rhyme (i.e., [o:n]), for *taaf* it consisted of the nucleus (i.e., [a:]), for *keeve* it consisted of the nucleus (i.e., [e:]) and for *perger* it consisted of the rhyme (i.e., [ɛR]).

Table 2. Overview of participant exclusion per dependent variable per analysis.

	questionnaire	2 SD accuracy LxCxD*	2 SD accuracy LxT**	2 SD RT*** LxCxD	2 SD RT LxT
total n excluded	4	3	2	6	5
total n included	39	36	37	33	34
Dutch					
n included	19	18	18	15	16
mean age (range)	45 (19-66)	44 (19-60)	44 (19-60)	46 (19-66)	44 (19-66)
n male	8	8	8	9	9
Limburgian					
n included	20	18	19	18	18
mean age (range)	49 (23-67)	50 (23-67)	50 (23-67)	48 (23-65)	48 (23-65)
n male	11	10	9	8	8

*LxCxD stands for the analysis on between-category variation trials only.

**LxT stands for the analysis comparing between-category and within-category variation trials.

***RT stands for the reaction time analysis.

5. Does native prosody affect pitch processing during word learning? Lexical encoding of tones by Limburgian and Dutch toddlers and adults

Based on: Ramachers, S., Brouwer, S., and Fikkert, P. (2017). How native prosody affects pitch processing during word learning in Limburgian and Dutch toddlers and adults. *Frontiers in Psychology*, <https://doi.org/10.3389/fpsyg.2017.01652>.

Abstract

In this study, Limburgian and Dutch 2,5- to 4-year-old children as well as adults took part in a word learning experiment. Following the procedure employed by Quam and Swingley (2010) and Singh, Hui, Chan, and Golinkoff (2014), 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. Since Limburgian is considered a restricted tone language, we expected that the pitch change would hinder word recognition in Limburgian, but not in non-tonal Dutch listeners. Contrary to our expectations, both Limburgian and Dutch children appeared to be sensitive to pitch changes in newly learned words, indicated by a significant decrease in target fixation in MP trials compared to CP trials. Limburgian and Dutch adults showed very strong word recognition in both trial types. The results are discussed against the background of the influence of the native prosodic system.

5.1 Introduction

Acquiring the sound structure of a language entails finding out which phonetic contrasts are meaningful in the native language (L1) and storing them as part of a word's lexical representation. Children need to learn to assign appropriate interpretations to many different sorts of phonetic variation, and separate variation that is lexically meaningful (i.e., phonemic variation) from variation that is not (e.g., speaker variation). Many studies have looked into the developmental perception of speech sound contrasts in the first year of life and into the way they are processed during word learning and recognition at later ages (e.g., Jusczyk, 1997; Kuhl, 2004; Stager & Werker, 1997; Swingley & Aslin, 2000; White & Morgan, 2008). This research has mainly focused on segmental contrasts, whereas approximately 60-70% of the world's languages employ pitch differences to distinguish words in addition to vocalic and consonantal contrasts (Yip, 2002). The aim of the present study is to add to the field of lexical tone acquisition by investigating the role of pitch contrasts during novel word learning. This is examined in child and adult speakers of Limburgian dialects of Dutch. Limburgian¹⁷ is a restricted tone language yielding an intriguing interaction between lexical and intonational tones. Limburgian participants' performance in a word learning experiment is compared to a control group of monolingual child and adult speakers of Dutch.

Pitch variation is meaningful in all languages of the world (Gussenhoven, 2004; Singh & Fu, 2016; Yip, 2002). Precisely how languages exploit variations in fundamental frequency (f_0) varies from one language to the next. In tone languages such as Mandarin Chinese, pitch patterns are used to distinguish words, similar to what phonemes do at the segmental level. Some tone languages make very extensive use of lexical pitch differences. In the densest case, in a tone language like Mandarin Chinese almost every syllable has to be pronounced with one of four tones (Duanmu, 2000). Other tone

¹⁷ Note that *Limburgian* is an umbrella term for many different dialects. See Chapter 2 of this thesis for more details.

languages are more restricted in their use of lexically contrastive pitch. These languages, for example Tokyo Japanese, have been referred to as either *PITCH-ACCENT LANGUAGES* or *RESTRICTED TONE LANGUAGES* (Hyman, 2009; Voorhoeve, 1973). There is some debate over whether there is a clear-cut distinction between tone languages on the one hand and restricted tone languages on the other. What they have in common is that pitch, be it to a greater or lesser extent, is necessary for determining the meaning of a word. Following Hyman's (2001, 2009) definition, we take the term 'tone language' to refer to languages that use pitch to distinguish between words, unless when referring to studies that used a different term.

Importantly, in non-tone languages like Dutch and English, pitch is not used to distinguish between words – except when it functions to enable the discrimination of a few very rare minimal pairs that differ in word stress (e.g., Dutch *VOORkomen* 'appear' vs. *voorKOMEN* 'prevent'), in which case pitch is only one of several correlated cues to stress. The fact that pitch is not used to signal lexical distinctiveness in non-tone languages might prevent speakers of these languages from distinguishing monosyllables that differ in pitch only (Schaefer & Darcy, 2014). Moreover, they might also refrain from encoding pitch information when building novel lexical representations due to the lack of experience with storing long-term memory representations of pitch (Braun, Galts, & Kabak, 2014).

Despite the above-mentioned functional differences, non-tone language listeners often show persistent sensitivity to non-native lexical tones throughout the lifespan (e.g., Hallé, Chang, & Best, 2004; Liu & Kager, 2014; Ramachers, Brouwer, & Fikkert, 2017, Chapter 3 of this thesis; Chapter 4 of this thesis; So & Best, 2010, 2014). Persistent sensitivity is mostly shown in purely perceptual tasks without lexical involvement (i.e., discrimination tasks; e.g., Broselow, Hurtig, & Ringen, 1987; Liu & Kager, 2014; Ramachers et al., 2017, Chapter 3 of this thesis; Chapter 4 of this thesis; Schaefer & Darcy, 2014; So & Best, 2008, 2010, 2014). Several factors have been put forward recently to

account for these findings, the most important one being the role of prosody in the L1.

The PERCEPTUAL ASSIMILATION MODEL FOR SUPRASEGMENTALS (PAM-S; So & Best, 2014) states that non-native pitch contrasts tend to be perceived according to their degree of similarity to native pitch patterns. The perceived distance between the non-native contrast and the closest native pattern(s) (if any) leads to differences in non-native contrast discriminability. Indeed, a number of studies on the perception of non-native pitch patterns have shown that prosodic experience from listeners' L1 guides their perception of non-native pitch patterns (e.g., Broselow et al., 1987; So & Best, 2008, 2010, 2014). For example, English listeners presumably discriminate Mandarin tone 4 (falling) due to assimilation to their statement intonation category (e.g., Broselow et al., 1987; So & Best, 2008), and Dutch listeners in Braun and Johnson (2011) probably perceived utterance-final Mandarin tone 2 (rising) as Dutch question intonation. Following these observations, the question thus no longer is whether non-native listeners discriminate lexical tones, but whether they interpret them as lexically relevant.

When acquiring a lexicon, tone-language learners need to learn to ascribe lexical relevance to pitch changes and encode tone lexically. Conversely, non-tone-language learners have to learn to disregard pitch changes that occur within words, despite the fact that they might still discriminate these pitch changes at lower levels of processing (e.g., in a purely perceptual task).

5.1.1 Integration of pitch into lexical representations

Recent work suggests that child and adult speakers of tone languages behave differently from non-tone language speakers in exploiting contrastive pitch contours when learning words. Tone language speakers attend to pitch information and exploit it during lexical access, whereas non-tone languages speakers do not, or at least to a lesser extent (e.g., Braun et al., 2014; Hay, Graf Estes, Wang, & Saffran 2015; Quam & Swingley, 2010; Singh et al., 2014). These

previous word learning studies primarily discussed the lexical integration of pitch by non-tone language listeners. Few of them looked at the interpretation of (non-)native pitch by listeners of tone languages, and if so, they focused on typically studied tone languages like Mandarin Chinese. However, within the family of tone languages, large differences exist.

First, tone languages differ with respect to the functional load of lexical tone. The functional load of tone depends on the tonal inventory (i.e., the number of tones, and, related to that, their information value), the distributional restrictions of tones (i.e., can they appear on any syllable?), the importance of tones for lexical disambiguation (i.e., how many minimal pairs are there in the language?), and the extent to which f_0 is the only cue to the tonal distinction (i.e., do duration or voice quality play a role?) (e.g., Kristoffersen, 2000; Pierrehumbert & Beckman, 1988; Tong, Francis, & Gandour, 2008; Wang, Bene, Jongman, & Sereno, 2004; Wu, Tu, & Wang, 2012). The smaller the inventory, the larger the amount of distributional restrictions, and the smaller the number of tonal minimal pairs, the more restricted a tone system is considered to be (Voorhoeve, 1973). The functional load of word-level pitch patterns in the L1 has been assumed to influence sensitivity to word-level pitch in speakers of these languages (e.g., Goss, 2015; Schaefer & Darcy, 2014; Wang et al., 2004; Wu et al., 2012).

There are some studies on the importance of pitch-accent for lexical access in Japanese, using for example lexical decision tasks. These studies indicate that Japanese speakers are sensitive to pitch as part of a word's phonological form (e.g., Cutler & Otake, 1999; Goss & Tamaoka, 2015). Yet, lexical pitch might not have equal priority as a cue to word recognition as vowels and consonants. This could be due to the relatively low number of accentual minimal pairs in Japanese and to the fact that, at least in a sentence context, there are other linguistic cues available for disambiguation (e.g., Tamaoka, Saito, Kiyama, Timmer, & Verdonschot, 2014).

A second difference within the family of tone languages lies in the complexity of their intonation systems. Typically, tone languages do not have complex intonation systems (e.g., Gussenhoven & van der Vliet, 1999) and, as a consequence, the pronunciation of a word with a certain lexical tone is rather stable across different contexts. In Standard Chinese, for example, different intonations only cause changes in pitch register, not in pitch contours (Wu, 2000). However, some more restricted tone systems like Norwegian, Swedish, and Limburgian, do show complex intonation systems. In these languages, intonation tones interact with lexical tones, causing variation in surface realizations (i.e., contours) of a lexical tone (e.g., Riad, 2013; Gussenhoven, 2000a). If a learner is confronted with many surface realizations, he may have a harder time abstracting away from all this variation to recover the underlying tone system. Indeed, it has been shown that surface variability in tone contours can delay the acquisition of lexical tone assignment (e.g., Demuth, 1995; Ota, 2003).

In the present study, we investigated lexical encoding of tone in Limburgian (for details on the Limburgian tonal system, see Chapter 2 of this thesis). By studying a language with a low functional load for a binary tone contrast embedded in a complex intonation system, this study widens our understanding of the influence of the functional load of tone and tonal surface variability on the acquisition and processing of a lexical tone system. By comparing speakers of Limburgian to a control group of non-tonal Dutch peers, we also address the influence that cross-linguistic differences in the functionality of pitch have on pitch processing. We will first review the existing literature on the lexical integration of pitch that typically studied non-tone language speakers and/or tone languages with a high functional load for tone.

Quam and Swingley (2010) tested recognition of newly learned words carrying a tone in a bimodal preferential looking experiment adopting a mispronunciation paradigm. The idea behind mispronunciation paradigms is that successful detection of form-meaning mismatches requires the prior establishment of novel representations that include the tonal or segmental

specification of interest. If the lexical representation of the newly acquired word is impoverished or incomplete with respect to for example its tonal specification, word recognition will not be hindered by tonal variability in the input signal.

In their study, English 30-month-old toddlers and adults were taught a novel pseudo-word as a label for a new toy. Subsequently, the target was either correctly pronounced (CP), i.e., with the trained tone, or mispronounced (MP), i.e. with a change in tone (tonal MP) or a change in vowel (vowel MP). Quam and Swingley (2010) showed that both children and adults interpreted the changes in accordance with their native phonology. Word recognition was hindered by a vowel change, as indicated by less target fixation in the vowel MP compared to when the target was pronounced correctly. However, word recognition was not hindered by a change in pitch. At least by 30 months of age, English children have thus learned to disregard pitch at the level of words, which is in line with the non-lexical function of pitch in their L1.

In a paradigm similar to that of Quam and Swingley (2010), Singh et al. (2014) showed that, at 18 months, mono- and bilingual English learners were equally sensitive to tone and vowel MPs, but at 24 months they no longer treated pitch as lexically contrastive, in accordance with their native phonology and in line with the behavior of the 30-month-olds and adults in Quam and Swingley's (2010) study. Mandarin English bilinguals¹⁸ who were dominant in Mandarin were sensitive to both vowel and tonal MPs at both ages. The authors suggest that, at 18 months, toddlers may over-assign weight to post-lexical pitch information in novel word learning due to its high attentional appeal and by virtue of having observed its linguistic significance, either at the post-lexical or at the paralinguistic level.

Similar findings come from a series of experiments by Hay et al. (2015). In an associative word learning task using the two-object switch

¹⁸ From personal communication with the authors, we know that the second language of the Mandarin bilinguals was English.

procedure (Stager & Werker, 1997), 14-month-old but not 17- and 19-month-old learners of English interpreted pitch differences as properties of words. According to Hay et al. (p. 10), between 14 and 17-19 months, children go through a phase of “interpretive narrowing”. With growing linguistic experience, they become more specific about what forms of words should be treated as lexically contrastive. Nevertheless, 17- and 19-month-olds continued to be sensitive to the difference between falling and rising pitch contours in a discrimination task that did not involve label-object mappings. To sum up, the studies above show that there is a shift in English children’s interpretation of the lexical relevance of pitch patterns in the course of the second year of life.

A study that compared the ability to store lexical tones (in this case Mandarin tones) among *adult* speakers of languages differing in their lexical and post-lexical use of prosody is reported in Braun et al. (2014). The languages under investigation (German, Japanese, French, and Mandarin) differed with respect to the *lexical* status of word-level prosody as well as the complexity of the *post-lexical* pitch system (i.e., the number of utterance-level contrasts). German has word stress and thus makes use of word-level prosody. Moreover, it has a relatively rich intonational system. French does not assign word stress to lexical items and would appear to have less pitch variability at the utterance-level. Japanese has word-level prosody in the form of pitch-accents. However, as in French, utterance-level pitch variability is more restricted. Speakers of Mandarin, Japanese, German and French had previously shown sensitivity to Chinese tones in purely perceptual tasks.

The aim in Braun et al. (2014) was to see if the ability to lexically encode pitch in a word learning paradigm depended on experience with lexical or post-lexical prosody. Participants’ recognition of newly learned words was tested in tonal and segmental mismatch conditions. As hypothesized, performance was modulated by the different prosodic structures of the participants’ L1. The Mandarin group outperformed all the other groups. More surprisingly, German participants significantly outperformed Japanese and French listeners. Japanese and French listeners did not differ significantly from

each other. The authors argue that the number of L1 *utterance-level* pitch contrasts, rather than the availability of *word-level* pitch contrasts, are beneficial for building long-term representations of lexical tone. However, German participants might have benefited both from their experience with f_0 as a cue to word stress and as a cue to post-lexical intonation. Importantly, the fact that f_0 is hardly used to signal *lexical* distinctiveness in German obviously does not prevent them from perceiving and lexically encoding non-native pitch information.

Much less is known about the lexical integration of pitch by speakers of more restricted tone languages like Limburgian. In this study, we ask whether pitch plays a role in novel word recognition for child and adult speakers of Limburgian in comparison to a Dutch control group.

5.1.2 The present study

As described in Chapter 2 of this thesis, Limburgian dialects of Dutch are restricted tone systems, distinguishing between a fall (accent 1) and a fall-rise (accent 2). Looking into the lexical encoding of tone in Limburgian can contribute to the discussion on the possible influence of functional load and phonetic variability on the lexical representation of tone. Studying Limburgian speakers' sensitivity to pitch changes within words could also provide evidence for Gussenhoven and Peters' (2008) claim that accent 2 is the underlyingly specified tone. The FEATURALLY UNDERSPECIFIED LEXICON MODEL (Lahiri & Reetz, 2002) can be used to formulate predictions on this matter. If the lexical representation of a newly acquired word is impoverished or incomplete with respect to its tonal specification, tonal features present in the input signal cannot mismatch with an underspecified (i.e. empty) slot in the lexicon. In this case, word recognition cannot be hindered by tonal variability in the input. If it is indeed the case that the Limburgian lexical tone contrast is privative, and accent 2 is the underlyingly specified accent, Limburgians would be sensitive to

mispronunciations of accent 2 (leading to a *mismatch*), but not or to a lesser extent to mispronunciations of accent 1 (leading to a *no-mismatch*).

Another characteristic of Limburgian speakers' input that could have an impact on the specificity of lexical representations is the fact that most Limburgians also speak Dutch and are considered bidialectal (Cornips, 2014). Hardly any studies on the mapping of sounds to meaning focused on children acquiring two languages, let alone on children acquiring multiple dialects or regional varieties of the same language. Extant studies involving both distant- and close-language-bilinguals (for a review, see Fennell, Tsui, & Hudon, 2016) have shown that learning novel minimal pair words in both mono- and bilinguals is favored when children listen to a speaker that sounds like people from their environment (e.g., Fennell & Byers-Heinlein, 2014; Mattock, Polka, Rvachew, & Krehm, 2010). In word recognition studies with known words, the use of cognates can hinder the detection of mispronunciations, at least in close-language bilinguals (e.g., Ramon-Casas & Bosch, 2010). As a consequence of the highly variable input Limburgians are exposed to (Durrant, Delle Luche, Cattani, & Floccia, 2015), the higher probability of hearing accented speech (e.g., Bosch & Ramon-Casas, 2011), and the large amount of lexical overlap in the input (e.g., Sebastián-Gallés & Bosch, 2009), Limburgian children may exhibit a greater acceptance of acoustic variation in phonetic categories.

In this study, we aimed to answer two questions. First, do children acquiring Limburgian encode pitch information as part of their lexical entries when learning novel words? Secondly, do they behave differently from Dutch age-matched peers in this respect? To see whether their interpretation of pitch is adult-like or not yet fully developed, we also tested Limburgian and Dutch adults.

Limburgian and Dutch 2,5- to 4-year-olds (Experiment 1) as well as adults (Experiment 2) participated in a bimodal preferential looking experiment (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Following the procedure employed by Quam and Swingley (2010) and Singh et al. (2014), 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.

In light of previous findings (Singh et al., 2014; Singh, Goh, & Wewalaarachchi, 2015), we expected Limburgians to be sensitive to MPs involving pitch. In light of the existing work on lexical access in pitch-accent languages like Japanese (e.g., Tamaoka et al., 2015), and regarding the relatively restricted nature of the Limburgian tonal system, a change in pitch might only hinder word recognition to a minor extent in Limburgian. Other factors pertaining to the Limburgian speakers' input that could lead to (temporarily) less stable representations, and thus weaker MP effects, are the large amount of surface variation in the contours of the Limburgian tones, phonetic variation due to exposure to multiple regional variants of a language (e.g., Durrant et al., 2015), and possibly also the fair amount of Dutch cognates without a tonal specification (but cf. van der Feest & Johnson, 2016).

As for our Dutch participants, Ramachers et al. (2017; Chapter 3 in this thesis) have shown that Dutch 6- to 12-month-old infants reliably discriminate the Limburgian tones in a discrimination task (see also Liu & Kager, 2014, and Chen & Kager, 2016, for similar results with Dutch infants on Chinese tones). Here we ask whether Dutch participants still attend to pitch in a higher-level task that requires lexical encoding of pitch. We expected that changes in pitch would not hinder their recognition of newly learned words. Previous studies with L1 English children and adults have shown that they ignore pitch information during word learning and recognition, which is in accordance with their native phonology (e.g., Hay et al., 2015; Quam & Swingley, 2010; Singh et al., 2014). However, speakers of Dutch have been shown to be more sensitive to suprasegmental information than speakers of English (Cooper, Cutler, & Wales, 2002). Moreover, adult speakers of German showed sensitivity to word-level pitch differences despite the fact that German has no lexical tone (Braun et al., 2014). It is thus possible that our Dutch participants behave differently from previously tested speakers of English when it comes to the lexical integration of

word-level pitch. Also, de Bree, van Alphen, Fikkert, and Wijnen (2008) showed that Dutch 36-month-olds were sensitive to miss-stressing. The fact that 3-year-old Dutch children appear to be sensitive to word-level suprasegmental properties might also facilitate their encoding of other word-level prosodic features, like lexical tone.

