ProPro Workshop 2017

PROCESSING PROSODY ACROSS LANGUAGES, VARIETIES, AND NATIVENESS

Booklet

SpeechNet BaWü
Stuttgart - Tübingen - Konstanz - Freiburg

AUGUST 31 - SEPTEMBER 1, 2017 | TÜBINGEN, GERMANY |
ALTE AULA | MÜNZGASSE 30

ORGANISED BY:
ANN KATHRIN GROHE, UNIVERSITY OF TÜBINGEN
NADJA SCHAUFFLER, UNIVERSITY OF STUTTGART
KATHARINA ZAHNER, UNIVERSITY OF KONSTANZ

Preface

Dear ProPro participants,

It is a pleasure to welcome you to the cosy city of Tübingen for the SpeechNet BaWü-Workshop on the "Processing of Prosody"! We received submissions from all over the world. This is great! Over the course of the next two days we are going to discuss research on how prosodic properties of speech are processed. Most projects are part of the presenters' dissertation, so this workshop enables junior researchers from the field to present their work and receive valuable feedback.

We welcome feedback from other junior researchers as much as from more experienced participants, such as our keynote and invited speakers: Anne Cutler from the MARCS Institute (Western Sydney University), Sarah Bibyk from the Vanderbilt University, Kiwako Ito from the Ohio State University, and Giusy Turco from the Sorbonne Nouvelle Paris. We are grateful that you have made it all the way to Tübingen and are looking forward to your talks and comments on students' work.

We also want to promote the SpeechNet BaWü[1] that we, the organisers are all part of - in the present workshop. The SpeechNet BaWü is a research network of (female) experimental linguists working on speech. The scientists are located at the southern German universities of Freiburg (JProf. Adriana Hanulíková), Konstanz (Prof. Bettina Braun), Stuttgart (Prof. Sabine Zerbian), and Tübingen (Prof. Andrea Weber) with close collaboration with Anne Cutler's research group at the MARCS Institute ("Cutler's Corner"). We acknowledge funding from the German Academic Exchange Service (DAAD) that made this collaboration possible. The SpeechNet-members have expertise in cross-language and native versus second language processing, with a special focus in prosody and in listening to variable and accented speech using different methodological approaches.

The network provides a platform for mutual exchange on experimental speech research in the region (and beyond), thereby enriching the scientific learning experience for early career researchers and fostering national and international collaborations. We are proud to present experimental work resulting from almost two years of SpeechNet collaboration in the present workshop. We have dedicated a special session just for SpeechNet in which the members present their projects. Moreover, we emphasise Anne Cutler's and Heather Kember's enormous support which has allowed for collaborative projects with researchers from the MARCS Institute, including research stays for SpeechNet members in Sydney and for MARCS members in Southern Germany. Thank you very much!

1) http://ifla.uni-stuttgart.de/index.php?article_id=218/speechnet-bawueamp;clang=0

Preface

Finally, we thank the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63) for the financial support, which has made the present workshop possible. Moreover, we cordially thank our student assistants Thanh Lan Truong and Maria Lazareva for supporting the workshop organisation!

We hope you enjoy the scientific and social exchange during the next two days and that you can use this opportunity to make valuable connections!

Cheers,
Ann-Kathrin, Nadja, and Katharina

Content

Programme

THURSDAY

08:45 - 09:15	REGISTRATION
MORNING SESSION	CHAIR: ANDREA WEBER
09:15 - 09:30	WELCOME
09:30 - 10:30	SARAH BIBYK (UNIVERSITY OF ROCHESTER): LISTENERS INTERPRET RISING AND FALLING INTONATION PRIOR TO THE FINAL BOUNDARY
10:30 - 11:00	COFFEE BREAK
11:00 - 12:30	SPECIAL SESSION: SPEECHNET BAWÜ
12:30 - 13:00	CONG ZHANG (UNIVERSITY OF OXFORD): IDENTIFYING INTERROGATIVE TUNES IN TIANJIN MANDARIN
13:00 - 14:00	LUNCH - INDIVIDUALLY IN THE OLD TOWN (SEE LUNCH RECOMMENDATIONS ON PAGE 52)
AFTERNOON SESSION	CHAIR: SABINE ZERBIAN
14:30 - 15:00	TALINA WEBER (UNIVERSITY OF KONSTANZ), MUNA SCHÖNHUBER & JANET GRIJZENHOUT: INTONATION OF POLAR QUESTIONS PRODUCED BY 2.5- TO 4-YEAR-OLDS
15:00 - 15:30	NÁDIA BARROS (UNIVERSITY OF LISBON) & SÓNIA FROTA: PROSODIC PHRASING IN EUROPEAN PORTUGUESE VARIETIES
15:30 - 16:00	SIMON WEHRLE (UNIVERSITY OF COLOGNE), TIMO B. ROETTGER & MARTINE GRICE: TRACKING THE PERCEPTUAL EFFECTS OF BACKCHANNEL BEHAVIOUR IN ASPERGER SYNDROME AND L2 SPEECH
16:00 - 16:30	COFFEE BREAK
16:30 - 17:30	KIWAKO ITO (THE OHIO STATE UNIVERSITY): INVESTIGATING DEVELOPMENT OF LINGUISTIC AND NON-LINGUISTIC PROSODY
19:00	WORKSHOP DINNER (ALTE KUNST, MARKTGASSE 8)

Special Session

THURSDAY
11:00 - 12:30

Members of SpeechNet BaWü present their work conducted in collaboration with Anne Cutler and Heather Kember (the research was supported by a DAAD research grant (Australia – Germany Joint Research Cooperation Scheme)).

Yuki Asano, Ann-Kathrin Grohe, Heather Kember, Andrea Weber: Perception of Australian English Uptalk

Heather Kember, **Ann-Kathrin Grohe, Katharina Zahner**, Bettina Braun, Andrea Weber, Anne Cutler: Similar prosodic structure perceived differently in German and English

Katharina Zahner, Heather Kember, Bettina Braun: Mind the peak – When museum is temporarily understood as musical in Australian English

Programme

FRIDAY

MORNING SESSION	CHAIRS: ANN-KATHRIN GROHE, KATHARINA ZAHNER
09:00 - 10:00	GIUSEPPINA TURCO (SORBONNE NOUVELLE PARIS) & SABINE ZERBIAN: TESTING FOR PROCESSING ADVANTAGES OF LINGUISTIC EFFECTS OF FOCUS IN L1 AND L2
10:00 - 10:30	AYA TAKEDA (UNIVERSITY OF HAWAII): INTONATIONAL MARKING OF INFORMATION STRUCTURE IN L2 ENGLISH
10:30 - 11:00	COFFEE BREAK
11:00 - 12:30	POSTER SESSION
12:30 - 14:00	LUNCH (PROVIDED)

AFTERNOON SESSION	CHAIR: YUKI ASANO
14:00 - 14:30	ANTJE HEY (BIELEFELD UNIVERSITY), KATHARINA NIMZ & PETRA WAGNER: COMPREHENSIBILITY IN L2 SPEECH: LEXICAL STRESS VERSUS SENTENCE ACCENT
14:30 - 15:00	JENNY YU (WESTERN SYDNEY UNIVERSITY), HEATHER KEMBER, ROBERT MAILHAMMER & ANNE CUTLER: PITCH AND SYNTACTIC DISAMBIGUATION IN ENGLISH AND GERMAN: AN EYE-TRACKING STUDY
15:00 - 15:30	COFFEE BREAK
15:30 - 16:30	ANNE CUTLER: WHAT IS SPECIAL ABOUT PROSODY IN PROCESSING?
16:30 - 16:45	FAREWELL
18:00 - 20:00	SOCIAL EVENT: STOCHERKAHNFAHREN

Poster Session

FRIDAY

11:00 - 12:30

1. **L. Ann Burchfield (Western Sydney University)** & Heather Kember: Effect of language background on L2 focus perception

2. **Hui-Ching Chen (University of Potsdam),** Stephen Crain, Barbara Höhle: Comprehension of prosodic and syntactic focus marking in Mandarin Chinese - Data from children and adults

3. **Syrine Daoussi Díaz (Universidad Autónoma de Barcelona)**, Lorraine Baqué & Marta Estrada: Are French speakers… "stress deaf"?

4. **Sichang Gao (Shanghai International Studies University)**: The Study of Prosodic Grouping by Chinese as a Second Language Learners

5. **Alastair Graham-Marr (Tokyo University of Science)**: Do prosodic differences between English and Japanese disadvantage Japanese learners of English?