For the adults, in principle the same expectations hold. However, due to accumulated linguistic experience, Limburgian adults might have learned not to rely on pitch alone during online language comprehension. We expected Limburgian adults to notice a change in tone, but it is an empirical question how strongly it will hinder word recognition. Dutch adults have been shown to be sensitive to Limburgian tones in an AXB discrimination task (see Chapter 4 of this thesis). In a word learning context, Dutch adults might also still show sensitivity to pitch differences by virtue of their accumulated linguistic experience with post-lexical intonation and word stress (but cf. Quam & Swingley, 2010).

First, we address the experiment with Limburgian and Dutch toddlers (Experiment 1). Secondly, the experiment with Limburgian and Dutch adults will be discussed (Experiment 2).

5.2 Experiment 1

5.2.1 Participants

A total number of 41 Limburgian toddlers were recruited via health care institutions and daycare centers in the city of Roermond, Limburg, the Netherlands. Twenty-three children with a mean age of 40.9 months ($SD = 5.9$ months; range = 31 months to 49 months; 6 boys) were included in the analysis. An additional 18 toddlers were tested but excluded from analysis because they failed to contribute sufficient data. For a detailed description of trial, block and participant exclusion criteria we refer to section 2.1.6 on data preprocessing and Table 1 in the Appendix.

Children in Limburg are often exposed to quite heterogeneous linguistic input. As a result, it is difficult to find toddlers who have only been exposed to one particular dialect, in our case Roermond Dutch. Children from the municipality of Roermond who were exposed to any East-Limburgian dialect (Bakker & van Hout, 2012), spoken by at least one parent or caregiver, were allowed to participate. We can assume that the realization of the word prosodic contrast within the East-Limburgian dialect region does not show much variation (Heijmans, 2003). Based on parental report (missing $N = 1$), using an adapted version of the PaBiQ (COST Action IS0804, 2011)¹⁹ administered during a telephone interview, the language input provided at home to 22 of the Limburgian children was as follows: (1) both parents speak a different East-Limburgian dialect ($N = 9$), (2) one parent speaks an East-Limburgian dialect, the other Standard Dutch ($N = 8$), (3) both parents speak the same East-Limburgian dialect ($N = 3$) and (4) one parent speaks an East-Limburgian dialect, the other a dialect from another Limburgian dialect region ($N = 2$). All children were reported to understand both Limburgian and Dutch. Moreover, 19 out of 22 children were reported to speak Limburgian, and all participants were reported to speak Dutch. All Limburgian toddlers thus picked up on Dutch, even if they were not addressed in it by (one of) their parents, but for example by friends or at daycare. All toddlers could thus be considered bidialectals. For language use in the home (input quantity) parents were asked a series of questions with rating scale responses about the languages used by each household member to the child. From this, a proportion of language use in the home was derived. The questionnaire also contained a language richness measure (input quality), as defined by the extent to which children were exposed to story-telling, either as read from books or produced spontaneously,

¹⁹ This questionnaire is a translation/adaptation of the Questionnaire for Parents of Bilingual Children (COST Action IS0804, 2011). It is the short version of a longer questionnaire piloted by research groups in several countries within COST Action IS0804, which was in part based on the ALEQ (Paradis, 2011) and the ALDeQ (Paradis, Emmerzael, & Sorenson Duncan, 2010).

the expression of feelings, educational games (e.g., counting, spelling), labeling new objects, and media (e.g., television, PC, tablet). 18 out of 22 children had higher input quantity scores in Limburgian than in Dutch. 17 out of 22 children had higher or equal input quality scores in Limburgian than in Dutch. See Table 2 in the Appendix for more details.

A total number of 40 Dutch toddlers were recruited from the subject pool of the Baby Research Center of Radboud University in Nijmegen, the Netherlands. All infants grew up in monolingual Standard Dutch-speaking families. Thirty-five toddlers with a mean age of 36.8 months ($SD = 1.8$ months; range = 34 months to 40 months; 13 boys) were included in the analysis. An additional 5 participants were excluded from the analysis for not contributing enough data ($N = 4$) and because one pair of children were twins ($N = 1$; the child contributing the least number of trials was excluded).

To make sure that none of the Dutch toddlers had substantial experience with a Limburgian dialect or any other tone language, their parents were asked questions related to the linguistic input of their child during an intake phone call. A child was regarded to have substantial experience with a tone language and thus not suitable for participation if (1) one of the parents or primary caregivers was a native speaker of a tone language, (2) the child had weekly contact with a native tone language speaker.

None of the participants had known developmental disorders or delays and none of them had substantial exposure to a language other than Limburgian or Dutch. Ethical approval for the study was obtained from the Ethiek Commissie Faculteit der Sociale Wetenschappen (ECSW) at Radboud University in Nijmegen. Caregivers signed an informed consent and received a picture book or a small monetary compensation for their participation.

5.2.2 Apparatus

Limburgian children were tested in a dimly lit office using a portable lab set-up in a daycare center in Roermond. They sat in front of a 24-inch LCD screen



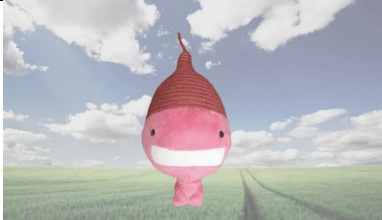

(Philips 249C4QHSB) and were recorded via a digital video camera (Sony HC40) mounted on a tripod below the table. Behind the monitor were two speakers (Logitech Z130). The video camera broadcast the recording to a 13-inch Apple MacBook Air. Recordings were made with the video software Vidi (version 0.4.7). The experiment was presented using the LOOK software (Meints & Woodford, 2008), run on a laptop (HP EliteBook Folio 9470m). During testing, the caregiver listened to masking music through noise-cancelling headphones (Sennheiser HME 110).

Dutch children were tested in a dimly lit room in the Baby Research Center at Radboud University, Nijmegen. The experiment was run in a test booth (size: 128 x 177cm), which is partly closed by black wooden partitions, left and right from the 47-inch television screen (LG 47LK530 ZC). A digital video camera (Sony Handycam DCR_HC85E PAL) was placed 30cm below the screen, hidden by a black curtain with an opening for the lens. The video camera provided a broadcast of the infant's behavior to a monitor behind the TV. Recordings of the infant for offline coding were made using Virtual Dub (Version 1.9.11). The experiment was controlled using the LOOK software (Meints & Woodford, 2008). The caregiver wore noise-cancelling headphones (Sennheiser HMEC 300) that played masking music.

5.2.3 Procedure

The procedure employed was the intermodal preferential looking paradigm (Golinkoff et al., 1987). The experiment lasted approximately ten minutes and consisted of two blocks, separated by a one-minute break. In each block, children would learn one novel word-object mapping. Subsequently, it was tested how they reacted to a pitch change in the newly learned word. Each child thus learned two new words, one with accent 1 and one with accent 2. Half of the participants learned the accent 1 word first and half learned the accent 2 word first. Each block featured a different pair of objects. A visual overview of a block is presented in Figure 4.

180 Setting the Tone

Phase	No. of trials	Example of visual stimuli	Example of auditory stimuli
Introduction	1		<p><i>Hey hallo! Zullen we een spelletje spelen? Let goed op!</i> ('Hey hello! Would you like to play a game? Pay attention!')</p>
Object Familiarization	1		<p><i>Kijk eens hier! Wat zijn dat? Die zijn mooi! Vind jij ze ook leuk?</i> ('Look! What are those? They look great! Do you like them too?')</p>
Learning	4		<p>2 x Target: <i>Kijk! Dit is een taaf₁. Een taaf₁. Zie je hem? Daar is de taaf₁... etc.</i> ('Look! This is a taaf₁! A taaf₁. Do you see it? There's the taaf₁...' etc.)</p>
			<p>2 x Distracter: <i>Ooo woow! Die ziet er leuk uit! Wat zou dat zijn?... etc.</i> ('Ooo woow! That looks great! What could that be?...' etc.)</p>



Test	8		<p>2 x CP: <i>Kijk naar de taaf₁! Zie je hem?</i> ('Look at the <i>taaf₁</i>! Do you see it?')</p> <p>2 x MP: <i>Kijk naar de taaf₂! Zie je hem?</i> ('Look at the <i>taaf₂</i>! Do you see it?')</p>
			<p>4 x Filler, e.g.: <i>Waar is de auto? Kun je hem vinden?</i> ('Where is the car? Can you find it?')</p>

Figure 4. Visual overview of an experimental block.

A block started with an encouraging introduction phase featuring a young girl waving and smiling at the participant, inviting him/her to play a game. In the following object familiarization phase, the child was familiarized with two novel toy objects appearing simultaneously at the far left and far right side of the screen. The objects were presented for nine seconds. The child heard (in Limburgian or in Dutch): "Look! What are those? They look great! Do you like them too?" One of these objects (henceforth: the target) would be labeled in the subsequent learning phase. The other one (henceforth: the distracter) would remain nameless. Target side during object familiarization was counterbalanced across blocks. The purpose of this phase was twofold: familiarization of stimuli prior to labeling usually boosts levels of retention (e.g., Hilton & Westermann, 2016) and it lowers the task demand (e.g., Fennell, 2012).

After object familiarization, the child proceeded to the learning phase. During this ostensive-labeling phase, participants were taught a new word carrying either accent 1 or accent 2. The phase consisted of four trials of 30 s

each. In the first and the third trial, the target appeared bouncing in front of a natural landscape and was labeled ten times in each trial in sentences like: “Look! This is a [target]! A [target]! Can you see it? There’s the [target]!” In total, the child heard twenty repetitions of the target label. Presenting a number of repetitions is in line with previous research on retention of novel word-object mappings (e.g., Quam & Swingley, 2010; Singh et al., 2014; Hilton & Westermann, 2016). Note that the target label always appeared in focus-final position in a declarative sentence. In this way, the phonetic realization of the Limburgian tones was held constant, and the child thus did not have to abstract away from different surface realizations. In trials two and four, the distracter object appeared in the same scenario and was talked about for an equal amount of time, but crucially, it did not receive a label. We tried to encourage the child to wonder what the name of the distracter was. The target and distracter object were presented for an equal amount of time to prevent a familiarity preference for one of both objects in the subsequent test phase. The order of trials was the same across blocks and participants.

Following the learning phase, the child entered the test phase that consisted of four test trials and four filler trials. In test trials, the target and the distracter toy appeared side by side on the screen. Children were asked to “Look at the [target].” The onset of the target was always at 2500 ms to enable children to inspect both objects before naming and to establish a baseline preference. To maximize engagement, a second sentence like: “Can you find it?” followed 1000 ms after target offset. Test trials lasted seven seconds.

In two of the test trials, the label for the target object was correctly pronounced (Correct Pronunciation (CP) trials), while in the other two, the label was mispronounced (Mispronunciation (MP) trials). This MP involved a change in pitch: A word taught with accent 1 was mispronounced with accent 2 and vice versa. Recall that during test trials the novel target item was paired with a novel, unlabeled distracter item. 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. This

presupposes the use of the principle of mutual exclusivity (ME; Markman, 1990). This principle guides people to map novel words to unfamiliar rather than familiar referents. The use of ME to identify referents of novel words has been reliably demonstrated in infants from 16 months of age (e.g., Halberda, 2003) and in monolingual, bilingual and bidialectal preschool children (e.g., Diesendruck & Markson, 2001; Durrant, 2014; Kalashnikova, Mattock, & Monaghan, 2015; Markman & Wachtel, 1988; Singh et al., 2014). The procedure with a novel target and a novel distracter object has been successfully applied in similar word learning studies with one-and-a-half to two-year-olds (Singh et al., 2014), two-and-a-half-year-olds (Quam & Swingley, 2010) and three- to five-year-olds (Singh & Quam, 2016).

Order of CP and MP trials was pseudo-randomized in such a way that the target would never appear on the same side more than twice in a row. Moreover, all children were presented at least one CP trial before the first MP trial. This resulted in three trial orders. To make sure children would remain engaged in the task, four filler trials involved correct pronunciations of four well-known words (e.g., Buckler & Fikkert, 2016; Singh et al., 2015). Test phases across all versions started with a filler trial to help children understand the nature of the task. Test and filler trials were presented in an alternating fashion.

Between blocks, children watched a one-minute video featuring farm animals and animal noises. The second block had the same structure as the first block but featured a new object-pair, one of which would receive a novel label. Object labels and tones were counterbalanced across participants. Each child was thus tested on his/her sensitivity to tonal MPs of accent 1 and accent 2 to test for asymmetries in tone sensitivity (e.g., Francis & Ciocca, 2003; Shi, Gao, Achim, & Li, 2017). Throughout the experiment, trials were preceded by a purple flashing light in the screen center and were initiated once the child fixated the attention getter.

5.2.4 Stimuli

For this experiment, we created two pseudo-word pairs: *taaf*_{1/2} [ta:f] and *moon*_{1/2} [mo:n].²⁰ We decided to teach each participant two words instead of one to reduce the possibility that any effects were idiosyncratic to a particular word. Moreover, in this way all participants could learn one word with accent 1 and one word with accent 2.

The segments and phonotactics of the target stimuli were equally compatible with Limburgian and Dutch, and both pseudo-word pairs were derived from existing tonal minimal pairs in Limburgian to ensure that they were legal with both tones. Additionally, we controlled for phonological neighborhood density, since the existence of phonological neighbors could hinder children from using their full phonological sensitivity (e.g., Swingley & Aslin, 2007; Swingley, Pinto, & Fernald, 1998) or from using the principle of ME (e.g., Jarvis, Merriman, Barnett, Hanba, & van Haitsma, 2004). We considered a word a phonological neighbor if the item differed from the novel word by substituting, adding or deleting a single phoneme (Luce & Pisoni, 1998; Swingley & Aslin, 2002). We only considered words from the Lexilijst Nederlands (Schlichting & Spelberg, 2002) that are supposed to be produced and known by 15- to 27-month-old Dutch children. *Taaf* had no phonological neighbors known to children of this age, whereas *moon* had one phonological neighbor for the Dutch participants (*maan* [ma:n], ‘moon’), and two for the Limburgian participants (*maon*₁ [mɔ:n], ‘moon’; *sjoon*₂ [ʃo:n], ‘shoe’).

Carrier sentences were recorded in Limburgian and Dutch. Target stimuli were recorded in and spliced from Limburgian carrier sentences to guarantee tone accuracy. All stimuli were recorded in a child-friendly way by a female native speaker of Dutch and of an East-Limburgian dialect spoken in the municipality of Roermond. She reported to be dominant in Limburgian, but was equally proficient in Dutch and was trained in speaking accentless Standard

²⁰ Subscripts indicate accents 1 and 2.

Dutch. For Limburgian children, pre-experimental instructions as well as the experiment itself were in Limburgian. For Dutch children, the entire procedure was in Dutch. Across language contexts, only the tokens of the target stimuli *taaf* and *moon* were the same. Care was taken that the Dutch and Limburgian stimuli were recorded with the same intent and enthusiasm. The target stimuli were recorded multiple times with accent 1 as well as accent 2 and always appeared in a declarative focus-final context to avoid differences in the phonetic realization of the tones. Recordings were made in a sound-attenuated booth using Adobe Audition (version CS6, 44.1 kHz). Stimuli were equalized for intensity to 65dB and prepared for the experiment using Praat (version 5.3.35; Boersma & Weenink, 2012). For stimuli excision we followed the guidelines presented in Turk, Nakai, and Sugahare (2006).

In total, 12 tokens of *taaf₁*, *taaf₂*, *moon₁* and *moon₂* were selected, based on intuition of a native speaker of Limburgian [the first author, SR] and careful listening by a trained phonetician [Carlos Gussenhoven]. Ten tokens were used in the learning phase, the remaining two in the test phase. For all tokens we measured maximum and minimum f_0 , f_0 range (max f_0 - min f_0), average f_0 and duration of the tone bearing portion as well as the duration of the entire token. Measurements were done manually, taking auditory as well as spectral properties into account. Independent *t*-tests revealed that accent 1 and accent 2 tokens differed significantly from each other with respect to minimum f_0 , maximum f_0 and f_0 range (see Table 3 in the Appendix).

The four filler trials involved correct pronunciations of known words. One filler pair consisted of a cow and a horse, and the other of a car and a ball. Items were chosen for their very high frequency in the productive vocabulary of the age group at test, according to the Lexilijst Nederlands (Schlichting & Spelberg, 2002).

The visual target stimuli consisted of four plush toy objects of an animate character (see Figure 5). All objects had different, vibrant colors (pink, blue, purple and yellow) and shapes. The pink and blue object (A and B in

Figure 5) were paired as well as the purple and yellow object (C and D in Figure 5). Pairs were matched in visual complexity, brightness and size. A paired-samples t-test comparing the mean proportion of looking time towards the target ($M = .51, SD = .08$) and the distracter object ($M = .50, SD = .08$) during the object familiarization phase showed that participants did not show a preference for the target object prior to the learning (i.e., labeling) phase ($t(57) = .59, p > .05$).

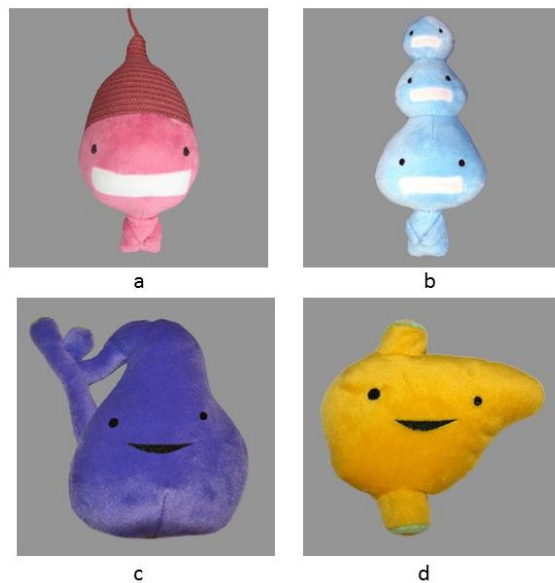


Figure 5. The visual target stimuli used in the experiment.

In the object familiarization phase and the test phase, the stimuli consisted of photographs of the objects against a grey background. During the learning phase, the objects bumped up and down against the background of a natural scene. Filler stimuli in the test phase consisted of photographs of a horse, a cow, a car and a ball against a grey background. Two different pictures per object were used across blocks to minimize boredom effects.

5.2.5 Data pre-processing and analysis

Children's video recordings were coded offline using ELAN (version 4.5.0; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) with a resolution of 40 fps. In test trials, target onset was always at 2500 ms. The 2500 ms phase prior to target onset was labeled the pre-naming window. The post-naming window lasted 2000 ms, starting 367 ms after target onset (e.g., Altvater-Mackensen, van der Feest, & Fikkert, 2013; Quam & Swingley, 2010; Singh et al., 2014; Swingley & Aslin, 2000). The coder was blind to trial type and target side. A random 20% of the videos were recoded by a second experienced coder. The correlation between two coders was very strong (Pearson's $r = .801$, $p < .001$).

To ensure that our analyses were based on clean data and to enable within-subject comparisons of CP vs. MP trials and of accent 1 vs. accent 2 words, we maintained a number of trial, block, and participant exclusion criteria. Table 1 in the Appendix provides a detailed overview of exclusion.