6. **Sergio Quiroz (Humboldt University of Berlin)**: Is intonation susceptible to the asymmetrical language switch cost?: A theoretical exploration

7. **Chikako Takahashi (Stony Brook University)**, Hyunah Baek, Sophia Kao & Alex HL Yeung: Processing and production of English focus prosody by native English vs. Mandarin speakers

8. **Thanh Lan Truong (University of Tuebingen)** & Yuki Asano: Udon Know Me: HL or LH? Pitch accent in spoken word recognition

9. **Sin Kwan Yuen (University of Tuebingen)** & Yuki Asano: Processing tonal contrast between native Chinese and English listeners

Social Event

Friday Evening

Stocherkahnfahren (only with prior registration)

Meeting point: Evangelisches Stift, Klosterberg 2,
72070 Tübingen

ANNE CUTLER

Anne Cutler studied languages and psychology at the Universities of Melbourne, Berlin and Bonn, taught German at Monash University, but embraced psycholinguistics as soon as it emerged as an independent sub-discipline, taking a PhD in the subject at the University of Texas. Postdoctoral fellowships at MIT and Sussex University followed, and from 1982 to 1993 a staff position at the Medical Research Council Applied Psychology Unit in Cambridge. In 1993, she became a director at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands, a post she held till 2013. She was also professor of comparative psycholinguistics at the Radboud University Nijmegen from 1995 to 2013, and, from 2006 to 2013, part-time Research Professor in MARCS Auditory Laboratories. In 2013, she took up a full-time position at the MARCS Institute.

SARAH BIBYK

Sarah Bibyk received a joint doctoral degree in Brain and Cognitive Sciences and Linguistics at the University of Rochester in 2016. Now, she is a postdoctoral researcher at Vanderbilt University, Peabody College, in the department of Psychology and Human Development. She works in the Communication and Language Lab, and her research interests include intonation, language processing, and language production.

KIWAKO ITO

Kiwako Ito received her PhD 2002 in Linguistics from the University of Illinois at Urbana-Champaign, advised by Jennifer S. Cole. Kiwako currently is a Senior Researcher at the Department of Linguistics, Ohio State University. In her research, she applies psycholinguistic methods to investigate acquisition and online use of prosody in children, L2 learners and individuals with developmental disorder.

GIUSEPPINA TURCO

Giuseppina Turco is a post-doctoral researcher at Laboratoire de Phonétique et Phonologie in Paris. She completed her PhD-thesis at the Max-Planck Institute for Psycholinguistics in Nijmegen in 2013. One of her current projects investigates non-local anticipatory effects from word medial geminates across three unrelated languages – Italian, Tashlhiyt Berber and Japanese.

Invited Speakers' Abstracts

INVITED SPEAKER

Anne Cutler

ABSTRACT

What is special about prosody in processing?

Prosodic processing differs across languages in some very important ways. For instance, the way listeners segment continuous speech signals varies as a function of the prosodic feature rhythm. Language learners are sensitive to this feature very early – even before birth, in fact. Moreover, the function of segmenting speech is a vital part of language development, and evidence that the ability to segment speech has been achieved is a significant predictor of language skills later in childhood. Thus prosodic processing (and language-specific prosodic processing at that) is one of the most crucial steps in our early life. Though prosodic processing continues to be important in language use, and is language-specific in interesting ways (at the word level, the sentence level and potentially more), there is surprisingly little targeted cross-language research in this area (though research on prosody itself, and on within-language prosodic processing, both seem to be growing). The most likely underlying reason is the relative lack of awareness of prosody in general, and it is time for us to consider whether there is something that can be done to rectify this, in particular by making prosodic knowledge more accessible both to language scientists of any kind, and to the general public.

INVITED SPEAKER

Sarah Bibyk

ABSTRACT

Listeners interpret rising and falling intonation prior to the final boundary

It is still not well understood how full intonational contours are processed and integrated during online sentence comprehension. Though there is a body of work investigating pitch accents specifically, comparatively less work has been dedicated to the processing of other parts of the contour, like boundary tones. Boundary tones are hypothesized however to be crucial in signaling various kinds of higher-level meanings, such as distinguishing questions from statements. We investigated how listeners process such contours in the context of a “targeted language game.” Participants played a card game where on critical trials the speaker produced an utterance that was structurally ambiguous (e.g. “Got an armadillo”) but was distinguished as a question or a statement by virtue of the intonational contour. Using eyetracking, we asked when in the contour do listeners interpret the utterance as a question or statement: it is only after they have heard the full contour (i.e. post boundary tone) or do they rely on earlier cues to distinguish the contours? Though an initial experiment suggested listeners need to wait until hearing the full contour, later experiments with more tightly controlled stimuli suggested that listeners will use acoustic information prior to the end of the boundary tone.

INVITED SPEAKER

Kiwako Ito

ABSTRACT

Investigating development of linguistic and non-linguistic prosody

In this seminar, I will talk about how to deal with the interaction between linguistic prosody and affect prosody, which has been traditionally considered as a non-linguistic component of communication. I will discuss how investigations of linguistic prosody and those of non-linguistic prosody have been separated theoretically and methodologically, and will propose that we should consider them simultaneously for understanding face-to-face daily oral communication. While the myth that emotion recognition is fundamental and easy and should precede the acquisition of language-specific prosody seems to have led the separation of research fields, affect processing and linguistic processing are equally complex and impact each other during communication. Potential experimental questions, tasks, measures and the associated challenges for future studies will be discussed.

INVITED SPEAKER

Giuseppina Turco

ABSTRACT

Testing for processing advantages of linguistic effects of focus in L1 and L2

Listeners exploit prior sematic context to determine focussed information but not prosodic context for the same purpose. A phoneme-detection task (Experiment 1) shows this finding in Sepedi, a language with no grammmaticalized prosodic expression of focus. Sepedi listeners detected phoneme targets more quickly when the phoneme-bearing words were focussed (than unfocussed) but not when these words were emphatic (vs non-emphatic). The lack of interaction between the two effects suggests that focus and prosodic emphasis do not share the same function (searching for focus) in Sepedi. Experiment 2 tested cross-linguistic differences on the processing of Sepedi learners of English (a language with grammaticalized focus-to-accent mapping). Non-natives detected phoneme-bearing words more quickly in focussed condition and in accented condition. Like in Sepedi, no interaction was found. The study suggests that non-natives acquire the L2 prosodic structure (even if it is very different from their L1) while they remain unaware of its underlying discourse representation.

Participants' Abstracts

Prosodic Phrasing in European Portuguese varieties

Nádia Barros, Sónia Frota
Center of Linguistics, University of Lisbon, Portugal
nadia.barros@campus.ul.pt, sonia.frota@mail.telepac.pt

Studies on prosodic and intonational phrasing across Romance languages and its varieties have shown variation, namely different phrasing patterns in sentences with Subject (S), Verb (V), and Object (O) structure, as well as differences in the syntactic and prosodic factors that affect these patterns (Cruz, 2013; D'Imperio et al., 2005; Elordieta et al., 2005; Frota, 2000, 2014; among others). Building on previous work on prosodic and intonational phrasing in European Portuguese (EP) (Cruz, 2013; Frota, 2000, 2014; Frota & Vigário, 2007; Vigário & Frota, 2003), as well as comparative studies between Portuguese and other Romance languages (D'Imperio et al., 2005; Elordieta et al., 2005; among others), the main goal of this study is to characterize Northern and Central-Southern varieties of European Portuguese, which have not been studied for prosodic phrasing.

The current study is part of a PhD project (with a grant funded by the Portuguese Science Foundation - FCT) that is being developed within the Interactive Atlas of the Prosody of Portuguese (InAPoP) project. The empirical database comprises data from five regions: Porto (Por) and Braga (Bra), two Northern varieties; Coimbra (Cob), Castelo Branco (CtB) and Évora (Eva), Central-Southern varieties. Data were collected in loco using InAPoP's methodology. A reading corpus of 76 neutral declarative sentences, controlled for syntactic and prosodic complexity (branching/non-branching constituents), as well as constituent length in number of syllables, was previously designed and used to study prosodic phrasing in Romance languages/varieties (Cruz, 2013; D'Imperio et al., 2005; Elordieta et al., 2005; Frota, 2000, 2014; Frota & Vigário, 2007; Vigário & Frota, 2003). The corpus was produced by female monolingual native speakers of each region aged between 20 and 45 years old. For the current analysis, two speakers for each region (2x5), and at least two renditions per speaker were considered. A prosodic and intonational analysis was made within the Prosodic Phonology and the Autosegmental Metrical approach to Intonational Phonology (Nespor & Vogel, 2007; Ladd, 2008; Frota, 2000, 2014; among others). Data were analysed in *Praat* (Boersma & Weenink, 2013), where nuclear contours and prosodic breaks were annotated according to P-ToBI (Frota et al., 2015), in a total of 1520 sentences.

The results confirm previous findings for EP: constituent branchingness and length have a different weight in the prosodic phrasing of SVO sentences depending on the variety, with the (S)(VO) phrasing pattern as the most frequent for the Northern and Central-Southern varieties (Figure 1), and (SVO) the most frequent pattern for Standard European Portuguese (spoken in Lisbon) and the Southern variety of Algarve (Cruz, 2013; Frota, 2000, 2014; Frota & Vigário, 2007; Vigário & Frota, 2003). By contrast, EP varieties do not seem to differ in the type of boundary cues used, with inner intonational breaks being usually marked by a high boundary tone (H%), with or without a pause, and the final IP edge showing the nuclear tonal configuration H+L* L%, the most frequent one for neutral declaratives (Frota, 2000, 2014).