Test trials were excluded if (1) a child looked less than 500 ms during the 2000 ms post-naming window (e.g., Quam & Swingley, 2010; Singh et al., 2014; Tsuji, Fikkert, Yamane, Mazuka, 2016), (2) the participant fixated only one of two objects during the 2500 ms pre-naming window (e.g., Buckler & Fikkert, 2016; Mani & Plunkett, 2011; Singh et al., 2015; White & Morgan, 2008), (3) an equipment or experimenter error occurred, and (4) if a participant refused to participate (e.g., by getting up and walking around) and the experiment had to be aborted.

A block was excluded if (1) a participant did not contribute at least one valid trial per condition (CP and MP) during the test phase (e.g., Buckler & Fikkert, 2016; Tsuji et al., 2016), and (2) total looking time during target and/or distracter learning trials was under 20 s out of a total of 60 s (e.g., Tsuji et al., 2016). The latter criterion is based on the assumption that children who pay more attention to the novel objects during learning should be better able to retain the novel word-object mapping (Hilton & Westermann, 2016).

Participants were excluded from the analyses if (1) at least one block had to be excluded, (2) an equipment failure or experimenter error occurred, and (3) other conditions were not met, e.g. if a participant's linguistic background was inappropriate or if we had twin participants.

Children's target recognition was inferred from the presence of a naming effect that is typically measured as an increase in target fixation upon hearing the target label relative to a baseline looking measure (e.g., Singh et al., 2015; Swingley & Aslin, 2000). There are several ways to compute a naming effect. In most cases, a proportion of target looking (PTL) measure is used. PTL is calculated by dividing the total amount of time children spent looking at the target (T) after target naming by the total amount of time children spent looking at the target and distracter (T+D) within the same time window.

To calculate the naming effect, the increase in PTL between the pre-naming and post-naming window of a test trial was calculated (i.e., $\text{Post-naming PTL}_{(T/[T+D])} - \text{Pre-naming PTL}_{(T/[T+D])}$), resulting in a difference score. Computing naming effects by taking each individual participants' pre-naming values into account serves to control for possible effects of preference for a particular stimulus (e.g., Mani & Plunkett, 2011; Quam & Swingley, 2010; Singh et al., 2015; White & Morgan, 2008).

A paired-samples *t*-test showed a small yet significant difference in PTL between object familiarization ($M = .51, SD = .08$) and pre-naming window ($M = .53, SD = .07$), $t(57) = -2.05, p = .045$, Cohen's $d = -0.27$. Moreover, a one-sample *t*-test showed that pre-naming PTL differed significantly from chance: $t(57) = 3.56, p = .001$, Cohen's $d = 0.47$. Thus, it appears that the target object had become slightly more interesting than the distracter after the learning phase due to repeated labeling (e.g., Schafer & Plunkett, 1998). To control for a possible effect of this target preference, we chose the post- minus pre-naming PTL measure (difference score) as our dependent variable, henceforth called the NAMING EFFECT.

Naming effects were calculated and compared for CP and MP trials. If children notice the MP, the naming effect will be significantly less strong in MP

than in CP trials. However, it is important to inspect the naming effect in MP trials more closely to gain insight into the strength of the MP effect. First, even if the naming effect in MP trials is significantly weaker than the naming effect in CP trials, it can still be positive and significantly above zero (as attested for one-feature segmental MPs in Mani & Plunkett, 2011, and White & Morgan, 2008). This indicates that target recognition is hindered to some extent, but that recognition still takes place. Secondly, the naming effect in MP trials might not differ significantly from 0, signaling uncertainty, meaning that target recognition is hindered to such extent that recognition fails (as attested for two- and three-feature segmental MPs in Mani & Plunkett, 2011, White & Morgan, 2008, and for tonal MPs in Singh et al., 2014, 2015). Thirdly, a significant negative naming effect would point to a preference for the distracter object and can be seen as evidence for the formation of a novel mapping between the auditory label and the distracter object based on ME (e.g., Mani & Plunkett, 2011; Swingley & Aslin, 2000; White & Morgan, 2008).

5.2.6 Results

Figure 6 shows naming effects for Limburgian and Dutch toddlers in the CP and MP condition.

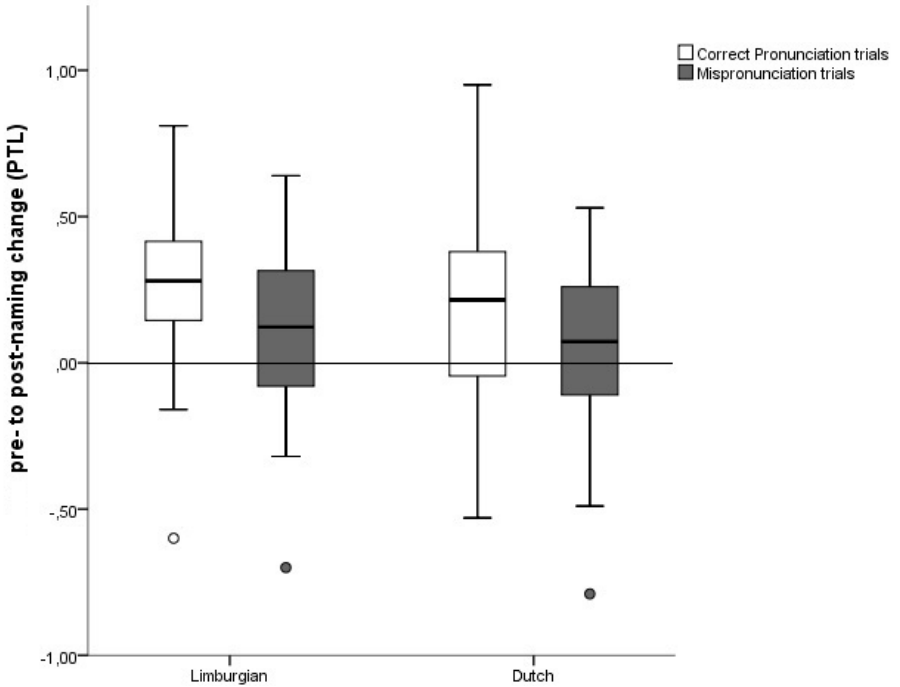


Figure 6. Boxplot showing the median and interquartile range of the pre- to post-naming change (post-naming PTL minus pre-naming PTL) in CP and MP trials for Limburgian and Dutch toddlers. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers. The solid horizontal line indicates chance level.

To ensure that word learning was successful, the naming effect in CP trials was compared to zero for each language group by means of a one-sample *t*-test. For both Limburgian and Dutch toddlers, there was a significant positive naming effect in CP trials (Limburgian: $M = .25$, $SD = .15$, $t(22) = 8.28$, $p < .001$, Cohen's $d = 1.73$; Dutch: $M = .18$, $SD = .23$, $t(34) = 4.60$, $p < .001$, Cohen's $d = 0.78$). From this we can conclude that both participant groups learned the novel word-object mapping.

Next, 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) as the between-subjects factor was conducted to evaluate the possible influence of language and pitch change on the naming effect. Results revealed a

significant main effect of Condition, $F(1,56) = 8.53$, $p = .005$, $\eta_p^2 = .13$, observed power = .82, with a significantly larger increase in PTL 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). Both Limburgian and Dutch children thus treated the pitch change as lexically relevant as indicated by a significantly weaker naming effect in MP trials compared to CP trials in both groups. Mean PTL values and standard deviations for pre- and post-naming windows per Condition and Language are listed in Table 1.

Table 1. Mean proportion of target looking in pre- and post-naming windows per group and condition for the toddlers. Standard deviations are in parentheses.

PTL (SD)	Limburgian	Dutch
CP Pre-naming	0.51 (.08)	0.51 (.10)
CP Post-naming	0.76 (.13)	0.69 (.20)
MP Pre-naming	0.58 (.10)	0.55 (.10)
MP Post-naming	0.70 (.20)	0.62 (.22)

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 $d = .37$). Thus, despite the naming effect being stronger in CP than MP trials, target recognition was still possible in MP trials. From this we can infer that the pitch change only hindered word recognition to a minor extent.²¹

²¹ Some previous studies found age-related differences in the sensitivity to pitch changes in tone language learning bilinguals (e.g., Singh et al., 2015). Since we also tested tone language learning 'bilinguals' spanning exactly this age range, we ran an additional mixed ANOVA on our Limburgian sample including Age as a within—subjects variable, comparing younger (31-38 months, $N = 11$) to older (42-49 months, $N = 12$) children.

We next tested Limburgian and Dutch adults in the same experiment to find out whether the sensitivity to pitch in both the Limburgian and Dutch children in Experiment 1 was adult-like or whether it reflected a not yet fully developed phonological system.

5.3 Experiment 2

As for the Limburgian children, we expected Limburgian adults to notice a change in tone, but it was an open question how strongly it would hinder word recognition. Adult speakers might have learned not to rely on pitch too much during online language comprehension because of the relatively low functional load of lexical tone and because pitch has no lexical relevance in their second L1, Dutch.

Speakers of Dutch were expected not to attend to pitch during the recognition of newly learned words. However, if the sensitivity exhibited by the Dutch children was dependent on their knowledge of pitch as a cue to word stress and/or intonation, Dutch adults might be even more sensitive to pitch differences by virtue of their accumulated experience with the native prosodic system.

5.3.1 Participants

Limburgian adults were recruited and tested in a public library in Roermond. The Limburgian listeners ($N = 14$, 5 males) ranged in age from 26 to 72 years ($M = 53.6$ years). An additional 10 participants were excluded from the analysis because 1) they reported to speak a dialect other than one from the East-Limburgian dialect region ($N = 4$), 2) they could only contribute one of two blocks due to exclusion of test trials ($N = 3$), or 3) they failed to learn the novel

The analysis yielded a main effect of Condition, $F(1,21) = 4.63$, $p = .04$. No other significant main effects or interactions were attested (all $ps > .05$). Due to the small sample sizes, this result should be interpreted with caution. However, the descriptives showed no trend towards an effect or interaction involving the variable Age.

word-object mapping in one or two blocks, signaled by a mean PTL equal or smaller than .50 in the post-naming phase of CP trials ($N = 3$). All included Limburgian participants were born and raised in the East-Limburgian dialect region and lived there at the time of test. All of them reported to actively use an East-Limburgian dialect. The Limburgian participants also had native command of Dutch, except for two participants who reported very good or good command. All of them can thus be considered bidialectals.

Dutch adults were recruited via a participant recruitment system at Radboud University, Nijmegen, and tested at the Baby Research Center of the same university. The Dutch listeners ($N = 22$, 7 males) ranged in age from 18 to 40 years ($M = 23$). None of them had weekly contact with people speaking a Limburgian dialect in their presence. Moreover, none of them grew up or lived in the province of Limburg. An additional 2 participants were excluded from the analysis due to the exclusion of one of both blocks.

All Limburgian and Dutch participants reported some degree of non-native command of one or more non-tonal languages (i.e., English, German, French, Spanish, Arabic, Polish) as indicated on a 6-point scale ranging from poor to native command, but none of them had experience with a tone language. All participants reported normal hearing and no speech, language, or attention deficits. Because of the fact that musical experience can have an influence on pitch processing (e.g., Burnham & Brooker, 2002; Burnham, Brooker, & Reid, 2015), we kept the number of musically trained individuals comparable across groups. Six of the Limburgian participants (43%) and eight of the Dutch participants (36%) reported to have had over three years of musical training. Ethical approval for the study was obtained from the Ethics Assessment Committee (EAC) of the Faculty of Arts at Radboud University, Nijmegen. Participants signed an informed consent and took part in the experiment either voluntarily or for a small fee.

5.3.2 Apparatus, procedure, and stimuli

The apparatus, procedure, and stimuli of the adult experiment were comparable to Experiment 1, as in Quam and Swingley (2010), who also tested children and adults under similar conditions. For the Limburgian adults we used the same portable set-up as the Limburgian children, but they were tested in a quiet, darkened room in a public library. To minimize external interference, stimuli were presented through noise-cancelling headphones (Sennheiser HME 110). Dutch adults were tested under the exact same conditions as the Dutch children.

Regarding the procedure, we added extra filler trials (16 instead of 4) to the test phase to distract adult participants' attention away from the purpose of the experiment, leading to a total number of 20 trials. Participants were told before the study that they would be helping to test an experiment designed for 3-year-olds.

A paired-samples *t*-test, comparing the mean PTL towards the target ($M = .51, SD = .05$) and the distracter object ($M = .49, SD = .05$) during the object familiarization phase, showed that adult participants did not show a preference for the target object prior to the learning (i.e., labeling) phase ($t(35) = .73, p > .1$). After the experiment, adults completed a language background questionnaire.

5.3.3 Data pre-processing and analysis

A random 20% of the adult videos were recoded by a second experienced coder. Inter-coder reliability was excellent (Pearson's $r = .937, p < .001$).

Post-naming PTL was calculated within a 1000 ms window, starting 367 ms after target onset. We could have shifted the analysis window for adults earlier in time, but since it has been shown that this does not have

consequences for the results (Swingley, 2009), we retained the starting point of 367 ms post-target onset.²²

As with the child data, we found a significant difference in PTL during object familiarization ($M = .51$, $SD = .05$) and pre-naming phase ($M = .56$, $SD = .12$), $t(35) = -2.73$, $p = .01$, Cohen's $d = -0.45$. Moreover, a one-sample t -test showed that pre-naming PTL differed significantly from chance: $t(35) = 3.16$, $p = .003$, Cohen's $d = 0.53$. Thus, it appears that also for the adults the target object had become more interesting than the distracter after the learning phase. We again chose the post-naming minus pre-naming PTL measure as our dependent variable.

5.3.4 Results

Naming effects for Limburgian and Dutch adults in CP and MP conditions are depicted in Figure 7.

²² A post-hoc inspection of the adults' looking behavior in an earlier time window, starting 200 ms post-target onset, revealed that adults were already on target. An additional analysis with a slightly earlier time window would thus probably not have changed the results.

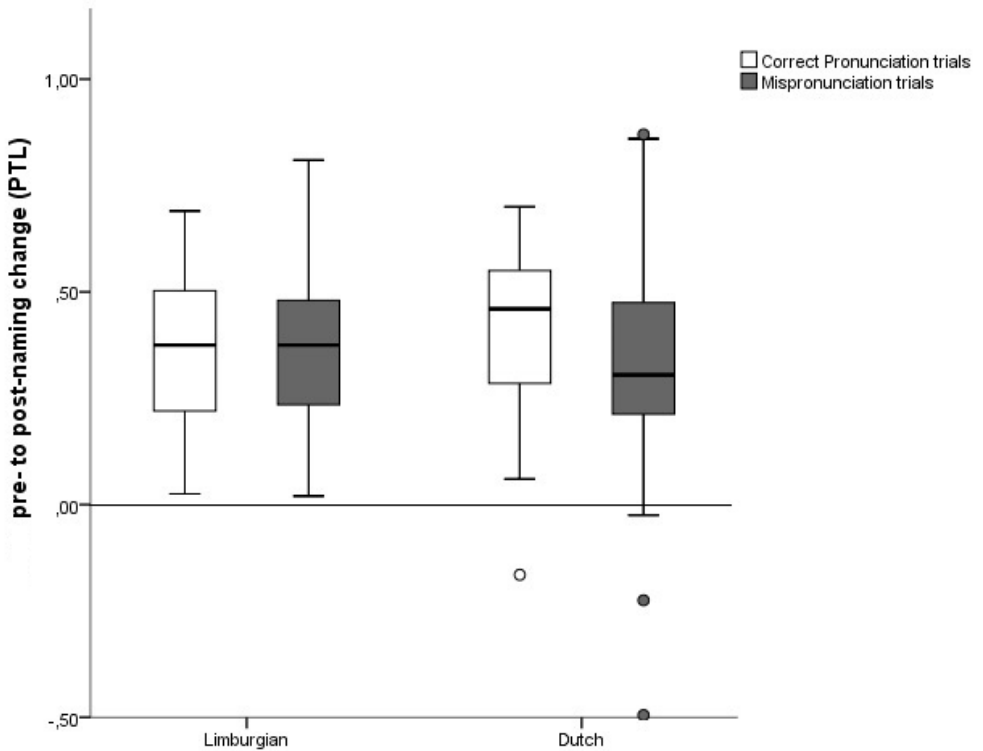


Figure 7. Boxplot showing the median and interquartile range of the pre- to post-naming change (post-naming PTL minus pre-naming PTL) in CP and MP trials for Limburgian and Dutch adults. Whiskers show the range between which the highest and lowest 25% of scores fall. Dots represent outliers. The solid horizontal line indicates chance level.

To ensure that the adult participants successfully learned the novel word-object pairings, the naming effect in CP trials was first compared to zero for each language group by means of a one-sample *t*-test. For both Limburgian and Dutch adults, there was a significant positive naming effect in CP trials (Limburgian: $M = .36$, $SD = .13$, $t(14) = 10.69$, $p < .001$, Cohen's $d = 2.86$; Dutch: $M = .41$, $SD = .14$, $t(22) = 14.28$, $p < .001$, Cohen's $d = 3.04$). From this we can conclude that both participant groups learned the novel word-object mappings.

Next, 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) as the between-subjects factor was conducted. The analysis yielded no main effects nor 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 $d = 1.53$).

The absence of an effect of Condition or Language is probably due to participants showing very strong naming effects in both CP and MP trials, as becomes clear from the PTL measures in Table 2.

Table 2. Mean proportion of target looking in pre- and post-naming windows per group and condition for the adult participants. Standard deviations in parentheses.

PTL (SD)	Limburgian	Dutch
CP Pre-naming	0.56 (.12)	0.52 (.13)
CP Post-naming	0.93 (.12)	0.94 (.10)
MP Pre-naming	0.59 (.17)	0.58 (.14)
MP Post-naming	0.94 (.10)	0.93 (.14)

As can be inferred from Quam and Swingley (2010), the procedure used is sensitive enough to yield a vowel MP effect. However, Quam and Swingley did not test native tone-language speakers and thus did not show whether the method is equally suited to yield sensitivity to a change in pitch. This means that we cannot rule out the possibility that our findings are due to a task effect.

Our adult data thus provide no evidence of an effect of pitch variation on the recognition of newly learned words.

5.4 Discussion

In this study, we asked whether pitch plays a larger role in novel word learning and recognition in children acquiring East-Limburgian compared to a control group of children acquiring Standard Dutch. To see whether their interpretation of pitch was adult-like or not yet fully developed, we also tested Limburgian and Dutch adults.

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 change in tone did not hinder word recognition to a great extent. Regarding our adult data, we can conclude that both Limburgian and Dutch adults succeeded in learning novel word-object mappings. However, we cannot draw conclusions about their interpretation of pitch changes due to very strong naming effects in both CP and MP conditions. In the next section, we will first discuss the findings from Experiment 1 with Limburgian and Dutch toddlers.

5.4.1 The lexical encoding of pitch in Limburgian and Dutch toddlers

The finding that Limburgian children were sensitive to MPs involving pitch was in line with previous word recognition studies with tone language learners (Singh et al., 2014; Singh et al., 2015). However, as signaled by the positive naming effect in MP trials, the pitch change did not fully inhibit target recognition. This pattern of results is in line with toddlers' responses to 1-feature segmental MPs in Mani and Plunkett (2011) and White and Morgan (2008). However, previous studies investigating Mandarin found no naming effects in tonal MP conditions (Singh et al., 2014; Singh et al., 2015), suggesting that pitch changes are more detrimental to word recognition in Mandarin than in Limburgian. We would like to suggest three explanations for this finding.

First, the fact that Limburgian children recognized the target word despite a tonal change might be due to the relatively low functional load of

tone. One of the factors contributing to the functional load of a contrast is the number of minimal pairs. The low frequency of tonal minimal pairs, plus the fact that listeners can mostly rely on sentence context for disambiguation, might mitigate the reliance on pitch in perceiving Limburgian. Similar explanations have been put forward by Cutler (1986) for the role of lexical stress in English and by Cutler and Otake (1999), Goss (2015), and Sekiguchi and Nakajima (1999) for the influence of pitch-accent on word recognition in Japanese. This reasoning is in line with the hypothesis that phonological category learning is driven by contrast in the vocabulary (e.g., Dietrich, Swingley, & Werker, 2007). However, Dietrich and colleagues argue on the basis of the results of a word recognition study that 18-month-olds' native-like performance cannot have been the result of top-down information from the lexicon. The tested age group did not seem to know many minimal pairs involving the distinctions at test. We thus cannot assume that children need minimal pairs to decide whether a contrast is phonologically meaningful or not.

A second explanation for the Limburgians' lenient treatment of MPs might be tonal surface variability. Recall that Limburgian listeners are confronted with a considerable amount of context-induced *allotonic* variation in lexical tone contours, but this variation cannot be ignored since it does signal meaningful information at the post-lexical level. In light of this variation, it could be a challenge to recover the underlying tone system, at least for young learners (Demuth, 1995; Ota, 2003; Rost & McMurray, 2010). A replication of our study with (monolingual) Swedish children could provide insight into the effect of surface variation on developing tonal representations.

A third factor that may have influenced our Limburgian participants' behavior is variation due to their exposure to multiple (closely related) linguistic varieties. Hardly any studies on the mapping of sounds to meaning focused on children acquiring two languages, let alone on children acquiring multiple dialects or regional varieties of the same language (for a review, see Fennell et al., 2016). One type of variation due to bidialectalism comes from

exposure to different dialects and Limburgian-accented Dutch. Evidence for the effects of dialect-related variation on the phonological representation of known words is scarce. Durrant et al. (2015) showed that variable phonological input as a result of dialect variation has an impact on the specificity of lexical representations in 20-month-old British English multidialectal toddlers. In a preferential looking paradigm, they were tested on their sensitivity to single feature MPs of monosyllabic known words. MPs involved changes of onset consonants or of the vowel nuclei that were phonemic in all the varieties at test. The authors' main finding was that multidialectal infants, other than monodialectal infants, did not treat MPs of familiar words differently from CPs, suggesting that long-term exposure to regional linguistic variation leads to a broadening of phonetic categories or poorer use of phonological information in word recognition.

Another type of variation due to bidialectalism stems from lexical overlap. Limburgians know many cognates that do not have a tonal specification in Dutch. As such, they receive mixed evidence for the lexical relevance of pitch. Possibly, this mixed evidence (temporarily) leads them to assign less weight to pitch as a lexically contrastive feature. The existing evidence points in another direction, though. Van der Feest and Johnson (2016) tested 24-month-old Dutch toddlers who received mixed distributional evidence for the lexical contrastivity of fricative voicing. Toddlers were exposed to Limburgian-accented Dutch (which maintains the fricative voicing contrast) and to Dutch as spoken in the Nijmegen region (where the fricative voicing contrast is neutralized). Children treated fricative voicing as lexically relevant only in a Limburgian-accented context. The authors conclude that toddlers who receive mixed evidence for a phonological contrast due to variation in accents in their input do not simply treat the contrast as allophonic, nor do they ignore the contrast. Rather, they appear to track two sets of statistics, one for each variant, as bilingual children have been argued to do (e.g., Sundara & Scutellaro, 2011). Studies showing that the presence of mixed distributional evidence for a lexical tone contrast does not lead to less specific lexical representations were

carried out by Singh et al. (2014) and Singh, Poh, and Fu (2016). Twelve- to 13-month-old Mandarin-English bilinguals who, like our Limburgian participants, received mixed evidence for the lexical relevance of pitch, noticed tonal MPs in a Mandarin version of the one-object Switch-task, but not in a non-tonal English version (Singh et al., 2016). In a preferential looking paradigm, also 18- and 24-month-old Mandarin-English bilinguals were sensitive to tonal MPs (Singh et al., 2014). From these findings we can probably infer that our Limburgian participants' lenient treatment of tonal MPs was not the result of their exposure to non-tonal cognates in Dutch. It could however be the case that their long-term exposure to dialect-related variation leads to a more general relaxation of phonetic boundaries, leading to less well specified lexical representations (e.g., Durrant et al., 2015). To investigate if the latter explanation holds, future studies should test Limburgians' responses to a variety of tonal and segmental MPs of familiar words, similar to the Durrant et al. study.

The fact that Dutch toddlers responded to pitch variation in a word learning task is not in line with previous studies on the lexical encoding of tone in non-tone language children (e.g., Hay et al., 2015; Quam & Swingley, 2010; Singh et al., 2014). These studies have shown that, from some point in development, English toddlers ignore pitch information during word learning. However, comparisons to these prior studies are difficult because these studies did not directly compare performance of tone and non-tone language learning children (i.e., in one statistical analysis). Moreover, prior studies testing non-tone language children have been restricted to learners of English, making it impossible to generalize their results to all non-tone language learners. We want to put forward three explanations for Dutch toddlers' sensitivity to word-level pitch.

First, Dutch toddlers could have interpreted the Limburgian pitch patterns as post-lexical intonation, as has already been put forward as an explanation for successful lexical tone discrimination in Dutch 6- to 12-month-old infants in Ramachers et al. (2017; Chapter 3 of this thesis). More

specifically, toddlers might over-assign weight to post-lexical factors in novel word learning tasks by virtue of having observed their communicative significance at other levels of linguistic structure (e.g., Hay et al., 2015; Singh et al., 2014). Similarly, Braun et al. (2014) proposed that extensive utterance-level prosody in the L1 is helpful for storing pitch information as part of novel mental representations. On the other hand, Frota, Butler, Correia, Severino, and Vigário (2012) showed that, by age 3, European Portuguese children do notice stress changes, but no longer treat intonation changes in newly learned words as lexically relevant.

A second possible explanation for the behavior of the Dutch toddlers also relates to L1 intonation. In a word recognition study, Fikkert and Chen (2011) showed that Dutch 24-month-olds have knowledge of appropriate native intonation patterns. Particularly in imperatives, Dutch toddlers strongly preferred a high-low pitch pattern combined with a strong-weak (trochaic) stress pattern. In our study, the target sentences in the test trials were always imperatives. Supposing that our Dutch toddlers were familiar with the intonation pattern of imperatives, their behavior could have been influenced by their expectations of what a well-formed imperative sounds like. An imperative that ends in a high-low pitch pattern (i.e., accent 1) could be preferred over an imperative ending in a low-high pitch pattern (i.e., accent 2). This would result in Dutch children structurally fixating the target *less* if pronounced with accent 2, *regardless* of the trained tone. However, in this case we should have found an interaction involving our variables Language and Tone. Since we attested no such interaction, our data provide no evidence for the suggestion that Dutch children's expectations regarding well-formed imperatives have influenced their behavior in our study.

The third explanation for the fact that Dutch toddlers noticed a pitch change in a novel word is that they might have perceived the Limburgian accent contrast as a quantity contrast rather than as a pitch contrast. Previous research has shown that the shape of a pitch pattern can indeed affect the perceived duration of the tone-bearing vowel (e.g., Gussenhoven & Zhou, 2013;

Lehiste, 1976; Pisoni, 1976; Yu, 2010). Despite the fact that the Limburgian tones' primary acoustic cue is pitch rather than duration, we think it is possible that speakers of Dutch perceived the pitch difference as a difference in duration. Previous research has shown that native and non-native speakers may give different degrees of attention to acoustic cues under the influence of the different functions and/or distributions of these cues in the L1 (Cebrian, 2006; Gandour & Harshman, 1978). For example, Gandour and Harshman showed cross-linguistic differences in the importance attributed to duration as a cue for tone perception, presumably reflecting the different linguistic status of vowel duration in their participants' L1s.

In Dutch, duration is one of the cues to word stress. The fact that our 2,5- to 4-year-old Dutch toddlers behaved differently from previously tested English peers (e.g., Quam & Swingley, 2010) could mean that Dutch listeners draw upon native prosodic cues to a greater extent than speakers of English. This is in line with Cooper et al. (2002), who have shown that Dutch listeners appear to process suprasegmental information more effectively than speakers of English. Moreover, in Dutch, duration is not only a cue to stress. As opposed to English, in Dutch vowel duration is lexically contrastive. Already at 18 months, Dutch children interpret vowel duration contrastively whereas Canadian English learners do not (e.g., Dietrich, 2006; Dietrich et al., 2007). Due to the use of duration as an acoustic cue to lexical contrast (i.e., word stress and vowel quantity) and Dutch children's early sensitivity to these contrasts (Dietrich et al., 2007; de Bree et al., 2008), we propose that the Dutch children in our study could have drawn upon their knowledge of this cue when perceiving a non-native tone contrast.

Anecdotal evidence with adult speakers of Dutch seems to strengthen this claim. Naïve speakers of Dutch who imitate the Limburgian tones tend to lengthen the stressed syllable of accent 2 words relative to accent 1 words. The impression that the citation form of accent 2 is longer in duration than the respective accent 1 form could be due to the more complex pitch pattern of

accent 2 (H*LH) compared to accent 1 (H*L), assuming that changes in f_0 can go hand in hand with a perceptual increase in duration (e.g., Rietveld & Gussenhoven, 1987; Lehiste, 1976; but cf. Gussenhoven & Zhou, 2013). In fact, Heijmans (2003) reports a formerly tonal dialect just outside the East-Limburgian area in which the tonal contrast was in large part reinterpreted as a length contrast. In future research, Dutch listeners could be presented tonal minimal pairs and explicitly judge which one sounds longer (e.g., Lehiste, 1976).

Until now, we have assumed different explanations for the behavior of the Limburgian and Dutch toddlers, despite their behavior being comparable. Lastly, we would like to mention the possibility that their behavior can be based on the same explanation. Recall that the only prosodic difference between Limburgian and Dutch is the fact that pitch is lexically relevant in Limburgian. Both languages make use of vowel duration, word stress and intonation. We therefore cannot exclude the possibility that the Limburgians might not have perceived the difference between accent 1 and accent 2 as a pitch contrast, but as an intonation or duration contrast.

Another finding that deserves some attention, especially in light of ongoing typological discussions about the phonological status of the Limburgian word prosodic contrast (e.g., Köhnlein, 2016, and references therein), is that Limburgian children were sensitive to MPs of both accent 1 and accent 2. Gussenhoven and Peters (2008) assume that accent 2 is the lexically specified tone, but our data provide no evidence for a perceptual asymmetry due to lexical (under)specification of one of the accents. It is possible that we did not attest an asymmetry due to a lack of power. However, an inspection of the means did not reveal a trend towards such an asymmetry. More research is needed to draw conclusions on this matter.

In the next section, we will turn to the findings from Experiment 2 with Limburgian and Dutch adults.

5.4.2 The lexical encoding of pitch in Limburgian and Dutch adults

In line with Quam and Swingley (2010), who used a very similar design, the Limburgian and Dutch adults in our study successfully learned novel word-object pairings. However, both groups showed very strong naming effects in both CP and MP trials, possibly masking effects of Condition and/or Language. The ceiling effect could either mean that the task was not sensitive enough (but cf. Quam & Swingley), or that our participants did not notice a pitch change within a word, or both.

Besides the pitch change condition, Quam and Swingley (2010) also included a vowel MP condition. In this condition, English participants exhibited a marginally significant negative naming effect, whereas they showed a significant positive naming effect in both the pitch MP and in the CP condition. Their effect of Condition thus rested on the significant negative naming effect induced by the vowel MP. They found no significant difference between the performance in pitch MP and CP conditions, which is in line with the behavior of our participants. In a future study, it would be valuable to include one or more segmental MP conditions in addition to a tonal MP condition (e.g., Quam & Swingley, 2010; Singh et al., 2014; Singh et al., 2015).

With respect to our Limburgian participants, it could be that lexical tone in Limburgian, relative to segments, does not share equal priority as a cue to word recognition. A similar claim has been made for Japanese (e.g., Goss, 2015). Since adult Limburgians have accumulated ample linguistic experience, they might have learned not to rely heavily on pitch during online language comprehension because of the relatively low functional load of lexical pitch and/or the amount of variability in their input. However, in light of Braun et al.'s (2014) finding, showing that adult speakers of German were very sensitive to Mandarin tone contrasts in a word learning paradigm, we strongly believe that the absence of effects in our study is due to task effects. To increase the demands on memory load, in a future study one could use disyllabic stimuli

and/or teach participants multiple tonal minimal pairs simultaneously (e.g., Braun et al., 2014).

Due to the lack of effects of Language, Condition or Tone in the adult study, we cannot draw conclusions on the phonological status of the Limburgian tone contrast. A lexical accent correctness judgment task (e.g., Goss & Tamaoka, 2015) or a lexical decision task with either phonological priming (e.g., Cutler & Otake, 1999) or semantic priming with tonal MPs could potentially advance our understanding of the lexical status of the Limburgian word prosodic contrast.

One important limitation that we want to mention at this point pertains to the input that both child and adult Limburgian participants were exposed to during the learning phase of the current experiment. Recall that they were presented multiple tokens of the target word, but that the prosodic context was held constant. That is, participants did not have to deal with the *allotonic* variation that they are usually confronted with. It would be interesting to see how Limburgian toddlers and adults would perform if this surface variation were included in the learning phase.

5.5 Conclusion

In conclusion, both Limburgian and Dutch 2,5- to 4-year-old children were sensitive to lexical pitch information in novel words. This indicates that they store pitch information as part of their novel lexical entries. Due to a lack of effects in our adult study, we cannot draw conclusions on the lexical encoding of pitch in Limburgian and Dutch adults. Since pitch is not contrastive at the word-level in Dutch, Dutch listeners should recognize words irrespective of their pitch pattern. Dutch toddlers' sensitivity to word-level pitch probably reflects their growing knowledge of the native prosodic system. They could either have perceived the different pitch patterns in terms of intonation (e.g., Singh et al., 2014), or in terms of vowel duration. The Limburgian toddlers' behavior was in line with our expectations since pitch is assumed to be part of Limburgian lexical representations. The fact that a pitch change only hindered

word recognition to a minor extent could be due to the specific input conditions that Limburgians are exposed to. Future studies should include speakers of similar restricted tone systems, for example Swedish, to corroborate that functional load and phonetic variability have an impact on lexical tone processing.

References

- Altwater-Mackensen, N., van der Feest, S., & Fikkert, P. (2013). Asymmetries in early word recognition: The case of stops and fricatives. *Language Learning and Development, 10*(2), 149-178. doi:10.1080/15475441.2013.808954.
- Bakker, F., & van Hout, R. (2012). De afgrenzing tussen de noordelijke en zuidelijke dialecten in Noord-Limburg. *Taal en Tongval, 64*(2), 159-193.
- Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program].
- Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics, 39*, 514-525.
- Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Language Research, 30*, 323-350.
- Braun, B. & E. Johnson. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics, 39*, 585-594.
- de Bree, E., van Alphen, P., Fikkert, P., & Wijnen, F. (2008). Metrical stress in comprehension and production of Dutch children at risk for dyslexia. In Chan, H., Jacob, H., & Kipia, E. (Eds.), *Proceedings of the 32nd annual Boston University Conference on Language Development (BUCLD 32)* (pp. 60-71).
- Broselow, E., Hurtig, R., & Ringen, C. (1987). The perception of second language prosody. In G. Loup & S. Weinberger (Eds.), *Interlanguage Phonology. The Acquisition of a Second Language Sound System* (pp. 350-364). New York: Newbury House.

- Buckler, H., & Fikkert, P. (2016). Dutch and German 3-year-old's representations of voicing alternations. *Language and Speech*, 59(2), 236-265.
- Burnham, D., & Brooker, R. (2002). Absolute pitch and lexical tones: tone perception by non-musician, musician, and absolute pitch non-tonal language speakers. *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2002)* (pp. 257-260).
- Burnham, D., Brooker, R., & Reid, A. (2015). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music*, 43(6), 881-897. DOI: 10.1177/0305735614546359.
- Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, 34, 372-387.
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, 25(5), 426-439.
- Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, 45(3), 207-228.
- Cornips, L. (2014). Socio-syntax and variation in acquisition: problematizing monolingual and bidialectal acquisition. In K.K. Grohmann (Ed.), *Three factors and beyond: Socio-syntax and language acquisition*. Special issue of *Linguistic Variation* 14(1), 1-25.
- COST Action IS0804 (2011). *Questionnaire for Parents of Bilingual Children (PaBiQ)*. <http://www.bi-sli.org>
- Cutler, A. (1986). *Forbear* is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29(3), 201-220.

210 Setting the Tone

- Cutler, A. & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, *105*(3), 1877-1888.
- Demuth, K. (1995). The acquisition of tonal systems. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 111-134). Hillsdale: Lawrence Erlbaum.
- Diesendruck., G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630-641.
- Dietrich, C. (2006). *The acquisition of phonological structure: Distinguishing contrastive information from non-contrastive variation*. Doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen.
- Dietrich, C., Swingley, D., & Werker, J.F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Science in the USA*, *104*, 16027-16031.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford University Press.
- Durrant, S. (2014). *The influence of long-term exposure to dialect variation on representation specificity and word learning in toddlers*. Doctoral dissertation, Plymouth University.
- Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, *42*, 447-465.
- van der Feest, S., & Johnson, E. (2016). Input-driven differences in toddlers' perception of a disappearing phonological contrast. *Language Acquisition*, *23*(2), 89-111. doi:10.1080/10489223.2015.1047096.
- Fennell, C. (2012). Object familiarity enhances infants' use of phonetic detail in novel words. *Infancy*, *17*(3), 339-353.

- Fennell, C., & Byers-Heinlein, K. (2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *International Journal of Behavioral Development*, 38, 309-316.
- Fennell C.T., Tsui, A.S., & Hudon, T.M. (2016). Speech perception in simultaneously bilingual infants. In E. Nicoladis & S. Montanari (Eds.) *Bilingualism across the Lifespan* (pp. 43-62). De Gruyter Mouton.
- Fikkert, P. & Chen, A. (2011). The Role of Word-Stress and Intonation in Word Recognition in Dutch 14- and 24-Month-Olds. *Proceedings of the 35th annual Boston University Conference on Language Development*.
- Francis, A., & Ciocca, V. (2003). Stimulus presentation order and the perception of lexical tones in Cantonese. *The Journal of the Acoustical Society of America*, 114(3), 1611-1621.
- Frota, S., Butler, J., Correia, S., Severino, C., & Vigário, M. (2012). Pitch first and stress next? Prosodic effects on word learning in an intonation language. *Proceedings of the 3rd annual Boston University Conference on Language Development* (pp. 190-201).
- Gandour, J., & Harshman, R. (1978). Cross-language differences in tone perception: a multidimensional scaling investigation. *Language and Speech*, 21, 1-33.
- Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M. & Gordon, L. (1987): The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23-45.
- Goss, S. (2015). *The effects of internal and experience-based factors on the perception of lexical pitch accent by native and nonnative Japanese listeners*. Dissertation, Ohio State University.

- Goss, S., & Tamaoka, K. (2015). Predicting lexical accent perception in native Japanese speakers: An investigation of acoustic pitch sensitivity and working memory. *Japanese Psychological Research*, 57(2), 143-154.
- Gussenhoven, C. (2000). On the origin and development of the Central Franconian tone contrast. In A. Lahiri (Ed.), *Analogy, Leveling, Markedness. Principles of Change in Phonology and Morphology* (pp. 215-260).
- Gussenhoven, C. (2004). Tone in Germanic: Comparing Limburgian with Swedish. In: G. Fant, H. Fujisaki, J. Cao, & Y. Xu (Eds.), *From Traditional Phonology to Modern Speech Processing* (pp. 129-136). Beijing: Foreign Language Teaching and Research Press.
- Gussenhoven, C., & Peters, J. (2008). De tonen van het Limburgs. *Nederlandse Taalkunde*, 13(1), 87-114.
- Gussenhoven, C. & van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.
- Gussenhoven, C., & Zhou, W. (2013). Revisiting pitch slope and height effects on perceived duration. *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013)* (pp. 1365-1369).
- Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87, B23-B34.
- Hallé, P. Chang, Y.-C., & Best, C. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- 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(1), 10-22.