Further investigation is in progress and will allow, by the end of the project, to identify syntactic and phonological constraints, as well as socio-phonetic aspects, which may impact on prosodic phrasing in other regions, contributing to deepen our knowledge of prosodic phrasing variation in EP and of the typology of prosodic phrasing patterns within and across Romance languages.

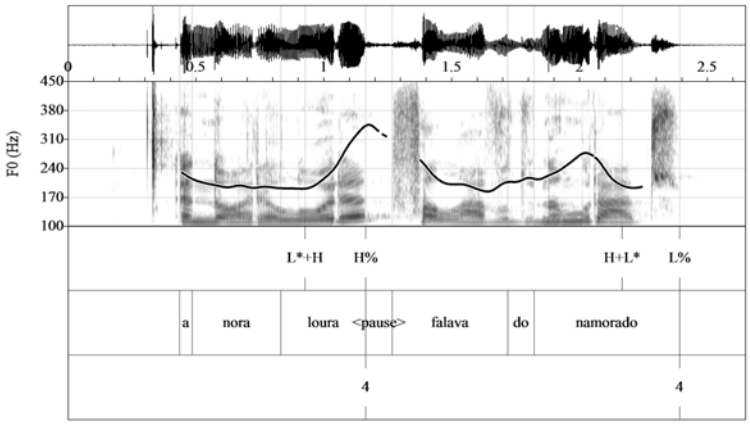


Figure 1. Intonational contour of the sentence *A nora loura falava do namorado* 'The blond girl spoke about the boyfriend', (S)(VO) pattern, produced by a speaker from Castelo Branco.

References

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* (Version 5.3.41) [Computer software]. [http://www.fon.hum.uva.nl/praat/]

Cruz, M. (2013). *Prosodic variation in European Portuguese: Phrasing, intonation and rhythm in* central-southern varieties (Doctoral dissertation), University of Lisbon, Lisbon, Portugal.

D'Imperio, M., Elordieta, G., Frota, S., Prieto, P., & Vigário, M. (2005). Intonational phrasing in Romance: The role of syntactic and prosodic structure. In S. Frota, M. Vigário, & M. J. Freitas (Eds.), *Prosodies* (pp. 59–97). Berlin/New York: Mouton de Gruyter.

Elordieta, G., Frota, S., & Vigário, M. (2005). Subjects, objects and intonational phrasing in Spanish and Portuguese. Studia Linguistica, 59 (23), 110-143.

Frota, S. (2000). *Prosody and focus in European Portuguese: Phonological phrasing and intonation.* New York: Garland Publishing.

Frota, S. (2014). The intonational phonology of European Portuguese. In S.A. Jun (Ed.), *Prosodic Typology II* (pp. 6-42). Oxford: Oxford University Press.

Frota, S. (Coord.) (2012-2015). *InAPoP Interactive Atlas of the Prosody of Portuguese* (Funded by Fundação para a Ciência e a Tecnologia FCT, PTDC/CLELIN/119787/2010) [http://labfon.letras.ulisboa.pt/InAPoP/].

Frota, S., & Vigário, M. (2007). Intonational phrasing in two varieties of European Portuguese. In T. Riad, & C. Gussenhoven (Eds.), *Tones and Tunes* (Vol. 1, pp. 265-291). Berlin: Mouton de Gruyter.

Frota, S., Oliveira, P., Cruz, M., & Vigário, M. (2015). *PToBI: Tools for the transcription of Portuguese prosody*. Lisboa: Laboratório de Fonética, CLUL/FLUL. [http://labfon.letras.ulisboa.pt/InAPoP/PToBI/]

Ladd, R. (2008). *Intonational Phonology* (2nd ed.). Cambridge: Cambridge University Press.
Nespor, M., & Vogel, I. (2007). *Prosodic phonology: With a new foreword.* Berlin/New York: Mouton de Gruyter.

Vigário, M., & Frota, S. (2003). The intonation of Standard and Northern European Portuguese. *Journal of Portuguese Linguistics*, *2* (2), 115-137.

Effect of Language Background on L2 Focus Perception

L. Ann Burchfield[1], Heather Kember[1,2]
[1]*The MARCS Institute, Western Sydney University, Australia*
[2]*The ARC Centre of Excellence for the Dynamics of Language, Australia*

Introduction

Listeners benefit from prosody in processing their native language (see Cutler, Dahan & Donselaar, 1997 for a review). Far less is known about L2 listeners' ability to use prosody, but evidence suggests that L2 listeners do not receive the same processing benefits afforded to native speakers. For example, English L2 speakers do not show the same benefits as English L1 speakers from prosodic focus either in processing speed in phoneme-monitoring (Akker & Cutler, 2003), or in recognition memory performance (Lee & Fraundorf, 2016). In the present study we ask if L2 listeners are even aware of English prosodic focus, and notice when it is incorrect for the context. We test 3 groups: L1 English listeners, L1 Mandarin listeners, and L1 Korean listeners.

Prosodic focus in Korean is achieved through accentual phrasing (Jun, 1998) while Mandarin can mark focus by increasing pitch range and duration of the focused word (Xu, 1999). While focus marking must co-occur with lexical tone in Mandarin, it still shares many of the same features as English, while Korean does not. Therefore, we predict that L1 Mandarin listeners will outperform L1 Korean listeners in perception of English focus.

Method

Stimuli included question-answer pairs recorded by native Australian English speakers. Each response was produced with two different focus placements and preceded by one of two possible questions. This resulted in four versions of each dialogue, two with correct focus placement (focusing new information) and two with incorrect focus placement (focusing given information). See Table 1 for an example. Twenty of these dialogues were constructed, resulting in 80 questions-answer pairs.

Participants will include 20 speakers from each language group. At present, data has been collected for 17 English speakers, 25 Mandarin speakers and 15 Korean speakers. Each participant heard 40 experimental sentences (counterbalanced across participants) as well as 40 filler items. Half of filler items were semantically odd and half were semantically natural. After each item, participants were asked to rate the appropriateness of the response (1-7 scale) and whether they would have answered the question differently ("yes" or "no"). These questions allowed us to probe whether listeners were aware of focus without the need for metalinguistic knowledge and vocabulary.

Results and discussion

Results for acceptability ratings and proportion of "yes" responses to the question "Would you have answered this question differently?" are given for each language group in Figures 1 and 2. L1 English, Mandarin Chinese and Korean participants all responded that they would have answered differently more often for incorrect versus correct focus placement, although the difference was larger for the L1 English group. L1 English and Mandarin Chinese participants gave lower acceptability ratings for items with incorrect focus placement versus correct focus placement, while Korean participants rated items with incorrect versus correct focus placement as roughly equivalent. These results suggest that while L2 English speakers are aware of prosodic focus, their L1 may affect their ability to perceive these differences.

Table 1: Example stimuli

Question	Correct Focus Placement	Incorrect Focus Placement
Who prevented the shelter from burning down?	The FIREMEN prevented the shelter from burning down.	The firemen prevented the SHELTER from burning down.
What did the firemen prevent from burning down?	The firemen prevented the SHELTER from burning down.	The FIREMEN prevented the shelter from burning down.

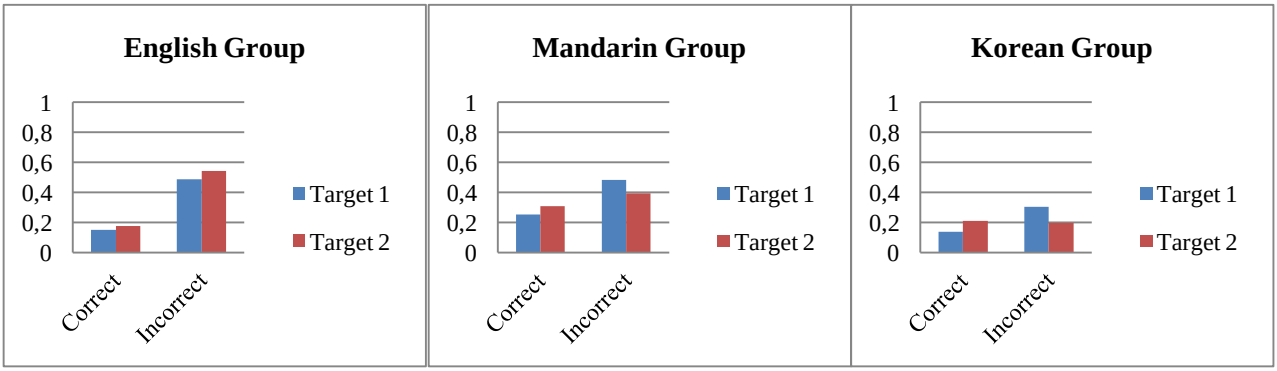


Figure 1: Proportion "Yes" responses to the question "Would you have answered this question differently?" for the L1 English Group (left) and L1 Mandarin group (right). Target 1 indicates focus earlier in the sentence (e.g. firemen) and Target 2 indicates focus later in the sentence (e.g. shelter).