- Heijmans, L. (2003). The relationship between tone and vowel length in two neighboring Dutch Limburgian dialects. In P. Fikkert & H. Jakobs (Eds.), *Development in Prosodic Systems* (pp. 7-46). Berlin/New York: Mouton de Gruyter.
- Hilton, M., & Westermann, G. (2016). The effect of shyness on children's formation and retention of novel word-object mappings. *Journal of Child Language*, doi:10.1017/S030500091600057X.
- Hyman, L. (2001). Tone systems. In M. Haspelmath, E. König, W. Östereicher & W. Raible (Eds.), *Language Typology and Language Universals: An International Handbook*. (2 vols.) Berlin: Walter de Gruyter
- Hyman, L.M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, 31, 213-238.
- Jarvis, L.H., Merriman, W.E., Barnett, M., Hanba, J., & van Haitsma, K.S. (2004). Input that contradicts young children's strategy for mapping novel words affects their phonological and semantic interpretation of other novel words. *Journal of Speech, Language and Hearing Research*, 47, 392-406.
- Jusczyk, P. (1997). *The discovery of spoken language*. MIT Press.
- Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition*, 18(4), 626-638.
- Köhnlein, B. (2016). Contrastive foot structure in Franconian tone-accent dialects. *Phonology*, 33, 87-123.
- Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.
- Kubozono, H. (1993). *The organization of Japanese prosody*. Tokyo: Kurosio.

214 Setting the Tone

- Kuhl, P. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews/ Neuroscience*, 5, 831-843.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, N. Warner, & T. Rietveld (Eds.), *Laboratory Phonology 7* (pp. 637-676). Berlin: Mouton.
- Lehiste, I. (1976). Influence of fundamental frequency pattern on the perception of duration. *Journal of Phonetics*, 4(2), 113-117.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133, 385-394.
- Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear Hear*, 19(1), 1-36.
- Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, 38(3), 606-627.
- Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14, 57-77.
- Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain meanings of words. *Cognitive Psychology*, 20, 121-157.
- Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: Comparing monolingual and bilingual infants. *Developmental Science*, 13, 229-243.
- Meints, K. & Woodford, A. (2008). *Lincoln Infant Lab Package 1.0: A new programme package for IPL, Preferential Listening, Habituation and Eyetracking*. [WWW document: computer software & manual]. <http://www.lincoln.ac.uk/psychology/babylab.htm>.

- Ota, M. (2003). The development of lexical pitch accent systems: an autosegmental analysis. *Canadian Journal of Linguistics*, 48(3/4), 357-383.
- Paradis, J. (2011). Individual Differences in Child English Second Language Acquisition: Comparing Child-Internal and Child-External Factors. *Linguistic Approaches to Bilingualism*, 1(3), 213-237.
- Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English Language Learners: Using Parent Report on First Language Development. *Journal of Communication Disorders*, 43, 474-497.
- Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Linguistic inquiry monograph 15. Cambridge: MIT Press.
- Pisoni, D. (1976). Fundamental frequency and perceived vowel duration. *Journal of the Acoustical Society of America*, 59. doi:10.1121/1.2002669.
- Quam, C., & Swingley, 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.
- Ramachers, S., Brouwer, S., & Fikkert, P. (2017). 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.
- Ramon-Casas, M., & Bosch, L. (2010). Are non-cognate words phonologically better specified than cognates in the early lexicon of bilingual children? In M. Ortega-Llebaria (Ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 31-36). Somerville, MA: Cascadilla.
- Riad, T. (2013). *The phonology of Swedish*. Oxford: Oxford University Press.

- Rietveld, A.C.M., & Gussenhoven, C. (1987). Perceived speech rate and intonation. *Journal of Phonetics*, *15*, 273-285.
- Rost, G.C., & McMurray, B. (2010). Finding the signal by adding noise: the role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608-635.
- Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*, 309-320.
- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, *5*(4), 489-522.
- Schlichting, J., & Spelberg, H.L. (2002). *Lexilijst Nederlands, een instrument om de taalontwikkeling te onderzoeken bij Nederlandstalige kinderen van 15-27 maanden in het kader van vroegtijdige onderkenning*. Manual. Lisse: Swets Test Publishers.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, *12*, 874-887.
- Sekiguchi, T., & Nakajima, Y. (1999). The use of lexical prosody for lexical access of the Japanese language. *Journal of Psycholinguistic Research*, *28*(4), 439-454.
- Shi, R., Gao, J., Achim, A., & Li, A. (2017). Perception and representation of lexical tones in native Mandarin-learning infants and toddlers. *Frontiers in Psychology*. doi: 10.3389/fpsyg.2017.01117.
- Singh, S., & Fu, C. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, *87*(3), 834-854. doi:10.1111/cdev.12512.

- 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.
- Singh, L., Hui, T.J., Chan, C., & Golinkoff, R. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, 17(1), 94-109.
- Singh, L., Poh, F.L.S., & Fu, C.S.L. (2016). Limits on monolingualism? A comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in Psychology*, 7(667). doi:10.3389/fpsyg.2016.00667.
- Singh L., & Quam, C. (2016). Can bilingual children turn one language off? Evidence from perceptual switching. *Journal of Experimental Child Psychology*, 147, 111-125.
- So, C., & Best, C. (2008). Do English speakers assimilate Mandarin tones to English prosodic categories? In *Proceedings of the 9th annual conference of the International Speech Communication Association (Interspeech 2008), incorporating the 12th Australasian International Conference on Speech Science and Technology (SST 2008)* (p. 1120). Baixas, France: International Speech Communication Association.
- So, C., & Best, C. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, 53(2), 273-293.
- So, C., & Best, C. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36, 195-221.

- Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381-382.
- Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, *39*(4), 505-513.
- Swingley, D. (2009). Onsets and codas in 1,5-year-olds' word recognition. *Journal of Memory and Language*, *60*, 252-269.
- Swingley, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.
- Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484.
- Swingley, D., & Aslin, R.N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99-132.
- Swingley, D., Pinto, J.P., & Fernald, A. (1998). Assessing the speed and accuracy of word recognition in infants. In C. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (pp. 257-277). Stamford, CT.
- Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R. (2014). Is pitch accent necessary for comprehension by Japanese speakers? – An ERP investigation. *Journal of Neurolinguistics*, *27*, 31-40.
- Tong, Y., Francis, A., & Gandour, J. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, *23*(5), 689-708.
- Tsuji, S., Fikkert, P., Yamane, N., & Mazuka, R. (2016). Language-general biases and language-specific experience contribute to phonological detail in toddlers' word representations. *Developmental Psychology*, *52*(3), 379-390.

- Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic Segment Durations in Prosodic Research: A Practical Guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter & J. Schließer (Eds.), *Methods in Empirical Prosody Research* (pp. 1-27). Berlin: De Gruyter.
- Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*, 4(1), 1-22.
- Wang, Y., Behne, D., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, 25, 449-466.
- Wetterlin, A. (2007). *The Lexical Specification of Norwegian Tonal Word Accents*. Doctoral dissertation, University of Konstanz.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. *Proceedings of the Fifth International Conference on Language Resources and Evaluation LREC*. Genoa: Italy.
- White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59, 114-132.
- Wu, Z. (2000). From traditional Chinese phonology to modern speech technology: Realization of tone and intonation in spoken Chinese. *Proceedings of the International Conference on Spoken Language Processing (ICSLP 2000)*, Vol. 1, B1-B12.
- Wu, X., Tu, J.-Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623-641.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

220 Setting the Tone

Yu, A.C.L. (2010). Tonal effects on perceived vowel duration. In C. Fougeron, B. Kuehnert, M. Imperio, & N. Vallee (Eds.), *Laboratory Phonology 10* (pp. 151-168). Berlin/New York: Walter de Gruyter GmbH.

Appendix

Table 1. Number and percentage of excluded trials, blocks and participants per language and age group for the child study.

	Dutch	Limburgian
<i>Trial exclusion n (%)</i>		
1: <500 ms LT post-naming	12 (7,5)	35 (21,3)
2: PTL pre-naming = 1	22 (13,8)	23 (14,0)
3: equip./exp. error	1 (0,6)	2 (1,2)
4: refusal to participate	4 (2,5)	14 (8,5)
Total excluded n (%)	38 (23,8)	74 (45,1)
<i>Total n trials</i>	160	164
<i>Block exclusion n (%)</i>		
1: not enough test trials	4 (5)	15 (18,3)
2: <20 s LT learning	-	4 (4,9)
Total excluded n (%)	4 (5)	19 (23,2)
<i>Total n blocks</i>	80	82
<i>Participant exclusion n (%)</i>		
1: 1 or 2 blocks excluded	4 (10)	18 (43,9)
2: equip./exp. error	-	-
3: other	1 (2,5)	-
Total excluded n (%)	5 (12,5)	18 (43,9)
<i>Total n participants</i>	40	41

222 Setting the Tone

Table 2. Means, standard deviations and ranges of proportions of input quantity and quality for the Limburgian children (missing $N = 1$).

PaBiQ measures	Mean (SD)	Range
Input quantity Limburgian	0.70 (0.24)	0.15 – 1
Input quantity Dutch	0.40 (0.22)	0.02 – 0.72
Input quality Limburgian	0.49 (0.11)	0.31 – 0.69
Input quality Dutch	0.39 (0.13)	0.19 – 0.69

Table 3. Acoustic measurements of the target stimuli.

Measures	Accent 1 (n=24)	Accent 2 (n=24)	p-value
Min F0 TBP* (Hz)	168.5 (8.1)	209.3 (22.4)	<.001
Max F0 TBP (Hz)	402.7 (32.1)	380.5 (20.0)	.007
Mean F0 TBP (Hz)	294.4 (38.2)	296.5 (22.7)	>.1
F0 range TBP (Hz)	234.2 (32.4)	171.2 (17.0)	<.001
Duration TBP (s)	.38 (.08)	.37 (.06)	>.1
Duration token (s)	.55 (.03)	.57 (.04)	>.05

*TBP stands for Tone Bearing Portion. Note that the TBP for *moon* stimuli consisted of the entire rhyme (i.e., [o:n]) whereas for *taaf* stimuli it consisted of the nucleus (i.e., [a:]).

224 Setting the Tone

6. General Discussion

This chapter summarizes and discusses the main findings of this thesis. After a summary of the main findings, the broader implications of this study are discussed in thematic blocks addressing both theoretical and methodological issues. Finally, suggestions for future research are given.

6.1 Summary

By studying the acquisition and processing of lexical tone in Limburgian, this thesis set out to widen our understanding of the influence of the functional load of tone and the phonetic variability in the realization of lexical tones on the acquisition and processing of a lexical tone system. Due to the opacity and the relatively low functional load of the Limburgian tone system, acquiring this system might be quite a challenge. By means of a series of behavioural experiments, we addressed how infants and children growing up with Limburgian²³ as well as adult speakers of Limburgian perceive and encode the tonal make-up of words, and compared them to control groups of speakers of Dutch, a language without lexical tone.

Chapter 3 reported on an experiment testing discrimination of the Limburgian lexical tones in 6-to-12-month-old Limburgian and Dutch infants using a hybrid visual habituation procedure (Houston, Horn, Qi, Ting, & Zhao, 2007). This study aimed to answer two questions:

1a. How does Limburgian infants' lexical tone perception develop in the first year of life?

²³ The term Limburgian is an umbrella-term for many different dialects. In this thesis, the focus is on East-Limburgian dialects. We refer the reader to Chapter 2 for more details on the linguistic situation in Limburg.

1b. Does Limburgian infants' perception of lexical tones develop differently from that of infants acquiring non-tonal Standard Dutch?

Based on previous studies on native tone and pitch-accent perception in the first year of life (Sato, Sogabe, & Mazuka, 2010a; Yeung, Chen, & Werker, 2013), we expected that 6-to-12-month-old Limburgian infants would successfully discriminate the Limburgian tones throughout their first year of life. For the Dutch infants, two outcomes were suggested. On the one hand, they could show an age-related performance decline (like the non-tone language listeners in Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; and Yeung et al., 2013). On the other hand, they could maintain their discrimination of the Limburgian tones, just like the Dutch infants in Liu and Kager (2014) and Chen and Kager (2016), who showed persistent sensitivity to certain (salient) Chinese tone contrasts.

Our results showed that both Limburgian and Dutch infants were sensitive to the Limburgian tonal contrast throughout their first year of life. We therefore could not conclude that Limburgian and Dutch infants attended differently to Limburgian lexical tones. The Dutch infants could have perceived the Limburgian tones as native intonation patterns (i.e., statement vs. question).

A follow-up experiment with Limburgian and Dutch adults was carried out to explore whether differences in lexical pitch processing between Limburgian and Dutch listeners appear at a later age. **Chapter 4** addressed the following research questions:

- 2a. Are Limburgian adults sensitive to word-level pitch?**
- 2b. Do Limburgian and Dutch adults differ in their sensitivity to word-level pitch?**
- 2c. Is Dutch listener's perception of Limburgian lexical tones influenced by their native intonation system?**

To address these questions, we ran a categorical AXB discrimination task with trials including *between*-category variation and trials with *within*-

category variation. Trials with between-category variation tested lexical tone discrimination (e.g., *moon*_{1a} – *moon*_{1b} – *moon*_{2a}; subscripts indicate tones and tokens). Within-category variation was added to test whether lexical pitch perception is categorical in Limburgian listeners (e.g., *moon*_{1a} – *moon*_{1b} – *moon*_{1b}). Limburgian participants were expected to outperform Dutch participants overall because of their experience with word-level pitch patterns. Moreover, if Limburgian listeners' perception is indeed driven by native lexical tone categories, Limburgian listeners were expected to do worse in perceiving within-category than between-category pitch differences, because they presumably have learned to abstract away from phonetic within-category variation (e.g., Hallé, Chang, & Best, 2004; Hoffmann, Sadakata, Chen, Desain, & McQueen, 2014; Qin & Mok, 2011; Xu, Gandour, & Francis, 2006). Dutch listeners' perception was not expected to be categorical, as they have no lexical tone categories to draw upon. They were expected to perform equally well (or badly) in between- and within-category variation trials.

To see whether Dutch adults' perception of Limburgian tones is influenced by native prosodic categories, we manipulated the position of the tones in our pseudo-word stimuli (e.g., Braun & Johnson, 2011). Half of the trials featured monosyllabic pseudo-words (*taag* and *moon*) where the tonal contrast is realized in utterance final position. In this position, the Dutch listeners were expected to interpret the tones as instances of native post-lexical intonation categories and show good discrimination. The citation forms of the Limburgian tones might be interpreted as Dutch statement and question intonation patterns. The other half of the trials consisted of disyllabic pseudo-words with trochaic stress (*keeve* and *perger*). In these pseudo-words, the tones were realized on the penultimate syllable and thus occurred in non-final position. Crucially, in this position, the tones supposedly do not signal a linguistically meaningful contrast for speakers of Dutch. As a consequence, we expected Dutch listeners to show poorer discrimination performance in these trials as compared to monosyllabic trials.

Overall, Limburgian listeners had higher accuracy scores than Dutch listeners. From this we conclude that speakers of Limburgian attend more closely to word-level pitch than speakers of Dutch as a result of the lexical distinctiveness of pitch variation in Limburgian. Moreover, Limburgian listeners performed equally well on discriminating between- and within-category pitch variation. Their high sensitivity to both between- and within-category pitch variation might be the result of their ample exposure to meaningful pitch variation due to the intricate tone-intonation interaction in Limburgian, a feature that sets Limburgian apart from typically studied tone languages.

Nevertheless, Dutch listeners performed above chance, indicating that they showed some sensitivity to Limburgian tones, in line with the Dutch infants' behavior in Chapter 3. However, we cannot conclude that this sensitivity was guided by their knowledge of Dutch intonation. Contrary to our expectations, Dutch listeners showed better discrimination in disyllables (pitch in non-final position) than in monosyllables (pitch in final position). On the basis of an acoustic analysis, we proposed that the success in the disyllables could be attributed to the fact that, in disyllables, the difference between accents 1 and 2 was more prominent (i.e., acoustically more salient).

Possibly, Dutch participants did not perceive the difference between stimuli as different native intonation patterns (e.g., 'this sounds like a question and this like a statement'), but rather as differing in their degree of foreign-accentedness (e.g. 'this sounds Limburgian and this sounds Dutch').

Despite the different functions of pitch in Limburgian and Dutch, both Limburgian and Dutch infants (Chapter 3) as well as adults (Chapter 4) showed persistent sensitivity to Limburgian lexical tones. After Chapters 3 and 4, the question thus no longer was whether Limburgian and Dutch listeners can discriminate lexical tones, but whether they interpret them as lexically meaningful.

In **Chapter 5**, we aimed to answer the following questions:

- 3a. Do child and adult speakers of Limburgian encode pitch information as part of their lexical entries when learning novel words?**
- 3b. Do child and adult speakers of Limburgian behave differently from Dutch peers with respect to their sensitivity to pitch in a word learning context?**

To answer these questions, Limburgian and Dutch 2,5- to 4-year-olds as well as adults participated in a bimodal preferential looking experiment (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Following the procedure employed by Quam and Swingley (2010) and Singh, Hui, Chan, and Golinkoff (2014), participants learned two novel word-object mappings. After training, word recognition (measured as an increase in the proportion of target fixation from pre- to post-naming phase) was tested in two conditions, involving correct pronunciation (CP) trials and mispronunciation (MP) trials featuring a pitch change.

Based on earlier studies with tone language listeners (Singh et al., 2014; Singh, Goh, & Wewalaarachchi, 2015), we expected Limburgians to be sensitive to MPs involving pitch under the assumption that their novel lexical representations include tonal information. However, no prior word learning and recognition studies investigated tone languages, whether or not ‘restricted’, that yield a considerable amount of surface variation in the realization of lexical tones. Additionally, the bidialectal status of our Limburgian participants contributed to the variability in their input. Therefore, we also raised the possibility that Limburgians might be less sensitive to tonal MPs than the tone language speakers in prior studies.

Because pitch is hardly or not used to signal lexical distinctiveness in Dutch, we expected speakers of Dutch not to encode pitch information when building novel lexical representations due to the lack of experience with storing

long-term memory representations of pitch (e.g., Hay, Graf Estes, Wang, & Saffran, 2015; Quam & Swingley, 2010; Singh et al., 2014).

Our results showed that both Limburgian and Dutch children fixated the target less in MP trials than in CP trials, suggesting that both groups encoded pitch lexically. The Limburgian toddlers' sensitivity to word-level pitch was in line with our expectations, but the Dutch toddlers' sensitivity was not. Dutch toddlers may have drawn upon their knowledge of native prosody (e.g., Singh et al., 2014).

Nevertheless, in MP trials, target looking also increased significantly from the pre- to the post-naming phase for both Limburgian and Dutch toddlers, meaning that word recognition was still possible despite the pitch change. This is not in line with previous studies testing tone language learners, who found no naming effects in tonal MP conditions (Singh et al., 2014; Singh et al., 2015). The Limburgians may have less stable representations of pitch due to their specific input conditions.

In line with Quam and Swingley (2010), the Limburgian and Dutch adults in our study successfully learned novel word-object pairings. However, both groups performed at ceiling in both CP and MP conditions, possibly masking any effects of condition and/or language. Due to their ceiling performance, we cannot draw conclusions about their interpretation of pitch changes. Our results thus did not support the hypothesis that Limburgian adults treat pitch differences within words differently from Dutch adults.

After a brief summary of the main results of this thesis, we now address the broader implications of the results.

6.2 Theoretical considerations

6.2.1 Perceptual reorganization for lexical tones

Experience with the ambient language and, at the same time, a lack of experience with non-native speech sounds is one of the factors guiding PERCEPTUAL REORGANIZATION, a process of perceptual attunement that leads

infants to perceive speech sound contrasts in accordance with their native phonology. In the course of the first year of life, perception of native contrasts has been shown to be maintained (e.g., Sato et al., 2010a) or even enhanced (e.g., Kuhl, Stevens, Hayashi, Deguchi, Kiritani, & Iverson, 2006), whereas perception of non-native contrasts has been shown to decline (e.g., Polka & Werker, 1994; Mattock et al., 2008; Tsushima et al., 1994; Werker & Tees, 1984). However, other studies have shown that discrimination of native and non-native contrasts is not only dependent on experience, but is also influenced by characteristics of the input other than frequency (e.g., Best, McRoberts, & Sithole, 1988; Chen & Kager, 2016; Liu & Kager, 2014; Mazuka, Hasegawa, & Tsuji, 2014; Narayan, Werker, & Beddor, 2010; Sato, Sogabe, & Mazuka, 2010b; Tsao, 2008, 2017). Some non-native tone contrasts remain perceivable throughout infancy (e.g., Chen & Kager, 2016; Liu & Kager, 2014) and even into adulthood (e.g., Hallé et al., 2004; So & Best, 2010).

Successful discrimination of non-native lexical tones by Dutch infants as reported in this thesis shows that PR is not only a process of perceptual *narrowing* based on input frequency (or linguistic experience). Infants' discrimination of native and non-native speech sound contrasts is rather a product of the interplay between language input, stimulus-specific characteristics (e.g., perceptual salience), and general perceptual maturation. Psycho-acoustic and experiential factors give way to multiple routes toward adult-like perception of speech sounds. An important experience-related factor that makes tones special is that pitch variation is used in all languages at the sentence level (e.g., Burnham & Mattock, 2007). As a consequence, every listener will have some amount of exposure to meaningful pitch variation. We proposed that the Dutch infants' success could be due to them perceiving the Limburgian tones as native Dutch intonation patterns (e.g., Best, 1994; Liu & Kager, 2014). Additionally or alternatively, the Dutch infants' sensitivity to Limburgian tones could be due to acoustic salience. In that case, probably any language learner should be successful in Limburgian tone discrimination. To

find out if the Limburgian tone contrast is acoustically salient, one could test discrimination of Limburgian tones in a group of language learners that make limited use of suprasegmental distinctions. What exactly determines the salience of a (tone) contrast, however, is another open question. For work that attempted to address this question for a different Limburgian dialect, see Fournier & Gussenhoven (2012).

Regarding the development of *native* tone perception in the first year of life, we have to consider the possibility that a more nuanced pattern of results, i.e., enhanced discrimination ability with age (e.g., Kuhl et al., 2006; Tsao, 2017), could have been masked due to the relatively small sample sizes within the Limburgian age groups. On the other hand, earlier work involving a restricted tone system did not show improvement in a behavioural discrimination task either (Sato et al., 2010a). Sato and colleagues found that 4- and 10-month-old Japanese infants showed equally successful discrimination of native pitch-accents. However, using Near Infrared Spectroscopy next to the behavioural paradigm, they showed that cerebral lateralization did change over time, changing from bilateral to left-dominant processing, indicating increased linguistic processing of pitch-accent. Future studies could use neurophysiological techniques to investigate whether Limburgian infants show an age-related difference with respect to cerebral lateralization.

6.2.2 Perceptual assimilation of lexical tones

The PERCEPTUAL ASSIMILATION MODEL FOR SUPRASEGMENTALS (e.g., So & Best, 2008, 2010, 2014) has been proposed to account for the repeated observation that non-tone language listeners are able to discriminate and categorize lexical tones (e.g., Braun & Johnson, 2011; Burnham et al., 2015; Francis, Ciocca, Ma, & Fenn, 2008; Hallé et al., 2004; Schaefer & Darcy, 2014; So & Best, 2010, 2014). On this account, listeners make use of native (e.g., post-lexical) prosodic categories when perceiving pitch patterns (e.g., lexical tones) in a foreign language. The degree of similarity between non-native and native categories

determines what type of assimilation takes place, which is in turn predictive of discrimination performance.

The Limburgian tone contrast in its citation form (i.e., in monosyllables) could be a case of two-category (TC) assimilation for Dutch listeners. TC assimilation occurs if two non-native categories (in this case accent 1, H*L, and accent 2, H*LH) are mapped onto two separate native categories (in this case statement intonation, i.e., H*L L%, and question intonation, i.e., H*L H%). If so, discrimination performance in monosyllables should be excellent.

However, Dutch listeners' discrimination performance in monosyllabic between-category variation trials did not imply a one-to-one mapping onto native intonation categories (i.e., TC assimilation), since TC assimilation normally leads to excellent discrimination (> .90). Moreover, against our expectations, Dutch listeners performed significantly better on disyllables than on monosyllables. We argued that this is probably due to stimulus-dependent acoustic characteristics. Discrimination of pre-final pitch movements by Dutch listeners has been found in previous work (Hoffman et al., 2014; but see Braun & Johnson, 2011). However, Hoffman et al.'s experiment, in contrast to ours, was a low memory demanding AX discrimination task that favored psychoacoustic processing. The performance of the Dutch listeners in their task was thus unlikely to be dependent on language-specific, phonological processing, and is therefore not directly comparable to our results.

One explanation that we would like to propose for the moderate performance of our Dutch listeners is based on suggestions made by Hallé et al. (2004) and Francis et al., (2008). According to Hallé and colleagues, among others, intonational categories might better be seen as a continuum of pitch variations instead of as strictly defined categories. If our Dutch participants have indeed called upon their native post-lexical intonation categories, their fairly poor performance could reflect that they were not able to unequivocally classify the perceived pitch patterns in our experiment as instances of some

finite set of contrastive categories (but see Braun, Kochanski, Grabe, & Rosner, 2006; Remijsen & van Heuven, 2003). The assimilation and discrimination predictions formulated in prior PAM studies were mostly applied to segmental contrasts, but they may not be directly transferable to suprasegmental contrasts. PAM's discrimination predictions might be overridden by the less categorical perception of intonation. So and Best (2014) also found rather weak categorization of Mandarin tones into common intonation categories by English and French speakers, which suggests a similar conclusion. More research is needed on the psychological reality and stability of post-lexical intonation categories.

Since we only administered a discrimination task to our participants, we can only speculate about possible mappings between Dutch and Limburgian. In order to determine the actual pattern of assimilation, it would be necessary to carry out a categorization/labeling task, possibly followed by a goodness-rating task (e.g., Best, McRoberts, & Goodell, 2001; Harnsberger, 2001; Højen & Flege, 2006; Polka, 1995; So & Best, 2008, 2014; Tyler, Best, Faber, & Levitt, 2014). In such a two-step task, participants would first be asked to categorize Limburgian tones into a set of predetermined native categories, for example 'statement', 'question', 'uncertainty', 'don't know' (e.g., So & Best, 2008). After choosing an answer category, participants rate the goodness of fit by judging how well the target tone matches the native category they have just chosen. Suppose both Limburgian tones would be categorized as 'statement', then the goodness of fit rating could make clear whether both tones are equally good or bad instances of the statement category (an instance of Single Category assimilation), or that one of the tones fits the native category better than the other (Category Goodness assimilation). These labeling test results can thus be used to classify a non-native contrast as one of the assimilation types proposed by PAM-S, which in turn are expected to correlate with a certain degree of discriminability.

Recall that from our results we could not tell if our Dutch participants proceeded through the experiment as a pitch-pattern discrimination task (e.g.,

statement vs. question) or as a regional-accent-discrimination task (e.g., Limburgian vs. Dutch). Probably two labeling tasks with two different groups of Dutch subjects should be run to see which group would show the most consistent labeling performance. Alternatively, a rating task in which Dutch listeners rate accent 1 and accent 2 stimuli on their degree of accentedness could be conducted.

A potential drawback of labeling tasks is that participants are forced to choose between pre-set labels (e.g., So & Best, 2008, 2014). The question remains whether their choices during such a task represent what happens during online non-native tone perception (for example in a discrimination task). More research investigating the relationship between labeling and discrimination results involving non-native prosodic categories could widen our understanding of how native intonation categories influence discrimination of non-native lexical tone categories.

Future studies should take into account that the possibility of perceiving a (prosodic) contrast in terms of regional accent variation might have consequences for testing the applicability of PAM(-S). Whereas all the aforementioned earlier studies used Chinese CV syllables to test tone identification and/or discrimination in non-tone language speakers (e.g., Broselow, Hurtig, & Ringen, 1987; Francis et al., 2008; Hallé et al., 2004; So & Best, 2008, 2010, 2014), our Dutch participants were presented with stimuli that sounded relatively more word-like and probably also more native-like. Because Limburgian and Dutch are closely related (in contrast to English/French and Chinese), we could create pseudowords that were equally compatible with both the languages at test. It is thus reasonable to assume that our Dutch participants perceived the stimuli as native pseudowords, an assumption that probably does not hold, or at least to a lesser extent, for the non-tone language listeners in prior studies.²⁴ The native-likeness of our

²⁴ For example, the English and French participants in So and Best (2010, 2014) were presented with Chinese syllables like /tʃi/, /tʃa/, and /tʃu/.

stimuli combined with our Dutch listeners' experience with regional accents might have led to a different listening strategy, and thus to different outcomes relative to prior studies. Future studies testing PAM(-S) should keep in mind that, possibly, listeners might not draw upon native categories if the contrast of interest can be perceived in terms of native regional variation, which is of course only possible if stimuli sound like native words.

6.2.3 Differences between tone languages and their influence on pitch processing

In this thesis we have shown that the performance of Limburgian listeners in a variety of tasks involving lexical tone perception was not always in line with previous studies testing tone-language listeners. We argue that these discrepancies are due to differences across tone languages. These differences pertain to the functional load (or 'usefulness') of tone and the phonetic variability in the realization of tones. It was largely unclear whether and how these differences influence the acquisition and processing of lexical tone. The following paragraphs discuss how our results contribute to answering this question.

In the discrimination task reported in Chapter 4, adult Limburgians showed equally successful discrimination of between- and within-category pitch variation. This was not in line with the frequently raised hypothesis that perception of within-category pitch differences should be more difficult than perception of between-category pitch differences because tone language listeners must ignore within-category variations to efficiently categorize f_0 contours into tonal categories (Hallé et al., 2004; Hoffmann et al., 2014; Qin & Mok, 2011; Xu et al., 2006). We argue that the Limburgians' high sensitivity to within-category pitch differences could be the result of their ample exposure to meaningful pitch variation due to the intricate tone-intonation interaction in Limburgian.

Another remarkable finding in Chapter 4 was the Limburgians' relatively low discrimination performance in between-category trials. Mandarin Chinese listeners in Xu et al. (2006) performed similar to our Limburgian listeners in an AX same-different task with Mandarin tones. However, in Hallé et al.'s (2004) non-categorical AXB two-step discrimination task, Mandarin listeners scored above 90% in trials featuring tone pairs at the category boundaries (i.e., between-category variation trials). In a categorical odd-one-out-paradigm, in which listeners heard three stimuli and had to indicate which one of them had a different tone, native Thai listeners in Wayland and Guion (2003) discriminated native Thai tones with 91% accuracy. Cantonese-English bilinguals in Tong, Lee, Lee, and Burnham (2015) discriminated Cantonese tones with an accuracy of 95% in a non-categorical AXB discrimination task. The first thing to note is that all these studies used slightly different methodologies, which could have invoked different listening strategies (i.e., more or less acoustic-phonetic or language-specific; for elaboration see section 6.3.1). Moreover, tone in Limburgian has a low functional load compared to tone in the languages tested in the aforementioned studies. The functional load of tone could thus also be an explanation for the Limburgians' relatively low discrimination performance.

In the word learning and recognition task reported in Chapter 5, we showed that Limburgians still preferred the target object over the distracter upon hearing a pitch change, meaning that word recognition was possible despite a tonal MP. Previous studies featuring Mandarin Chinese toddlers found no naming effects in tonal MP conditions (Singh et al., 2014; Singh et al., 2015). The discrepancy between surface forms and underlying tones in Limburgian could make it challenging for a learner of Limburgian to unravel the underlying tone system (e.g., Demuth, 1995; Ota, 2003; Rost & McMurray, 2010). As a result, tonal representations in Limburgian might be (temporarily) less stable.

One important limitation that we want to mention at this point pertains to the input that both child and adult Limburgian participants were

exposed to during the learning phase of our word learning experiment. Recall that they were presented multiple tokens of the target word, but that the prosodic context was held constant. That is, during the experiment, participants did not have to deal with surface variation due to the tone-intonation interaction that they are usually confronted with. It would be interesting to see how Limburgian toddlers and adults would perform if this surface variation were included in the learning phase.

Future studies should compare tone sensitivity in speakers of restricted tone systems, for example Swedish, to speakers of more dense tone systems like Mandarin Chinese to find out whether functional load and phonetic variability have an impact on category strength and, as a result, on discrimination and/or categorization and lexical encoding of tones. Teasing apart the relative influence of functional load and phonetic variability is probably difficult, since those tone languages that have smaller tonal inventories typically also yield more complex intonation systems.

6.2.4 The influence of variation due to bidialectalism

What complicates the matter is that most of our Limburgian participants had exposure to multiple regional varieties of one language. Being bidialectal, Limburgians have two rich prosodic systems at their disposal: Limburgian and Dutch. The Limburgians' experience with two linguistic systems that make extensive use of pitch might positively affect their perception of pitch in general and may have contributed to their sensitivity to within-category pitch differences in Chapter 4.

Another type of variation due to bidialectalism comes from the experience with multiple dialects and Limburgian-accented Dutch. Evidence for the influence of dialect- or regional accent-induced variation on lexical specificity is scarce. Durrant, Delle Luche, Cattani, and Floccia (2015) showed that variable phonological input as a result of dialect variation has an impact on the specificity of lexical representations in 20-month-old multidialectal

toddlers. They showed that multidialectal children, in contrast to monodialectals, were not able to detect a range of single feature MPs of familiar words.

Another potential influencing factor due to experience with multiple languages is the fact that Limburgian has many cognates in Dutch without a tonal specification. Thus, Limburgians receive mixed evidence for the lexical relevance of pitch. Possibly, this mixed evidence (temporarily) leads them to assign less weight to pitch as a lexically contrastive feature. However, earlier results point in another direction. Van der Feest and Johnson (2016) showed that 24-month-old Dutch toddlers who received mixed distributional evidence for a phonological contrast due to exposure to two regional accents of Dutch treated this contrast as lexically relevant if listening to a speaker maintaining the contrast. In another study, 12- to 13-month-old Mandarin-English bilinguals who received mixed evidence for the lexical relevance of pitch noticed tonal MPs in a Mandarin version of the one-object Switch task, but not in a non-tonal English version. The authors suggest that the mere presence of mixed evidence, or conflict, across the languages of a child may highlight relevant properties of each language (Singh, Poh, & Fu, 2016). From these latter studies we can probably conclude that mixed evidence for the lexical relevance of a contrast does not lead to less stable lexical representations.

From our study, it is not possible to tell apart the influence of the native tonal system on the one hand (i.e., low functional load and surface variation) and the influence of variability caused by bidialectalism on the other. The question thus remains what type(s) of variation has caused our Limburgian participants' results. Future research should control for whether participants are monolingual, bidialectal, or bilingual. A replication of our study with monolingual Swedish children could provide more insight into the effect of functional load and surface variation on developing tonal representations, as Swedish is also a restricted tone system showing substantial surface variation in the realization of word-level pitch contrasts (e.g., Bruce, 1977; Riad, 2013).

6.3 Methodological considerations

6.3.1 Comparability across studies

Throughout this thesis, it has been pointed out that the comparison of our results to results of prior studies was difficult. On the one hand, this was because our population of interest, speakers of Limburgian, represented a special case with respect to the linguistic input they are exposed to (i.e., the Limburgian tonal system and the variability due to bidialectalism). On the other hand, we also encountered numerous cases of methodological differences across studies. We want to highlight some of these differences in this section.

The AXB task reported in Chapter 4 was designed to invoke language specific (i.e. phonological) listening rather than psycho-acoustic (i.e., phonetic) perception by including relatively long ISI's (> 1000 ms; e.g., Harnsberger, 2001; Polka, 1995; Strange & Shafer, 2008; but cf. Dupoux, Peperkamp, & Sebastián-Gallés, 2001), and by including token variation (e.g., Dupoux et al., 2001; Højen & Flege, 2006).

In the previous section, it has become evident that interpreting our Limburgian listeners' behavior in the AXB task was a complicated endeavor. This interpretation difficulty was also due to the fact that prior studies on lexical tone discrimination (e.g., Hallé et al., 2004; Tong et al., 2015; Wayland & Guion, 2003; Xu et al., 2006) all used slightly different paradigms which could invoke different listening strategies (i.e., more or less acoustic-phonetic or language-specific). To ensure validity and comparability of future studies, it is crucial to enhance awareness with respect to the consequences of changing parameters such as ISI and token variability. More systematic investigations of the effects of these parameters on linguistic processing could also lead to better insight into the most suitable method to assess a given research question.

Related to this concern, studies greatly differ with respect to their attempts to improve the ecological validity of their tasks. Within the scope of a laboratory experiment, our AXB experiment was set up to resemble natural listening conditions as much as possible. Unlike prior AXB discrimination

experiments (e.g., Hallé et al., 2004; So & Best, 2014; Tong et al., 2015; Wayland & Guion, 2003), we added a substantial number of filler trials containing vowel and consonant differences to direct participants' attention away from the purpose of the experiment (i.e., lexical tone perception). Moreover, we included both between- (i.e., phonological) and within-category (i.e., phonetic) pitch differences. Also, we did not tell our participants beforehand that they should pay attention to pitch differences, whereas, rather surprisingly, most prior studies did explicitly direct their participants' attention to the contrast in question (e.g., Harnsberger, 2001; Polka, 1995; So & Best, 2014; Wayland & Guion, 2003). These participant instructions could have influenced the listening strategies adopted by participants. We think that awareness for the potential consequences of particular instructions on task behavior is of great importance.

Another methodological issue that we would like to address concerns the inclusion of a control group. Throughout this thesis, the performance of tone language speakers was statistically compared to a control group of non-tone language speakers. We think this is crucial in determining whether the tested behavior is specific to a particular group. However, prior studies either did not statistically compare performance of tone and non-tone language learning children (e.g., Singh et al., 2014), or they did not include a control group in the first place (e.g., Hay et al., 2015; Quam & Swingley, 2010). Both these latter publications showed that, from a particular age onward, English toddlers disregard pitch changes during word recognition. However, in the absence of a comparison with a native tone language learning group of toddlers, it may not be legitimate to conclude that the English toddlers' performance is a result of attunement to their native language phonology. Their conclusion might have been different if they had also included Mandarin Chinese children and added language as an independent variable in their analysis. The latter also holds for Singh et al. (2014), who did test both tone- and non-tone language speakers, but did not statistically compare these groups. That is, these studies might not have led to the same conclusions if they had

included language as a factor in their analyses, and, conversely, our conclusions might have been less complex if we had only tested Limburgian speaking subjects.

6.3.2 Choosing the distracter object: Familiar or novel?

In the word learning study reported in Chapter 5, participants were presented with pairs of novel objects. Both the target and the distracter object were unknown to child and adult listeners. The idea behind this was that 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. As such, a distracter preference in MP trials is a qualitatively different response from an absence of a preference for either target or distracter and thus offers the participant another option in interpreting the mispronounced label (e.g., Mani & Plunkett, 2011; Singh et al., 2015). This possibly provides a more sensitive measure of the strength of an MP effect, or, put differently, the degree of inhibition on target recognition. Indeed, White and Morgan (2008) for example showed that infants aged 1;7 incrementally increased their attention to a novel distracter over the familiar target with graded increases in the phonological distance between the novel word they heard and the familiar object label. Singh et al. (2015) were interested in whether the presence of a familiar or novel distracter would impact toddlers' behavior and put this question to the test in a word recognition paradigm with *familiar* words. They ran two experiments, one including a familiar distracter and the other with a novel distracter. The experiments yielded highly similar results, but the authors did not undertake a direct statistical comparison between the two. Moreover, they did not only change the familiarity of the distracter object across experiments, but also made changes to other variables, making it difficult to draw definite conclusions on the influence of distracter familiarity. The question how distracter familiarity might influence naming effects and sensitivity to MPs in

word recognition tasks with *newly learned words* remains an open question since there are no systematic comparisons of word learning studies using familiar and novel distracters yet.