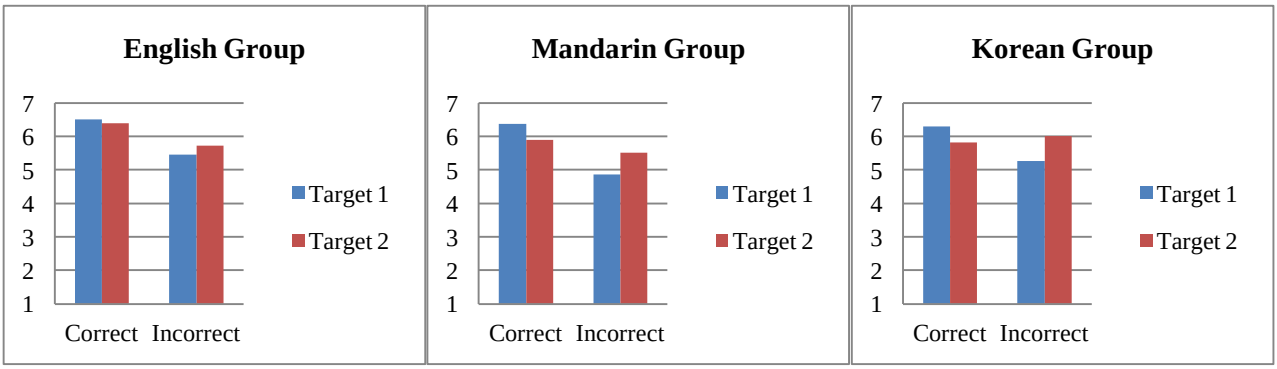


Figure 2: Mean appropriateness ratings for the L1 English group (left) and L1 Mandarin group (right). Target 1 indicates focus earlier in the sentence (e.g. firemen) and Target 2 indicates focus later in the sentence (e.g. shelter)

References

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and non-native listening. *Bilingualism: Language and Cognition*, *6*(02), 81-96.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. Language and Speech, 40(2), 141-201.

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*, 189-226.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. Journal of Phonetics, 27(1), 55-105.

Comprehension of prosodic and syntactic focus marking in Mandarin Chinese- Data from children and adults

Hui-Ching Chen[123], Stephen Crain[2], Barbara Höhle[1]

1. *Department of Linguistics, University of Potsdam, Potsdam, Germany*
2. *Department of Cognitive Science, Macquarie University, Sydney, Australia*
3. *International Doctorate for Experimental Approaches to Language and Brain*

The questions how information structure (IS) affects the structure of verbal utterances and how listeners interpret the cues related to IS have been explored for decades but it still remains unclear how children implement IS to understand focus crosslinguistically. Focus indicates the presence of alternatives in the current discourse (Rooth, 1992) and it can be marked by syntactic or prosodic means. In Mandarin - as a tone language - the relevance of prosody to mark IS is questionable. Previous research indicates that Mandarin speakers (adults and children) apply acoustic parameters i.e. fundamental frequency, duration and amplitude to mark IS in sentence production (Ouyang and Kaiser, 2014; Yang and Chen, 2014). Most interestingly, Chen found that Mandarin-speaking children relied more heavily on stress than adults to identify focus in a sentence-picture verification task. She assumed that children switch their strategy from a prosodic one to a structural one based on word order during language acquisition (Chen, 1998).

In this study, a newly designed task- a sentence-picture verification task- adjusted from Szendrői et al. study (2017) was employed to investigate how Mandarin-speaking children understand focus structures, i.e. the cleft construction and the canonical sentence with an element bearing focal stress (Examples 1-3). The pictures were designed so that participants could correct either the Subject NP or the Object NP, depending on the assignment of the focus element. During the task, not only the verbal responses were recorded but also their eye movements were collected by the Tobii eye-tracker TX300. There were three conditions, one prosody condition (Subject-accented), and two syntax conditions (Cleft-Subject and Pseudocleft-Object). 52 5-year-old Mandarin-speaking children and 56 Mandarin-speaking adults were tested. The results showed that both Mandarin-speaking adults and children were not sensitive to prosodic information but relied on syntactic information to identify the focus (Fig. 1). More, the conclusions of the eye-tracking results are similar to the ones of the response data (Fig. 2). The findings of the present study are in striking contrast to those of Chen (1998), suggesting that children learning Mandarin Chinese have an adult-like stronger weighting of syntactic over prosodic cues as markers of focus from early on.

Examples:

1. Subject-accented
   XIAONIAO you shueping, shi ma?
   bird have bottle Aux Q
   'The BIRDYF has the bottle, is that right?'
2. Subject-cleft
   Shi XIAONIAO you shueping, shi ma?
   SHI bird have bottle Aux Q
   'It is the BIRDYF who has the bottle, is that right?'
3. Object-pseudocleft
   Xiaoniao you de shi SHUEPING, shi ma?
   bird have DE SHI bottle Aux Q
   'What the birdy has is the BOTTLEF, is that right?'

Fig. 1

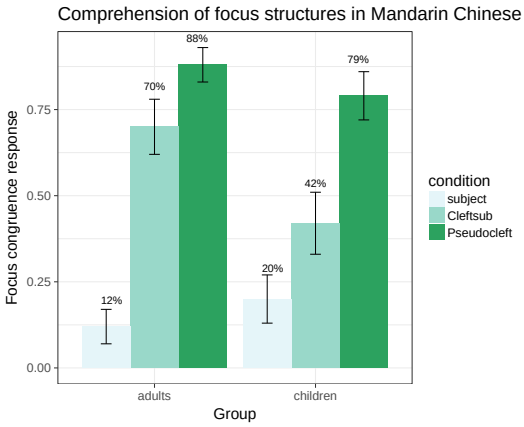
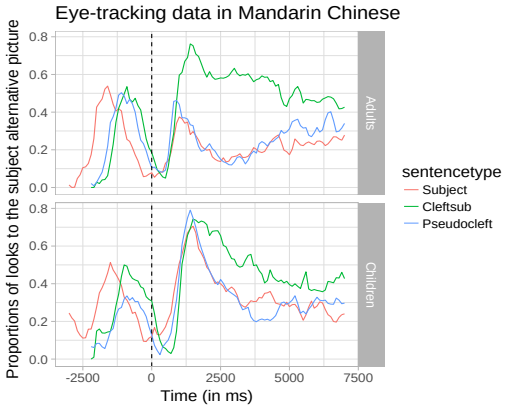


Fig. 2



References

Rooth, M. (1992). A theory of focus interpretation. Natural language semantics,1(1), 75-116.

Yang, A., & Chen, A. (2014). Prosodic focus marking in Chinese four-and eight-year-olds. In Speech prosody (Vol. 7, pp. 713-717).

Chen, S. H. E. (1998). Surface cues and the development of given/new interpretation. Applied Psycholinguistics, 19(04), 553-582.

Ouyang, I. C., & Kaiser, E. (2014). Prosody and information structure in a tone language: an investigation of Mandarin Chinese. Language Cognition and Neuroscience, 30(1-2), 57-72. doi:10.1080/01690965.2013.805795

Szendrői, K., Bernard, C., Berger, F., Gervain, J., & Höhle, B. (2017). Acquisition of prosodic focus marking by English, French, and German three-, four-, five-and six-year-olds. Journal of child language, 1-23.

Are French learners… "stress deaf"?
A comparative study between intermediate and advanced French learners of Spanish.

Syrine Daoussi ; Lorraine Baqué Millet ; Marta Estrada Medina
Syrine.Daoussi@uab.cat; Lorraine.Baque@uab.cat; Marta.Estrada@uab.cat
Universitat Autònoma de Barcelona

Although there is a growing body of research on the role of prosody in oral comprehension in L2, it is still a subject of debate (Cutler, 2012). The two languages considered here (Spanish and French) are classified as "syllable-timed" languages but differ regarding their stress pattern. Indeed, Spanish stress is free, distinctive and encoded at the lexical level of processing, whereas in French stress is fixed, not-distinctive and generally considered to be computed at the post-lexical level. A hypothesis has been advanced that these differences may lead to a persistent "stress deafness" for French learners (Dupoux, E., Peperkamp, S., 2001; Dupoux, E.; Sebastián-Gallés, N.; Navarrete, 2008; a contrario Astésano, C., Bertrand, R., Espesser, R., & Nguyen, 2012; Muñoz, M., Panissal, N., Billières, M., & Baqué, 2009). The aim of this preliminary study is to see whether "persistent stress deafness" (Dupoux, E.; Sebastián-Gallés, N.; Navarrete, 2008) equally appears between advanced French learners of Spanish in immersion and intermediate learners while detecting incoherencies through an oral comprehension task. Our hypothesis stands that several factors may have an influence on perception of accentual contrasts as participants' mother tongue, their Spanish level, error type (morphologic vs vocalic) and items' complexity. Participants who took part in this study were on the one hand 38 French students of Spanish as foreign language (aged between 19 and 25) with an intermediate level of Spanish (b1-b2) and on the other hand 20 French learners (aged between 30 and 45) who are living in an immersion context in Spain and are expected to show higher linguistic abilities (c1-c2). We also counted with 20 Spanish speakers as control participants. All of them were asked to listen and assess both grammatical and semantic linguistic acceptability of items at three complexity levels. They were randomly presented several items in Spanish, such as isolated words (lavo), sentences: (Ese día, Sara lavó la taza) and short texts built with a couple of coherent sentences: (Yo, cada mañana, lavo la taza, miro la tele y me voy a trabajar). The obtained data were analysed by means of linear mixed-effects regression models, in which participants were introduced as random variables, mother tongue and type of error (segmental and stress related, lexical or morphological), and items' complexity as independent factors. The dependent variables were the Signal Detection Theory measures of sensibility (loglinear A') and of response bias (loglinear B'') (Stanislaw & Todorov, 1999). Results show that the French learners' performance is poorer than natives in all tasks, especially regarding stress versus vocalic errors. For both groups of foreign learners, errors are more difficult to identify in shorts texts compared to sentences or isolated words. If error detection is more difficult in texts for French intermediate learners, they showed no sensitivity differences regarding their position (first or second verb) while we observed that advanced learners' sensibility decreases when errors fall on the second verb. However, French living in an immersion context detected better errors with a paroxytone pattern. The native controls Spanish learners reacted better to oxytone pattern, i.e the less common pattern in their language.

key words: prosody- French- Spanish- stress deafness- oral comprehension

References

Astésano, C., Bertrand, R., Espesser, R., & Nguyen, N. (2012). Perception des frontières et des proéminences en français. In *Actes de la conférence conjointe JEP-TALN-RECITAL* (pp. 353–360).

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge: MA: MIT Press. http://doi.org/10.1353/lan.2014.0014

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

Dupoux, E.; Sebastián-Gallés, N.; Navarrete, E. . P. S. (2008). Persistent stress "deafness": the case of French learners of Spanish. *Cognition*, *106*(2), 682–706.

Muñoz, M., Panissal, N., Billières, M., & Baqué, L. (2009). ¿La metáfora de la criba fonológica se puede aplicar a la percepción del acento léxico español? Estudio experimental con estudiantes francófonos. In C. L. Bretones & E. Al. (Ed.), *La lingüística aplicada actual: Comprendiendo el lenguaje y la mente*. (pp. 489–500). Universidad de Almería, AESLA.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. Retrieved from http://people.brandeis.edu/~sekuler/stanislawTodorov1999.pdf

Comprehensibility in L2 speech: lexical stress versus sentence accent

Antje Hey, Katharina Nimz & Petra Wagner
Faculty of Linguistics and Literary Studies
Bielefeld University

This study investigates which prosodic features in L2 speech effect comprehensibility the most and should be prioritized in pronunciation training. Two types of prosodic features were investigated: lexical stress and sentence accent. Both contribute to the comprehensibility of a word, phrase, or sentence. Because of the specific role of lexical stress in word retrieval (Cutler 2005, Field 2005, Van Donselaar et al. 2005, Aitchison 2012) it was hypothesized that incorrect lexical stress has worse consequences for comprehensibility than incorrect sentence accent. It was further hypothesized that sentences with both correct lexical stress and sentence accent are easier to understand than sentences that contain incorrect lexical stress and/or incorrect sentence accent. Finally, it was hypothesized that sentences with both incorrect lexical stress and sentence accent are more difficult to understand than sentences that exhibit just one of the two types of mistakes. In this study, 'incorrect sentence accent' refers to a mismatch between the anticipated and realized placement of prosodic focus given a short discourse context. Such contexts preceded the acoustically presented stimuli assessed by native listeners in a rating task. For both lexical stress and sentence accent, only the presence or absence of prosodic prominence is investigated, not its specific acoustic realization.

The hypotheses were tested in an online survey. Seventy-eight native speakers of German rated 12 sentences on a 5-point rating scale according to how well they had understood what the speaker had said. The stimulus material was based on recordings of a native Italian speaker from Central Italy with good knowledge of German, i.e. B2 level and additional pronunciation training. Each sentence was recorded in four prosodically different versions. Version 1 did not contain any categorical prosodic mistakes. Version 2 contained both incorrect lexical stress and incorrect sentence accent. Version 3 contained correct lexical stress and incorrect sentence. Version 4 contained incorrect lexical stress and correct sentence accent. Six of the words tested for correct or incorrect lexical stress were nominal compounds (first constituent stressed correctly and second constituent stressed incorrectly (e.g. **Koch**buch/*Koch**buch**)) and the other six words were adjectives preceded by the negation particle *un-* (e.g. **un**freundlich/*unf**reund**lich). The first part of each sentence contained the correctly or incorrectly stressed compound or adjective and the second half always contained the correctly or incorrectly placed sentence accent. In order to ensure that the four versions differed only in location of prominence realization, the recordings were edited using a splicing method. For each element of a sentence (the part of the sentence that included either the correct or incorrect lexical stress or the correct/incorrect sentence accent) the best candidate (in terms of overall pronunciation quality) out of at least 3 recordings was selected and then combined with the best candidate of the other part. This way, the first part of a sentence was identical in two of the four different versions of a sentence: The first half a sentence with the correctly stressed compound or adjective, for example, was used in the completely correct version (version 1) and the version with correct lexical stress and incorrect sentence accent (version 3). To avoid learning effects, we used a block design where each participant heard and rated only one of the four possible versions of a sentence; with randomized orders for each participant. Each sentence was embedded in a context that specified a certain constituent as the exponent of prosodic focus, i.e. the correct location of sentence accent. A Kruskal-Wallis test confirmed a significant difference between the four "error" groups *($H=33.5$, $df = 3$, $p<0.001$)*, and posthoc Dunn tests revealed that the hypotheses could partially be corroborated. Sentences that did not contain any prosodic errors interfered less with comprehension than versions with errors (*p-values: 1.5e-07, 0.03262,*

0.00032). Sentences with correct lexical stress were rated significantly better than sentences with both incorrect lexical stress and incorrect sentence accent *(p=0.03154)*.

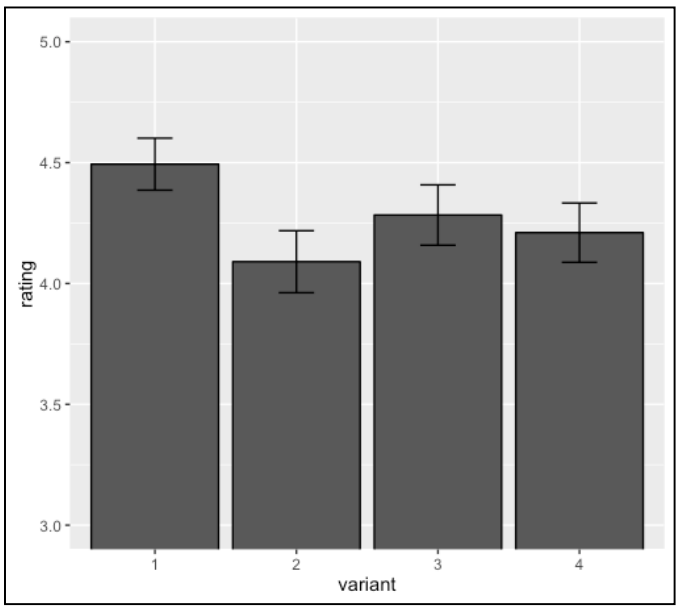


Figure 1: Mean ratings and error bars for the 4 prosodic error versions (1 =all correct, 2 = all incorrect, 3 = correct lexical stress only, 4 = correct sentence accent only)

Sentences with correct sentence accent (and incorrect lexical stress) however were not rated better than sentences with both types of errors (see Figure 1). It seems that, in this study, correct lexical stress facilitated comprehension, but sentence accent did not. The hypothesis that sentences with both types of errors are generally rated worse than all other versions could not be confirmed as there is no significant difference between versions 2 and 4. Also, gender seems to be a factor in this study that contributes to the rating of comprehensibility. Women rated sentences overall better than men *(p=3.784e-05)*, but this could be an effect of the different group sizes (22men, 56 women).

Further research could investigate lexical stress errors in different types of words and whether the acoustic means of realizing lexical stress play a role in comprehensibility ratings. From a cross-linguistic point of view, it would be interesting to investigate other language pairs with the same paradigm.

References

Aitchison, J. (2012). *Words in the Mind: An Introduction to the Mental Lexicon* (4th ed.). Pondicherry: Wiley-Blackwell.

Cutler, A. (2005). Lexical Stress. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264–289). Oxford u.a.: Wiley-Blackwell.

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly, 39*(3), 399–423.

van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology, 58*(2), 251–273.

Do prosodic differences between English and Japanese disadvantage Japanese learners of English?