As pointed out by Tsuji (2014), most infant and child researchers have quite strong intuitions as to what kinds of manipulations are suitable for answering specific questions, but systematic studies testing these intuitions are rare. Meta-analyses comparing effects across studies using different paradigms are gaining popularity in the field and are of great importance. Such an analysis comparing effects of distracter familiarity in novel word learning studies would be of great interest, or alternatively a novel experimental study directly comparing this influence across similar groups (i.e., similar with respect to language background, age, and the contrast at test).

6.3.3 Assessing the linguistic input of bidialectal children

Because chances were high that our Limburgian sample was exposed to quite heterogeneous input, we assessed the linguistic background of the Limburgian toddlers in Chapter 4 by administering a parental questionnaire that was designed for parents of *bilingual* children, the PaBiQ (COST Action IS0804, 2011; for more information, see e.g., Tuller, 2015).²⁵ Among other things, this questionnaire provides information on the quantity and quality of linguistic input and on early milestones. We adapted the questionnaire somewhat to our specific purposes, asking parents explicitly about their use of a Limburgian dialect relative to Dutch, and/or relative to any other languages.

Administering this questionnaire turned out to be a challenging task. In this section, we would like to mention some of the difficulties we have

²⁵ This questionnaire is a translation/adaptation of the Questionnaire for Parents of Bilingual Children (COST Action IS0804, 2011). It is the short version of a longer questionnaire piloted by research groups in several countries within COST Action IS0804, which was in part based on the ALEQ (Paradis, 2011) and the ALDeQ (Paradis, Emmerzael, & Sorenson Duncan, 2010).

encountered, hoping that these will feed future discussions about the reliability and validity of bilingual/bidialectal exposure questionnaires.

First, several questions with respect to the amount of language input had to be answered along a 5-point Likert scale ranging from *nooit* ('never'), *zelden* ('rarely'), *soms* ('sometimes'), *meestal* ('usually'), to *altijd* ('always'). These kinds of scales are quite common in social and psychological research (e.g., Schwarz, Bless, Bohner, Harlacher, & Kellenbenz, 1991). But, what does 'sometimes' mean? Since it is the middle value on the Likert scale, it could be interpreted as '50% of the time'. Intuitively, we found the difference between 'sometimes' and 'usually' quite big, whereas the distance between 'rarely' and 'sometimes' is less clear. Informal communication with other users of the questionnaire revealed that these kinds of questions are being treated and interpreted differently across researchers. To reduce the possibility of parents interpreting these qualitative labels differently, we decided to link them to percentages, ranging from 0, 25, 50, 75, to 100%, hoping that this would be a more reliable measure. Still, it remained unclear if the responses to each language then should add up to 100%, or whether a parent could for example answer 75% for both languages. Future investigations should definitely pay attention to both the reliability and validity of these kinds of scales.

It also turned out to be difficult for parents to judge their child's current skills in Limburgian. One of the questions for example was how well the child spoke Limburgian in comparison to age-matched peers. The difficulty could have arisen from the fact that Limburgian is a non-standardized regional variety that, other than Dutch, is not explicitly taught and hardly written. People are probably more used to talking about skills in Dutch, the standard language used at daycare and at school. Moreover, there is much variation between dialects. Exposure to this variation may result in less reliable judgments of what is correct and what is not.

A few questions in the PaBiQ assessed the amount of language mixing by the child and in the home. These were very demanding for Limburgian parents too. They were for example asked how often it occurred that the child

would start a sentence in Limburgian and end the sentence in Dutch (or vice versa), how often the child would use a Dutch word when speaking Limburgian (and vice versa), or how often Limburgian and Dutch were mixed within the home. We suggest that the difficulty in answering these questions resides from the possibility that Limburgian and Dutch might not be experienced as very distinct languages (or at least as less distinct than for example English and French). This is probably due to the fact that Limburgian and Dutch are varieties of one language.

In the end, we decided to disregard all questions that parents found hard to answer. Since the PaBiQ was designed for *bilingual* rather than *bidialectal* children, it might not be that surprising that some questions turned out not to be suitable. Future studies should definitely further look into language separation and/or perceived language distance in bidialectals and their ramifications for other language-related tasks. In fact, research on bidialectal language acquisition and processing is scarce. We do not know yet whether it makes a difference if children grow up with two closely related languages, e.g., Spanish and Catalan or Swedish and Norwegian (also called *close-language bilinguals*; Fennell, Tsui, & Hudon, 2016), compared to children growing up with two varieties of the same language, for example Limburgian and Dutch (i.e., *bidialectals*). Is the distinction even relevant to language processing, or does it just mirror a political distinction? We leave these questions for future research.

6.4 Conclusion

This thesis provided a cross-linguistic investigation into the developmental perception and representation of lexical pitch in tone and non-tone language listeners. By focusing on Limburgian, a restricted tone system, we addressed the potential influence of functional load and phonetic variability on lexical tone acquisition and processing. Whereas Limburgian and Dutch 6- to 12-month-old infants discriminated Limburgian tones equally well, adult speakers

of Limburgian showed greater sensitivity to word-level pitch patterns than adult speakers of Dutch. Moreover, Limburgian adults were also highly sensitive to within-category tone variation, possibly due to their exposure to a considerable amount of meaningful pitch variation.

With respect to the lexical encoding of pitch, we have shown that both Limburgian and Dutch 2,5- to 4-year-old toddlers responded to pitch changes in newly learned words. Nevertheless, word recognition was not inhibited completely by the pitch change. We assume that the lenient treatment of tonal MPs by Limburgians can also be ascribed to variability in the Limburgian input that has been put forward repeatedly in this thesis. The lack of an effect of language was surprising, since pitch is not lexically contrastive in Dutch. Unfortunately, the word learning study with Limburgian and Dutch adults could not clarify this issue as participants performed at ceiling, obscuring any possible effects. As a consequence, our results provided no evidence for the hypothesis that Limburgian and Dutch adults treat word-level pitch differently. The observation that child and adult speakers of Dutch showed sensitivity to non-native word-level pitch variation confirms the conclusion of previous researchers that the mere presence or absence of lexical tone contrasts in the native language in itself can not sufficiently explain cross-language perception of lexical tones.

Comparable word recognition studies comparing monolingual child and/or adult speakers of more and less restricted tone systems could further our understanding of the influence of functional load and surface variability on the acquisition and processing of a lexical tone system.

References

- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In H. C. Nusbaum (Ed.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.
- Best, C., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, *109*(2), 775-794.
- Best, C., McRoberts, G., & Sithole, N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345-360.
- Braun, B. & E. Johnson. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, *39*, 585-594.
- Braun, B., Kochanski, G., Grabe, E., & Rosner, B.S. (2006). Evidence for attractors in English intonation. *JASA*, *119*(6), 4006-4015.
- Broselow, E., Hurtig, R., & Ringen, C. (1987). The perception of second language prosody. In G. Loup, S., Weinberger (Eds.), *Interlanguage phonology. The acquisition of a second language sound system* (pp. 350-364). New York: Newbury House.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattasanone, N., Schwarz, I.-C., & Webster, D. (2015). Universality and

- language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, 36(6), 1459-1491. doi. <http://dx.doi.org/10.1017/S0142716414000496>
- Burnham, D. & Mattock, K. (2007). The perception of tones and phones. In O.-S. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 259-280). Amsterdam: John Benjamins.
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, 25, 426-439.
- COST Action IS0804 (2011). *Questionnaire for Parents of Bilingual Children (PaBiQ)*. <http://www.bi-sli.org>
- Demuth, K. (1995). The acquisition of tonal systems. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 111-134). Hillsdale: Lawrence Erlbaum.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness”. *Journal of the Acoustical Society of America*, 110(3), 1606-1618.
- Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants’ representation of familiar words. *Journal of Child Language*, 42, 447-465.
- van der Feest, S.V.H., & Johnson, E. (2016). Input-driven differences in toddlers’ perception of a disappearing phonological contrast. *Language Acquisition*, 23(2), 89-111.
- Fennell C.T., Tsui, A.S., & Hudon, T.M. (2016). Speech perception in simultaneously bilingual infants. In E. Nicoladis & S. Montanari (Eds.) *Bilingualism across the Lifespan* (pp. 43-62). De Gruyter Mouton.

- Fournier, R., & Gussenhoven, C. (2012). Measuring phonetic salience and perceptual distinctiveness: The lexical tone contrast of Venlo Dutch. *Revista Diadorim / Revista de Estudos e Linguísticos Literários do Programa de Pós-Graduação Vernáculas em Letras da Universidade Federal do Rio de Janeiro*, 12, 54-90.
- Francis, A., 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.
- Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987): The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23-45.
- Hallé, P. Chang, Y.-C. & Best, C. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- Harnsberger, J. (2001): On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, 110(1), 489-503.
- 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(1), 1-13.
- Hoffmann, C., Sadakata, M., Chen, A., Desain, P., & McQueen, J. (2014). Within-Category Variance and Lexical Tone Discrimination in Native and Non-Native Speakers. *Proceedings TAL 2014*, 45- 49.
- Højen, A., & Flege, J. (2006). Early learners' discrimination of second-language vowels. *Journal of the Acoustical Society of America*, 119(5), 3072-3084.

250 Setting the Tone

- Houston, D., Horn, D., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy, 12*(2), 119-145.
- Kuhl, P., Stevens, E., Hayashi, A., Degushi, 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.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition, 133*, 385-394.
- Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language, 38*(3), 606-627.
- Mattock, K. & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization. *Infancy, 10*, 241-265.
- 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.
- Mazuka, R., Hasegawa, M., & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology, 56*(2), 192-209. doi:10.1002/dev.21193.
- 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*(3), 407-420.
- Ota, M. (2003). The development of lexical pitch accent systems. *Canadian Journal of Linguistics, 48*(3/4), 357-383.

- Paradis, J. (2011). Individual Differences in Child English Second Language Acquisition: Comparing Child-Internal and Child-External Factors. *Linguistic Approaches to Bilingualism*, 1(3), 213-237.
- Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English Language Learners: Using Parent Report on First Language Development. *Journal of Communication Disorders*, 43, 474-497.
- Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America*, 97(2), 1286-1296.
- 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.
- Qin, Z., & Mok, P. (2011). Discrimination of Cantonese tones by Mandarin, English and French speakers. *Proceedings of the psycholinguistic representation of tone* (pp. 50-53). Hong Kong: Causal Productions.
- Quam, C., & Swingley, 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.
- Remijsen, A., & van Heuven, V. (2003). On the categorical nature of intonational contrasts. An experiment on boundary tones in Dutch. In J. van de Weijer, V. van Heuven, & H. van der Hulst (Eds.), *The Phonological Spectrum - Volume II: Suprasegmental Structure* (pp. 225-246). Amsterdam: John Benjamins.
- Riad, T. (2013). *The phonology of Swedish*. Oxford University Press.
- Rost, G.C., & McMurray, B. (2010). Finding the signal by adding noise: the role of noncontrastive phonetic variability in early word learning. *Infancy*, 15(6), 608-635.

252 Setting the Tone

- Sato, Y., Sogabe, Y., & Mazuka, R. (2010a). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, *22*(11), 2503-2513.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010b). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-119.
- Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, *5*(4), 489-522.
- Schwarz, N., Bless, H., Bohner, G., Harlacher, U., & Kellenbenz, M. (1991). Response scales as frames of reference: The impact of frequency range on diagnostic judgments. *Applied Cognitive Psychology*, *5*(1), 37-49.
- 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.
- Singh, L., Hui, T.J., Chan, C., & Golinkoff, R. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, *17*(1), 94-109.
- Singh, L., Poh, F.L.S., & Fu, C.S.L. (2016). Limits on monolingualism? A comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in Psychology*, *7*(667). doi:10.3389/fpsyg.2016.00667.
- So, C., & Best, C. (2008). Do English speakers assimilate Mandarin tones to English prosodic categories? In *Proceedings of the 9th annual conference of the International Speech Communication Association (Interspeech 2008), incorporating the 12th Australasian International Conference on Speech Science and Technology (SST 2008)* (p. 1120). Baixas, France: International Speech Communication Association.

- So, C., & Best, C. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, *53*, 273–293.
- So, C., & Best, C. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, *36*, 195-221.
- Strange, W., & Shafer, V. (2008). Speech perception in second language learners: Re-education of selective perception. J. Hansen Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-191). Amsterdam: John Benjamins.
- Tong, X., Lee, S.M.K., Lee, M.M.L., & Burnham, D. (2015). A Tale of Two Features: Perception of Cantonese Lexical Tone and English Lexical Stress in Cantonese-English Bilinguals. *PLoS ONE*, doi:10.1371/journal.pone.0142896.
- 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*(2), 111-124.
- Tsao, F.-M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in Psychology*, doi:10.3389/fpsyg.2017.00558.
- Tsuji, S. (2014). *The Road to Native Listening. Language-General Perception, Language-Specific Input*. Doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.
- Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C.T. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: language-specific developmental changes in speech

perception abilities. *1994 International Conference on Spoken Language Processing* (pp. 1695-1698). Yokohama, Japan.

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: disentangling bilingualism from Language Impairment* (pp. 301-330). Bristol: Multilingual Matters.

Tyler, M., Best, C., Faber, A., & Levitt, A. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4-21.

Wayland, R., & Guion, S. (2003). Perceptual discrimination of Thai tones by naive and experienced learners of Thai. *Applied Psycholinguistics*, *24*, 113-129.

Werker, J., & Tees, R. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*(6), 1866-1878.

White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

Xu, Y., Gandour, J., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch contrasts. *Journal of the Acoustical Society of America*, *120*(2), 1063-1074.

Yeung, H.H., Chen, K.H., & Werker, J. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123-139.

7. Samenvatting in het Nederlands

7.1 Inleiding

Iedere taal op de wereld is opgebouwd uit een set klanken. Die set verschilt per taal. Klanken, zoals bijvoorbeeld de medeklinkers [p] in *pak* en [t] in *tak*, zijn de bouwstenen van woorden. Een verschil in klank kan tot een ander woord leiden. Als de [p] in het woord *pak* wordt vervangen door een [t] ontstaat er een ander woord: *tak*.

Veel Limburgse dialecten hebben naast klinkers en medeklinkers nog een derde ‘trucje’ om woorden van elkaar te laten verschillen: *lexicale toon*. Als iemand uit Roermond bijvoorbeeld achtereenvolgens het woord *been* en daarna het woord *benen* in het Limburgs zegt, dan zal opvallen dat deze woorden qua klinkers en medeklinkers identiek zijn, namelijk *bein*; het verschil zit hem in hun melodie. *Benen* heeft stoottoon (accent 1, een dalend toonhoogteverloop), *been* heeft sleeptoon (accent 2, een dalend-stijgend toonhoogteverloop). Andere *toontalen* zijn bijvoorbeeld het Chinees, veel talen in Afrika en, iets dichterbij, het Noors en het Zweeds. Naar schatting zijn zo’n 60-70% van de talen op de wereld toontalen.

Binnen die grote familie van toontalen bestaan echter ook aanzienlijke verschillen. Enerzijds met betrekking tot de functionaliteit, of het *gewicht* van toon (in het Engels zegt men *functional load*), anderzijds met betrekking tot hoe veranderlijk de uitspraak van bepaalde melodieën is onder invloed van andere talige eigenschappen (bijv. zinsklemtoon, positie in de zin). We weten echter nog weinig over de mogelijke invloed die deze verschillen hebben op het leren van een toontaal, omdat de meeste onderzoeken zich hebben gericht op Aziatische toontalen. Met onderzoek naar de verwerving en verwerking van toon in het Limburgs vergroten we onze kennis van de invloed van functionaliteit en uitspraakvariatie op de verwerving en verwerking van lexicale toon.

Wat het Limburgs ten opzichte van veel andere toontalen zo interessant maakt is dat de melodieën van accent 1 en accent 2 sterk variëren afhankelijk van de *prosodische context* waarin het betreffende woord wordt gesproken. Als het woord in isolatie wordt uitgesproken, of als het woord aan het einde van een bevestigende zin staat en de zinsklemtoon heeft, gelden de zojuist beschreven melodieën. Maar in andere gevallen, bijvoorbeeld in een vraagzin, worden accent 1 en accent 2 anders uitgesproken. Die variatie in uitspraakmogelijkheden zou voor een kind dat Limburgs leert wel eens een uitdaging kunnen zijn, omdat het uit al die verschillende realisaties moet opmaken dat ze eigenlijk allemaal tot maar twee categorieën behoren, namelijk accent 1 of accent 2.

Een tweede eigenschap waarin het Limburgs zich onderscheidt van reeds onderzochte toontalen is dat toon relatief weinig functionaliteit of gewicht heeft. Eerder onderzoek met volwassen toontaalsprekers heeft uitgewezen dat het gewicht van toon gevolgen kan hebben voor hoe sprekers van zo'n taal tonen verwerken, d.w.z. hoe goed ze bijvoorbeeld tonen kunnen waarnemen of hoe bewust ze zich van tonen zijn. In het Mandarijn Chinees heeft bijna elke lettergreep één van vier mogelijke tonen. Zo zijn er ook veel woorden die alleen maar door de toon van elkaar verschillen (denk aan het *bein* voorbeeld uit het Limburgs). Zulke gevallen noemen we *minimale paren*. In het Limburgs zijn er relatief weinig van zulke tonale minimale paren. Het Limburgs heeft bovendien 'slechts' twee tonen, en daarnaast kunnen die tonen alleen op bepaalde lettergrepen (bijv. met hoofdklemtoon) voorkomen. Al deze 'beperkingen' zorgen ervoor dat toon in het Limburgs relatief minder belangrijk of aanwezig is. De vraag is wat voor gevolgen dat heeft voor het leren van toon door baby's en kinderen, en ook voor de verwerking van toon door volwassenen.

Met behulp van een reeks gedragsexperimenten hebben we onderzocht hoe baby's en peuters en volwassen sprekers van het Limburgs toon waarnemen en interpreteren. We vergeleken ze telkens met sprekers van het Nederlands, een taal die wel intonatie gebruikt maar geen lexicale toon.

7.2 Experiment 1: Toonperceptie bij Limburgse en Nederlandse baby's

Kinderen leren al in de loop van het eerste levensjaar uit welke klanken hun moedertaal is opgebouwd. Aanvankelijk zijn baby's heel flexibel: ze zijn in principe in staat om bijna alle klankcontrasten die er op de wereld bestaan waar te nemen. Welke klankverschillen ze in de loop van het eerste levensjaar blijven waarnemen hangt er onder andere vanaf in welke taalomgeving ze opgroeien. Rond zes maanden weten ze welke klinkers belangrijk zijn en welke niet, en rond twaalf maanden welke medeklinkers er in hun taal gebruikt worden. Dit proces wordt ook wel *perceptual reorganization* genoemd.

Naar de verwerving van klinkers en medeklinkers is relatief veel onderzoek gedaan; de verwerving van tonen heeft tot nu toe minder aandacht gekregen. Sommige studies laten zien dat kinderen rond de negen maanden hebben ontdekt of tonen binnen woorden betekenisvol zijn of niet. Andere studies lieten een ingewikkelder patroon zien. Afhankelijk van de tonen die er getest werden waren baby's meer of minder gevoelig.