Alastair Graham-Marr
gmarr@rs.kagu.tus.ac.jp
Department of Liberal Arts, Tokyo University of Science

This poster presents ongoing research into some of the difficulties that Japanese learners of English face when trying to comprehend naturally spoken English (Graham-Marr, 2015, 2017). Given that many of the prosodic features of English are not often taught in the classroom, many Japanese learners of English struggle to understand naturally spoken English. The prosodic features of English, its sounds, rhythms, volume and tempo, differ considerably from many Asian languages. Many Asian languages are syllable timed, while English is a stress-timed language. In English, vowel length routinely varies in order to maintain equal intervals between two stressed syllables. However, such vowel length variance is absent from many Asian languages. Moreover Japanese, being a mora-timed language is even one phonological step further removed from the prosody of stress-timed languages.

Although the exact definition of a mora is still the subject of some debate, for our purposes, a mora can simply be defined as a sound unit. While a syllable is a unit of sound having one vowel, in Japanese a syllable can have 1 or 2 moras. For example, the word Tokyo has 2 syllables, to-kyo. However, it has four mora, in that, both syllables have extended vowels. In Japanese Tokyo is written as (とうきょう), to-u-kyo-u. Mora-timed languages tend to have an even rhythm, that is the time interval between each mora is roughly equal. So, while Japanese has both long vowels and short vowels, differences in vowel length result in different words being rendered. That is, vowel length changes meaning. For example, the words strengthen and permission have a similar pronunciation in Japanese. The difference is in vowel length. Strengthen is pronounced kyouka with an extended initial vowel, and permission is pronounced kyoka with a short initial vowel. Both words have two syllables. However kyouka has a long vowel and therefore has 3 moras, while kyoka has a short vowel and therefore only 2 moras. So, the time needed to enunciate the word kyouka is roughly one third longer than the time needed to enunciate the second word kyoka, as each mora is usually given more or less equal weight.

This poster reports on many of the difficulties that Japanese learners have to cope with when trying to understand naturally spoken English and presents student reactions to a pedagogical approach that explicitly point out many of the more common prosodic features of English. The findings of these studies are in accord with other studies done in Asian contexts which have found a positive effect for explicitly teaching prosodic features, (see for example, Wong, 2017; Goh and Taib, 2006). However, while students appreciated having prosodic features pointed out, and understood the mechanisms, they nonetheless continued to struggle with comprehension tasks, suggesting that students need many long hours of practice before they are able extemporaneously process and comprehend naturally spoken English. That is knowledge alone is insufficient.

References

Goh, C. C. & Taib, Y. (2006). Metacognitive instruction in listening for young learners. ELT Journal, 60(3), 222-232

Graham-Marr, A. (2017). Impediments to EMI and ESP in an Asian context. The Journal of Teaching English for Specific and Academic Purposes, 5 (2), 285-294 ISSN 2334-9182

Graham-Marr, A. (2015). Elevating the development of listening skills to foster SLA in an Asian context. Journal of Foreign Language Teaching and Applied Linguistics, 2 (2). ISSN 2303-5528

Wong, S., et al. (2017). Perception of Native English Reduced Forms in Chinese Learners: Its Role in Listening Comprehension and Its Phonological Correlates. TESOL Quarterly, Vol. 51, No. 1, pp 7-31

Is intonation susceptible to the asymmetrical language switch cost? A theoretical exploration

Sergio Quiroz
Berlin School of Mind and Brain
Humboldt Universitat zu Berlin

When a bilingual speaker uses one language, the non-target language is not switched-off: it remains active (e.g. Thierry & Wu 2007). To prevent interference, therefore, the non-target language is inhibited across linguistic features such as syntax and phonology (Meuter & Allport 1999). I argue here that inhibition is also applied to intonational aspects, something not yet shown through this paradigm. As such, this paper focuses on theoretical and methodological principles at the crux of intonation, multilingualism and cognitive functions by reviewing existing literature. Additionally, I present original data from monolingual German speakers and compare it to previous published data from monolingual Polish speakers (Arvaniti et al. 2017) in order to build a construct under which to measure the effects of language switching on intonation.

Meuter and Allport (1999) note that an asymmetrical switch cost goes into effect when we switch between a less habitually used language (herein the “secondary language” or “L2”) and a more highly used language (“primary language” or “L1”). In particular, they find that going from the L2 to the L1 results in longer naming latencies of stimuli than when going from L1 to L2. This asymmetrical switch cost has also been seen at the segmental level with regards to the voice onset timing (VOT) associated with plosive consonants. Olson (Olson 2013) observed that L1-Enlglish speakers produced shorter English-VOTs when going from their L2-Spanish to English (in comparison to not switching: L1-English to L1-English), while L1-Spanish speakers produced longer Spanish-VOTs (when compared to not switching). A comparable effect has not been shown for intonation. One possibility for this is the lack of a construct which could show such asymmetrical cost switch effects.

To this end, I focus on the intonational phonology of the vocative chant and look in particular into the boundary tones (represented as % in the Autosegmental Metrical Phonology notation). The vocative chant is best conjured by imaging an adult sweetly calling out to a child for dinner. It is said to be produced similarly across many European languages (Ladd 2008): a high tone followed by a second high, but reduced, tone with the second tone occurring as a plateau. While this is certainly true of German phonologically (Grice et al. 2005) as well as phonetically (Niehbur 2015, Quiroz and Zygis forthcoming), recent acoustic evidence shows that Polish does not plateau at the second high tone, instead it rises (Arvaniti et al. 2017). Phonologically the calls are represented as (L+)H* !H-% for German (Grice) and as L+H* !H-H% for Polish (Arvaniti et al. 2017).

I identify in theory that this difference in the boundary tones with respect to F0 pitch height serves as a good candidate for a construct to measure parallel effects to those seen in the VOT. The hypothesis is that an L1-Polish speaker would produce a lower boundary tone when switching (German-Polish) vs. not-switching (Polish-Polish) and an L1-German speaker would produce higher boundary tones when switching (Polish-German) vs. not-switching (German-German).

References

Arvaniti, A., Żygis, M., & Jaskuła, M. (2017). The phonetics and phonology of the Polish calling melodies. Phonetica, 73(3-4), 338-361.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German Intonation in Autosegmental Metrical Phonology. In: Jun, S-A. (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, 55-83. Oxford: Oxford University Press.

Meuter, R.F.I., Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. J. Mem. Lang. 40, 25–40.

Niebuhr, O. (2015). Stepped intonation contours: A new field of complexity. In: Skarnitzl, R. & Niebuhr, O. (Eds.) *Tackling the complexity of speech*. Prague: Charles University Press.

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, *41*(6), 407-420.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30): 12530 12535.

The study of prosodic grouping by Chinese as a second language learners

Sichang Gao
Shanghai International Studies University

This paper investigated how Korean learners of Chinese incorporate prosodic specifications such as boundaries and breaks to group or chunk the language content into prosodic units, under the effect of syntax and semantics. Thirty participants who were divided into two different language levels were included in the research and fifteen native speakers' performances were recorded as the baseline. Participants' speech production was elicited by sentence memorization tasks and the sentences were transcribed and analyzed by Praat (5.4).

This study consisted of three experiments: (1) The first experiment compared the prosodic grouping characteristics between learners of Chinese and the native speakers on the aspects of boundary production rate, boundary occurence, boundary strength, and the length of prosodic unit. The results showed that the learners of Chinese had demonstrated significant less competence on prosodic grouping abilities than the native speakers. Moreover, different learners' groups did not manifest significant differences, a finding demonstrated that the length of learning had not contributed to the improvement of Chinese prosodic grouping abilities; A second result showed that Chinese native speakers were inclined to pause more when the subject and the object are comparative longer than the predicate. But the pause rate has no significant difference between the length of subject and the object. The higher-level learners have demonstrated the same pattern under the same circumstance. (2) The second experiment looked at the influence that syntactic complexity, syntactic constituencies and the constituent length had on the boundary occurrence and boundary strength in learner's sentence production, while the influence on the native speaker is little; (3) The third experiment tried to discover the effect that semantic prosody had on language learners' prosodic groupings. However, no significant effect was found on both language learners and native speakers. No matter how proficient the speakers were, they did not show the phonological coherence on the semantic preferred collocations.

Current results demonstrated inconsistencies between syntactic and semantic influences on different levels' learners, which surmises that some features of prosodic groupings can be incrementally promoted with increased learning time, but some features such as semantic phonological coherence did not shown the same effect.

Keywords: *prosodic grouping*, *speech production*, *Chinese as a second language*

References

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, *100*(2), 233.

Frazier, L., Clifton, C., & Carlson, K. (2004). Don't break, or do: prosodic boundary preferences. *Lingua*, *114*(1), 3-27.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive psychology*, *15*(4), 411-458.

Grosjean, F., & Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, *36*(2), 98-114.

Jun, S. A. (2003). Prosodic phrasing and attachment preferences. *Journal of psycholinguistic research*, *32*(2), 219-249.

Keating, P. A. (2006). Phonetic encoding of prosodic structure. *Speech production: Models, phonetic processes, and techniques*, 167-186.

Keating, P., Baroni, M., Mattys, S., Scarborough, R., Alwan, A., Auer, E., & Bernstein, L. (2003). Optical phonetics and visual perception of lexical and phrasal stress in English. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)* (pp. 2071-2074).

Keating, P. & Shattuck-Hufnagel, S. (2002) "A Prosodic View of Word Form Encoding for Speech Production" *UCLA Working Papers in Phonetics* 101 112-156.

Lin, P. M. S. (2010). The phonology of formulaic sequences: A review.*Perspectives on formulaic language: Acquisition and communication*, 174-193.

Lin, P. M. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics*, ams017

Processing and production of English focus prosody by native English vs. Mandarin speakers

Chikako Takahashi, Hyunah Baek, Sophia Kao, and Alex HL Yeung
Stony Brook University

Our study addresses two questions: whether native and non-native speakers use prosodic focus cues similarly and whether there is a relationship between speakers' processing and production of focus prosody. While both English and Mandarin (at least some dialects) have been reported to exhibit in-focus expansion (expanded F0/intensity range and longer duration) and post-focus compression (F0/intensity drop and compressed range) (Cooper et al., 1985; Xu, 1999), the location of such prosodic landmarks differs in the two languages. Kao et al. (2016) demonstrated that while a good deal of variability was found in the prosody of both native and non-native speakers, Mandarin speakers' misalignment of the pitch peak and failure to utilize intensity cues affected the perceived naturalness of their English focus intonation.

Our current study, which is focused on the processing of prosodic cues, involved two groups: 21 advanced non-native speakers of English whose L1 is Mandarin (MS) and who had been living in the US as graduate students for less than two years; and 21 native English speakers. The participants were first shown three objects (e.g., *mittens, necklace, sweater*) in varied colors or patterns (e.g., *ivory, purple, flowered, dotted*). Each adjective and noun was bisyllabic with initial stress. Participants heard the instruction "Click on the Adj + Noun" (e.g., "Click on the purple mittens" as shown in Fig.1.). After they responded, four new pictures appeared with the instruction, "Now click on the Adj + Noun" (Fig. 2.). The new instruction contained a contrast in either the adjective, the noun, or both, and was spoken with either natural or unnatural prosody, as shown in Table 1. The participants' reaction time (RT) was measured from the offset of the crucial word (Adj or Noun) to the time of the response.

Preliminary results suggest that the RTs of native speakers were faster overall for all three prosodic conditions shown in Table 1. Furthermore, while both English and Mandarin speakers' RTs show an effect of prosodic naturalness, with longer RTs in the unnatural condition compared to natural focus prosody, the effect was greater for English speakers than Mandarin speakers. This indicates that English speakers were able to use prosodic cues more efficiently than Mandarin speakers.

The same participants took part in a production task designed to elicit contrastive focus intonation in which they instructed an experimenter to place colored objects on a white board. We will compare the results of this task with the processing task results.

Table 1. Processing task conditions

Context: *Click on the purple mittens. Now click on the...*			
		Adjective contrast	Noun contrast
Conditions	Natural prosody + Narrow focus	*SCARLET mittens*	*purple SWEATER*
	Unnatural prosody + Narrow focus	*scarlet MITTENS*	*PURPLE sweater*
	Natural prosody + Broad focus	*SCARLET NECKLACE*	

Figure 1. Processing task: instruction 1 (*Click on the purple mittens*)

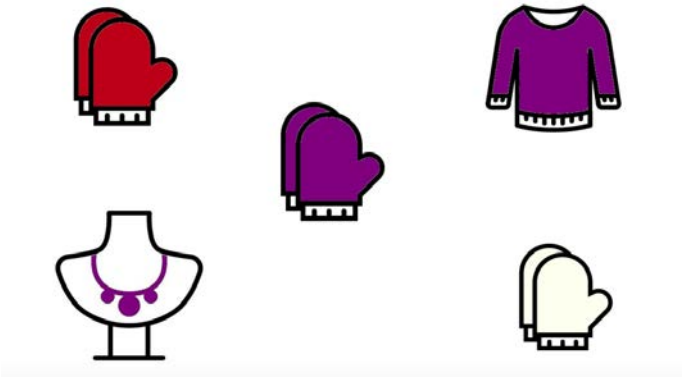


Figure 2. Processing task: instruction 2 (*Now click on the purple sweater*)

References:

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America*, *77*(6), 2142-2156.

Kao, S., Hwang, J., Baek, H., Takahashi, C., & Broselow, E. (2016). International teaching assistants' production of focus intonation. *Proceedings of Meetings on Acoustics 26* (1), 1 - 13.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27, 55-105.

Intonational marking of Information Structure in L2 English

Aya Takeda
University of Hawaii

English pitch accents signal information status of discourse referents. Contrastive and new information are accented, whereas given information is often deaccented/null (Bock & Mazzella 1983; Dahan et al. 2002; Terken & Nooteboom 1987). Furthermore, contrastive and new information are said to be marked with different pitch accent types—L+H* and H*, respectively (Ito & Speer 2008; Watson et al. 2008). In Japanese, new vs. given information is typically differentiated morphosyntactically rather than prosodically, while contrastiveness is signaled by pitch range expansion (Ito et al. 2012), which is phonetically similar to L+H*. In light of the differences between English and Japanese, this study investigates whether Japanese second language learners (L2ers) of English can acquire three prosodic patterns—L+H*, H*, and deaccentuation—and their link to contrastive, new, and given information status. If L1 prosody plays a dominant role in L2 processing as previously reported (e.g. Braun & Tagliapietra 2011), the L+H*- contrastive mapping should be easier than either the H*-new mapping or the null accent-given mapping for L1-Japanese L2ers of English.

Production experiment: In a dyadic, animal-coloring speaking task, English natives (n=70) and Japanese L2ers (n=64) saw a slide that featured colored drawing tools and animals (Fig. 1) and gave instructions to the confederate (e.g. *Use the green paintbrush to color the cow. Now, use the blue paintbrush to color the dolphin.*). The information status of colors and types of drawing instruments was manipulated to create three discourse conditions: both new (*red pen–blue paintbrush*), adjective contrastive (*red paintbrush–blue paintbrush*), and both given (*blue paintbrush–blue paintbrush*). Acoustic analyses on the adjective region (Fig. 2) revealed that the L1 group used higher pitch for the adjective-contrastive condition than for the new condition (b=5.99, p<.01), and the pitch in these two conditions was in turn higher than the pitch in the given condition (b=10.29, p<.001). The L2ers used higher pitch in both new and adjective-contrastive conditions than in the given condition, suggesting that they were able to indicate new/contrastive referents vs. given referents prosodically in production.

Eye-tracking listening experiments: Participants from the production experiment were assigned to either the new–given session (25 natives, 26 L2ers) or the new–adjective contrastive session (45 natives, 38 L2ers). On each trial, they saw a display containing drawing tools and animals; upon hearing the pre-recorded instructions (e.g. *Use the blue paintbrush to color the dolphin. Now, use the green crayon to color the cow.*), they clicked on the appropriate drawing tool and animal. In the new-given session, the second instruction was either both new or both given, crossed with 2 pitch accent types—H* or null accent on the color adjective. In the new-contrastive session, the second instruction was either both new or adjective contrastive, crossed with H* or L+H*. An analysis window of –100ms-900ms was used for an empirical logit analysis (Barr 2008). In the new-given session (Fig. 4), the pitch-by-discourse interaction was significant for the L1 group (b=.41, p<.05), and marginally significant in earlier trials for the L2 group (b=.42, p=.09), which indicated a facilitative effect of null accent in the given context for both natives and L2ers. In the new-adjective contrastive session (Fig. 5), the L1 group showed a significant pitch-by-discourse interaction (b=.46, p<.01) that indicated a facilitative effect of L+H* in the contrastive context, but the L2 group did not.

The study provides a hint that L1-Japanese L2ers of English can acquire the given-null accent mapping. Contrary to our predictions, the L+H*-contrastive mapping turned out to be challenging despite the fact that similar prosodic marking (contrastive marking via pitch expansion) exists in their native language.