Om te achterhalen of Limburgse en Nederlandse baby's²⁶ van 6, 9 en 12 maanden het verschil tussen accent 1 en accent 2 kunnen horen, gebruikten we de *visuele fixatieprocedure* (*visual fixation/habituation procedure*). In ons experiment kregen baby's herhalingen van het nonsens-woordje *taag* (afgeleid van *daag*, 'dag(en)') te horen. De helft van de proefpersonen met accent 1, de andere helft met accent 2. Ondertussen kijken ze naar een bewegend plaatje op een scherm. De mate van aandacht van de baby's leiden we af uit hun kijkgedrag: zolang de baby naar het scherm kijkt, kunnen we aannemen dat ze naar de geluiden luistert die er worden aangeboden. Aanvankelijk is de situatie heel interessant. Echter, hoe vaker ze hetzelfde hoort (*taag, taag, taag, ...*), hoe

²⁶ Met 'Limburgs' en 'Nederlands' wordt hier consequent verwezen naar taalachtergrond, niet naar nationaliteit. De onderzochte sprekers van het Limburgs hadden ook allemaal de Nederlandse nationaliteit.

saaier het wordt. Op een zeker moment treedt er gewenning op: de kijktijd naar het beeldscherm neemt af. Zodra er gewenning optrad lieten we de baby's het woordje *taag* met de andere toon horen. Als de baby's horen dat er iets nieuws wordt aangeboden zal het kijken weer toenemen. Als een baby het verschil tussen accent 1 of accent 2 echter niet kan horen, zal haar aandacht na het aanbieden van het woord met de andere toon verder afnemen.

Wat we hebben gevonden is dat zowel Limburgse als Nederlandse baby's gedurende hun eerste levensjaar gevoelig zijn voor het verschil tussen accent 1 en accent 2. Onze resultaten leveren geen bewijs voor een leeftijdsgebonden verandering in die gevoeligheid. Voor de Limburgse baby's hadden we die gevoeligheid voorspeld. Voor het gedrag van de Nederlandse baby's zijn er meerdere verklaringen. Het zou kunnen dat het tooncontrast akoestisch heel opvallend is (en dat dus eenieder het zou kunnen waarnemen), maar het zou ook kunnen dat Nederlanders de Limburgse tonen waarnemen als intonatie. We noemen dit *perceptuele assimilatie*. De intonatie-melodieën van bevestigende en vragende zinnen vertonen gelijkenis met het toonhoogteverloop van respectievelijk accent 1 en accent 2.

Om deze twee verklaringen verder te onderzoeken hebben we ook waarnemingsexperiment met Limburgse en Nederlandse volwassenen uitgevoerd.

7.3 Experiment 2: Toonperceptie bij Limburgse en Nederlandse volwassenen

Om te weten te komen of Limburgse en Nederlandse volwassenen wél van elkaar verschillen m.b.t. hun gevoeligheid voor toon, en om te onderzoeken of het succes van de Nederlandse baby's in experiment 1 te verklaren is door akoestische opvallendheid of door perceptuele assimilatie, namen een groep Limburgse en Nederlandse volwassenen deel aan een taalwaarnemingsexperiment: Een *AXB discriminatietaak*.

Tijdens dit experiment hoorden deelnemers meerdere keren opeenvolgingen van drie nonsens-woorden, bijvoorbeeld *taag₁ – taag₂ – taag₂*

(het getal duidt de toon van het woord aan). Na zo'n drietal moesten de deelnemers telkens beslissen of het tweede woord (*taag₂*) meer op het eerste (*taag₁*) of meer op het derde woord (*taag₂*) leek door zo snel mogelijk op een knop ("1" of "3") te drukken. Het enige verschil tussen de woorden was de toon. Een hoog percentage correcte antwoorden betekent dus dat iemand het verschil tussen accent 1 en accent 2 goed kan horen.

Om te kunnen achterhalen of perceptuele assimilatie een rol speelde in de waarneming van de Nederlanders gebruikten we verschillende typen stimuluswoorden. Eenlettergrepige woorden (*taag* en *moon*) waarin het toonhoogteverloop noodzakelijkerwijs op de *laatste* lettergreep van het woord gerealiseerd wordt, en tweelettergrepige woorden (*keeve* en *perger*) met klemtoon op de eerste lettergreep, waarin het toonhoogteverloop op de *voorlaatste* lettergreep van het woord gerealiseerd wordt. Bovendien worden accent 1 en accent 2 in niet-finale positie iets anders gerealiseerd dan in finale positie: In niet-finale positie heeft accent 1 een dalend toonhoogteverloop en wordt accent 2 als een onveranderlijke, stabiele (*level*) toon gerealiseerd. Eerder onderzoek had laten zien dat Nederlanders minder gevoelig zijn voor toonverschillen als die toonverschillen op een plek binnen het woord voorkomen waar in het Nederlands geen betekenisvolle intonatieverschillen zitten, bijvoorbeeld op een voorlaatste lettergreep. Als perceptuele assimilatie een rol speelt verwachtten we dat Nederlandse volwassenen meer correcte antwoorden zouden geven na drietallen met eenlettergrepige woorden dan na drietallen met tweelettergrepige woorden.

Naast deze drietallen voegden we ook nog drietallen toe met drie keer dezelfde toon (bijv. *keeve_{1a}* – *keeve_{1a}* – *keeve_{1b}*). In deze drietallen moesten deelnemers kleine verschillen *binnen* een tooncategorie (bijv. accent 1) onderscheiden om het goede antwoord te kunnen geven. In dit specifieke geval zou het juiste antwoord "1" zijn. Met behulp van deze drietallen wilden we onderzoeken of de toonperceptie van Limburgse volwassenen meer *categoriaal* is dan die van Nederlanders. Eerder onderzoek had namelijk laten zien dat

sprekers van bijvoorbeeld het Mandarijn Chinees beter zijn in drietallen waarin ze *tussen* tooncategorieën moesten onderscheiden dan in drietallen waarin ze *binnen* tooncategorieën verschillen moesten horen, terwijl sprekers van het Engels in beide gevallen hetzelfde presteerden. De verklaring hiervoor is dat de waarneming van de Chinezen als het ware gestuurd wordt door de tooncategorieën waarover ze beschikken. Hun kennis van categorieën *filtert* hun waarneming. Voor hen is het belangrijk om die categorieën te kunnen onderscheiden. Verschillen *binnen* die categorieën spelen geen rol en kunnen genegeerd worden. Sprekers van het Engels hebben geen lexicale tooncategorieën en luisteren dus op een andere manier, ze letten meer op akoestische verschillen en worden niet of minder beïnvloed door kennis van bepaalde categorieën binnen hun taal.

Onze resultaten laten zien dat Limburgse volwassenen over het algemeen significant gevoeliger zijn voor lexicale toon dan Nederlandse volwassenen. Ze waren significant accurater in het horen van het verschillen *tussen* tooncategorieën. Tegen verwachting in vonden we in de Limburgers geen verschil in accuraatheid tussen *binnen*- en *tussen*-categorie discriminatie. Limburgers waren dus ook heel gevoelig voor verschillen *binnen* tooncategorieën. Wij stellen voor dat dit te maken heeft met het feit dat Limburgers ervaring hebben met een grote verscheidenheid aan toonhoogteverschillen. Ten eerste doordat de uitspraak van accent 1 en accent 2 aanzienlijk kan verschillen onder invloed van de prosodische context. Dit is een van de eigenschappen van het Limburgse toonsysteem waarin het Limburgs duidelijk verschilt van eerder onderzochte toontalen. Dit verschil zou kunnen verklaren waarom onze resultaten afwijken van die van eerder onderzoek. Anderzijds konden onze Limburgse deelnemers ook allemaal Nederlands. Aangezien het Nederlands ook over een uitgebreid intonatiesysteem beschikt, hadden onze deelnemers dus allemaal ervaring met twee talen die uitgebreid gebruik maken van toonhoogteverschillen. Deze rijke ervaring leidt wellicht tot een grotere gevoeligheid voor toonhoogteverschillen in het algemeen.

Met betrekking tot de Nederlandse volwassenen waren de resultaten minder duidelijk. Tegen verwachting in waren ze accurater in tweelettergrepige drietallen dan in eenlettergrepige drietallen. Nadere akoestische analyse van onze stimuli doet vermoeden dat accent 1 en accent 2 in de tweelettergrepige woorden akoestisch meer verschillend waren en dat die opvallendheid de waarneming van de Nederlanders bepaald heeft. Er is meer onderzoek nodig om te kunnen bevestigen of de gemeten akoestische verschillen inderdaad algemeen kenmerkend zijn voor accent 1 en accent 2 in niet-finale positie. In *tussen*-categorie discriminatie waren Nederlanders significant slechter dan in *binnen*-categorie discriminatie. Ook dit week af van onze verwachtingen. Als Nederlandse deelnemers de accent 1 en accent 2 woorden 1 op 1 met Nederlandse intonatiepatronen hadden vergeleken had hun accuraatheid in *tussen*-categorie discriminatie hoger moeten zijn dan we hebben vastgesteld. Ondanks dat waren ze wel significant beter dan kans, wat betekent dat ze toch een zekere gevoeligheid voor de Limburgse tonen lieten zien. Meer onderzoek is noodzakelijk om te kunnen bepalen wat het gedragspatroon van de Nederlanders heeft veroorzaakt.

De vraag was niet langer of Limburgers en Nederlanders de Limburgse tonen kunnen waarnemen (het antwoord voor Limburgers: Ja; voor Nederlanders: Ja, maar minder goed dan Limburgers), maar welke functie zij aan die tonen toekennen. Beschouwen ze tonen als relevante eigenschappen van woorden?

7.4 Experiment 3: De lexicale representatie van toon in Limburgse en Nederlandse peuters en volwassenen

In het Limburgs vormen tonen belangrijke eigenschappen van woorden. Het woord *boom* 'boom' moet met accent 2 worden uitgesproken; met accent 1 zou het vreemd klinken. Het omgekeerde geldt voor *sjaop* 'schaap', dat met accent 1 wordt uitgesproken. Concreet betekent dit dat een Limburger woorden in het langetermijngeheugen zou moeten *kenmerken* met een bepaalde toon. De vraag

is of dat inderdaad gebeurt en of een spreker van het Limburgs raar op zou kijken als *sjaop* met accent 2 in plaats van met accent 1 wordt uitgesproken. Zal hij het woord nog wel herkennen? Engelse kinderen ondervinden bijvoorbeeld ook enige hinder als het woord *car* als *dar* wordt uitgesproken. Die problemen tijdens de woordherkenning zijn terug te voeren op het feit dat de waargenomen vorm van het woord niet overeenkomt met de vorm die in het geheugen is opgeslagen. Een Nederlander hoeft níet te weten welke melodie bij welk woord hoort en hoeft dat dus ook niet in zijn mentale lexicon op te slaan.

Met ons derde experiment wilden we onderzoeken of Limburgse en Nederlandse peuters tussen de 2,5 en 4 jaar en volwassenen toon opslaan als eigenschap van woorden in het mentale lexicon. Om dit te achterhalen ontwierpen we een woordleerexperiment waarbij we het *kiezend-kijken-paradigma* (*preferential looking paradigm*) gebruikten. Tijdens dit experiment leerden deelnemers twee nieuwe nonsens-woorden (*taaf* en *moon*), één woord met accent 1 en één woord met accent 2, en testten we vervolgens of ze gevoelig waren voor veranderingen (*mispronunciations*) in de toon van het woord. Zou een deelnemer het opmerken als een woord geleerd met accent 1 met accent 2 zou worden uitgesproken, of omgekeerd? Om de woorden te leren werd deelnemers 20 herhalingen van het woord in combinatie met een plaatje van een onbekend object (de *target*) gepresenteerd. Vervolgens kreeg de deelnemer naast elkaar twee plaatjes te zien, namelijk de *target* en een afleider (*distracter*). Tegelijkertijd hoorde hij bijvoorbeeld *Kiek nao de taafi! Zuus-ten'm?* ('Kijk naar de taaf. Zie je-n'm?'). Het kijkgedrag van de deelnemer naar aanleiding van de aangeboden taaluiting kan verschillende patronen vertonen. Als de deelnemer het woord *taafi* herkent zal hij relatief langer naar de *target* kijken dan naar de *distracter*. Als de kijktijd naar beide plaatjes níet verschilt betekent dat dat de deelnemer het woord niet heeft herkend. Alle deelnemers, zowel kinderen als volwassenen, waren succesvol in het leren van de twee nieuwe woorden.

Vervolgens wilden we weten of Limburgers en Nederlanders verschillend zouden reageren als het geleerde woord met de 'verkeerde' toon

zou worden uitgesproken. Daarom hoorden deelnemers niet alleen zinnen waarin het woord goed werd uitgesproken ('correcte uitspraak conditie'), maar ook zinnen waarin het woord met de andere toon werd uitgesproken ('incorrecte uitspraak conditie'). Uit eerder onderzoek weten we dat de herkenning van een woord belemmerd (of zelfs onmogelijk) wordt als het foutief wordt uitgesproken. Als Limburgers hebben onthouden dat *taaf* een accent 1-woord is zullen ze relatief minder lang naar de *target* kijken in de incorrecte uitspraak-conditie (*taaf* uitgesproken met accent 2) dan in de correcte uitspraak-conditie (*taaf* met accent 1). Op basis van eerder onderzoek met toontaalsprekers verwachtten we dat de herkenning van *taaf* in de incorrecte uitspraak-conditie voor Limburgers (enigszins) belemmerd zou worden doordat dat wat de deelnemer hoort (accent 2) niet overeenkomt met wat er in het mentale lexicon is opgeslagen (accent 1). We verwachtten dat het kijkgedrag van Nederlanders in de correcte en incorrecte conditie niet significant zou verschillen. De aanname is immers dat Nederlanders woorden in hun mentale lexicon niet markeren voor toon. Wat ze ook horen, het kan niet 'botsen' met wat ze hebben opgeslagen.

Een andere vraag was hoe ernstig het gebruik van een foutieve toon voor de woordherkenning van een Limburger zou zijn. Zou het net zo ernstig zijn als *poes* uitspreken als *toes*, bijvoorbeeld? Vanwege de relatief lage functionaliteit van toon in het Limburgs en vanwege de variatie in uitspraak van accent 1 en accent 2 naar aanleiding van de prosodische context zou het kunnen dat met name kinderen die de taal nog aan het leren zijn niet zoveel hinder ondervinden van een verandering in toon.

Onze resultaten voor de 2,5- tot 4-jaar oude kinderen laten zien dat zowel Limburgse als Nederlandse kinderen reageren op een verandering in toon. Ze kijken relatief minder lang naar de *target* in de incorrecte conditie dan in de correcte conditie. Dit betekent dat hun herkenning van het geleerde woord in zekere mate beïnvloed wordt door de toon waarmee het woord wordt uitgesproken. Desalniettemin keken de kinderen wel nog steeds significant

meer naar de *target* dan naar de *distracter*, wat betekent dat woordherkenning, ondanks een verandering in toon, nog steeds mogelijk was. Voor de Limburgers was deze gevoeligheid voor toon naar verwachting. Het feit dat de incorrecte toon relatief minder impact had op woordherkenning dan in eerder onderzoek met bijvoorbeeld kinderen die Mandarijn Chinees leren kan weer terug te voeren zijn op de eerder genoemde verschillen die het Limburgse toonsysteem zo uniek maken ten opzichte van de toonsystemen die eerder onderzocht zijn. Een ander factor die aan dit ‘milde’ effect van de toonverandering zou kunnen bijdragen is het feit dat Limburgers in hun omgeving met dialectvariatie in aanraking komen. Die variatie zou ook kunnen leiden tot minder stabiele representaties van toon in het langetermijngeheugen. Studies met kinderen die een vergelijkbaar toonsysteem leren, bijvoorbeeld Zweeds, zouden meer uitsluitsel kunnen geven over de invloed van functionaliteit en uitspraakvariatie op de ontwikkeling van lexicale tooncategorieën.

De gevoeligheid die de Nederlandse kinderen lieten zien was onverwacht. Mogelijk werden ze ook in dit experiment door hun kennis van Nederlandse intonatiepatronen beïnvloed. Een andere optie is dat de Nederlandse kinderen het tooncontrast als een verschil in klinkerduur hebben waargenomen. Eerder onderzoek heeft laten zien dat sprekers van een niet-toontaal het verschil tussen lexicale tonen in een vreemde taal als een duurverschil kunnen opvatten, mits duurverschillen in hun moedertaal betekenisvol zijn. Klinkerduur is voor Nederlanders een bekend concept. Bovendien speelt duur een rol in de realisering van klemtoon.

Wat de volwassenen betreft leveren onze resultaten geen bewijs voor de hypothese dat Limburgers gevoelig zijn voor toonverschillen noch voor de hypothese dat Limburgers en Nederlanders zich in dit opzicht verschillend zouden gedragen. De volwassenen keken in beide geteste condities significant langer naar de *target*. Dit zou een taakeffect kunnen zijn; misschien was het experiment te eenvoudig / niet gevoelig genoeg om een verschil te kunnen vaststellen. Het toevoegen van een ‘incorrecte uitspraak-conditie’ met bijvoorbeeld een medeklinkerverandering (bijv. *taaf* → *paaf*) zou in de

toekomst ook meer uitsluitsel kunnen geven over het relatieve belang van toon in woordherkenning ten opzichte van klinkers en medeklinkers.

7.5 Conclusie

Dit proefschrift heeft laten zien dat Limburgse en Nederlandse baby's in het eerste levensjaar nog niet verschillen in hun gevoeligheid voor (Limburgse) lexicale tonen. Beide nemen ze het verschil tussen accent 1 en accent 2 waar. Bij Limburgse en Nederlandse volwassenen zien we wel een verschil: Limburgers zijn significant gevoeliger dan Nederlanders. Ook waren Limburgse volwassenen heel gevoelig voor verschillen binnen tooncategorieën. Wij vermoeden dat dit het resultaat is van het feit dat Limburgers ervaring hebben met een complex prosodisch systeem. Uit het toonhoogteverloop van elke lettergreep met hoofdklemtoon moeten zij afleiden welke lexicale toon deze lettergreep heeft én welk intonatiepatroon er gerealiseerd wordt. Beide zijn immers, op verschillende niveaus van taal (woord vs. zin), betekenisvol.

Met betrekking tot de lexicale representatie van toon hebben we laten zien dat zowel Limburgse als Nederlandse 2,5- tot 4-jaar oude kinderen gevoelig zijn voor toonverschillen binnen nieuwe woorden. Ondanks het toonverschil was woordherkenning wel nog steeds mogelijk. Deze relatief milde invloed van toonverschillen op woordherkenning zijn mogelijk te wijten aan de hoeveelheid toonhoogtevariatie waaraan Limburgers dagelijks worden blootgesteld. Dat ook Nederlandse kinderen reageerden op toonverschillen binnen woorden is wellicht te verklaren door hun rijke ervaring met prosodie in het Nederlands. Doordat het experiment met volwassenen geen effecten opleverde kunnen we hier geen conclusies aan verbinden.

Meer onderzoek is noodzakelijk om de precieze oorzaak van de onverwachte gevoeligheid voor lexicale toon binnen de Nederlanders vast te kunnen stellen. Vergelijkbaar onderzoek onder kinderen en volwassen sprekers van bijvoorbeeld het Zweeds zou kunnen bijdragen aan ons begrip

266 Setting the Tone

van de rol die functionaliteit en uitspraakvariatie spelen in de verwerving en verwerking van een lexicaal toonsysteem.

Curriculum Vitae

Stefanie Ramachers was born on November 11, 1987 in Geleen, the Netherlands. She has an MA in German language and culture and an MA in Linguistics, both from Radboud University, Nijmegen.

After graduation in 2011, the Netherlands Organization for Scientific Research granted her funding for her individual PhD project on the acquisition of lexical tone in Limburgian dialects at Radboud University.

At the same time, she started working as a lecturer at Radboud University's department of German language and culture. Here, she developed and taught courses in German linguistics, ranging from phonetics and phonology, syntax, semantics and pragmatics to language acquisition, experimental research, and foreign language didactics.

In 2013 Stefanie was granted Radboud University's Frye Stipend for promising female researchers. For her endeavours as a lecturer, her students nominated her for the 2015 Faculty of Arts Teaching Award. In 2017, Stefanie won the scientific pitch competition Radboud Talks.

In 2016 and 2017, Stefanie also worked freelance for the Dutch publisher Malmberg. She was editor-in-chief of a digital adaptive grammar trainer for German as a foreign language that is now used in Dutch secondary education.

Stefanie currently works as an assistant professor in German linguistics and as a post-doctoral researcher in applied linguistics and knowledge valorization at Radboud University.