Production Experiment

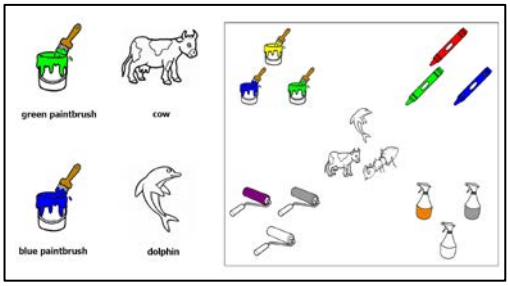


Fig. 1: Sample visual display

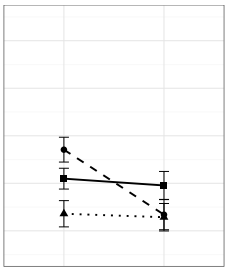


Fig. 2: Mean pitch in ADJ+N regions

Eye-tracking Listening Experiment

New-given session

(1a) Both new—H*
blue paintbrush→green$_{H^*}$ *crayon*

(1b) Both new—null accent
blue paintbrush→green$_{H}$ *crayon*

(1c) Both given—H*
green crayon→green crayon

(1d) Both given—null accent
green crayon→green crayon

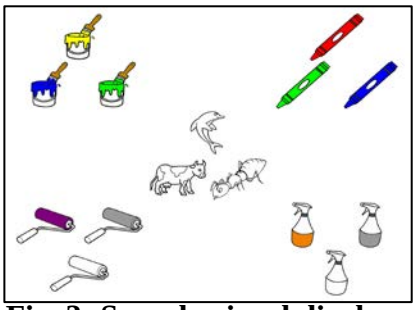


Fig. 3: Sample visual display

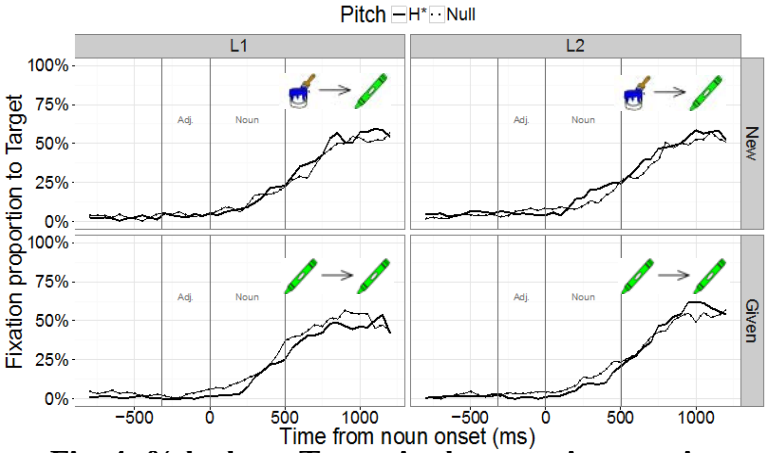


Fig. 4: % looks to Target in the new-given session

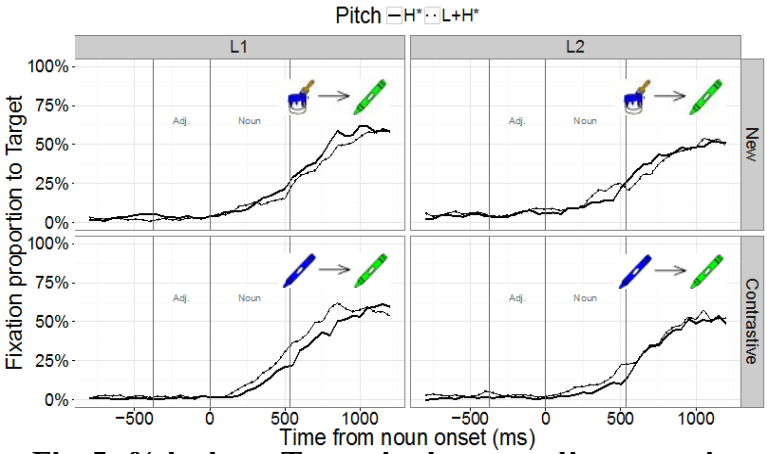


Fig. 5: % looks to Target in the new-adj. contrastive session

New-adj. contrastive session

(2a) Both new—H*
blue paintbrush→green$_{H^*}$ *crayon*

(2b) Both new—L+H*
blue paintbrush→green$_{L+H^*}$ *crayon*

(2c) Adj contrastive—H*
blue crayon→green $_{H^*}$ *crayon*

(2d) Adj contrastive—L+H*
blue crayon→green $_{L+H^*}$ *crayon*

References

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of memory and language*, *59*, 457-474.

Braun, B., & Tagliapietra, L. (2011). On-line interpretation of intonational meaning in L2. *Language and Cognitive Processes*, *26*, 224-235.

Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory and Cognition*, *11*, 64–76.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*, 292-314.

Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, *66*, 265-284.

Ito, K., & Speer, S. R. (2008). Anticipatory effect of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*, 541-573.

Terken, J. M. B., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, *2*, 145–163.

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. (2008). Interpreting pitch accents in on-line comprehension: H* vs. L+H*. *Cognitive Science*, *32*, 1232-1244.

Udon know me: HL or LH? Pitch accent in spoken word recognition

Thanh Lan Truong and Yuki Asano
thanh-lan.truong@uni-tuebingen.de; yuki.asano@es.uni-tuebingen.de
University of Tübingen

During a casual conversation with my supervisor about her research, who is a Japanese native speaker, I noticed that she almost always confused the minimal pair words such as *ame* which can be either accented (realised as a high-low pitch) meaning *rain*, or unaccented (as a low-high pitch with an initial low), meaning *candy* (Vance, 1987). This often led to misunderstandings, because these words are identical in form but different in meaning and in pitch accent production. Inspired by this observation, the current study explored the question whether or not pitch accent information is an integral part of Japanese spoken word recognition.

Previous research has revealed numerous effects of pitch accent in word recognition, but Cutler and Otake (1999) is to date the only study which used the priming method for the investigation of pitch accent (e.g., Asano, 2016; Honda, 2007; Minematsu & Hirose, 1995; Tamaoka, Saito, Kiyama, Timmer, & Verdonschot, 2014; but see Cutler & Otake, 1999). Motivated by the findings of Cutler and Otake (1999), the present study replicated the third experiment. Form priming was used as the method and it has been found that pitch-accent information plays an important role in Japanese word recognition. Form priming is one widely used method for testing word recognition, and it is with no doubt the most direct way to test this phenomenon (Cooper, Cutler, & Wales, 2002); however, this research design is primarily concerned with the activation of lexical form representation, but the authors intended to go into a much deeper level, that is the access of word meaning (Moss & Gaskell, 1999). The current study, therefore, explored the effect of pitch-accent information by using associative priming as the research method. Additionally, two inter-stimulus interval conditions (henceforth, ISI) as manipulation were included: 250 ms for the short ISI condition and 2000 ms for the long one (e.g., Asano, 2016). It is claimed that ISI influences the levels of speech processing. Discussions on human memory have argued that acoustic information, such as pitch, is temporarily stored for 250 ms in the sensory memory after the speech sound has been perceived. It then reaches the working memory, but the listener cannot hold the original acoustic information but only the phonological one, which disappears within two seconds (Baddeley, 2000). ISI, therefore, will also provide a closer insight into the processing of pitch-accent information, meaning whether pitch is phonetically (250 ms) or even phonologically processed (2000 ms) (Cowan & Morse, 1986).

Twelve native speakers of Japanese, who were tested in the laboratory of Tuebingen University, listened to a prime word which was followed by a target word. Prime and target words were presented auditorily. Both words were interrupted by a short or long ISI. Participants were instructed to make word or non-word decisions by button press. Participants' reaction times (=RT) and response accuracy were measured. It was predicted that semantic association effect should be observable for both ISI conditions. Moreover, RTs might vary between the short and long ISI if pitch is not phonologically processed in Japanese spoken word recognition, suggesting that phonetic pitch facilitates word recognition. However, if pitch is also phonologically processed, then there should be no RT difference between short and long ISI.

Data have been collected, and we are currently analysing data. What can be revealed so far is that participants reacted faster when prime and target words shared semantic features, which confirmed the assumption; nonetheless, participants' RTs between both ISI conditions were different. Participants reacted faster in the short ISI condition compared to the long ISI one. This indicates that pitch accent information, stored in the sensory memory, is used in

word recognition. By contrast, pitch remaining as phonological information in the working memory did not influence the RT. Interestingly, Japanese listeners also showed advantage in recognising accented words over unaccented ones. More can be said after a closer examination of the data. This project is still work in progress; hence, a detailed explanation of the findings will be discussed in the talk.

References

Asano, Y. (2016). *Localising foreign accents in speech perception, storage and production.* (Ph.D. Doctoral Thesis), University of Konstanz, Konstanz.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn Sci, 4*(11), 417-423.

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

Cowan, N., & Morse, P. A. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of Acoustical Society of America, 79*(2), 500-507.

Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *J Acoust Soc Am, 105*(3), 1877-1888.

Honda, M. (2007). The role of prosody in Japanese: The use of pitch information in spoken word recognition by L1 and L2 speakers. *CamLing*, 96-103.

Minematsu, N., & Hirose, K. (1995). Role of prosodic features in the human process of perceiving spoken words and sentences in Japanese. *Journal of the Acoustical Society of Japan, 16*, 311-320.

Moss, H. E., & Gaskell, M. G. (1999). Lexical semantic processing during speech comprehension. In S. Garrod & M. Pickering (Eds.), *Language Processing* (pp. 59-99). UK: Psychology Press.

Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R. G. (2014). Is pitchaccent necessary for comprehension by native Japanese speakers? - An ERP investigation. *Journal of Neurolinguistics, 27*, 31-40.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany, NY: State Univ. of New York Pr.

Intonation of polar questions produced by 2.5- to 4-year-olds

Talina Weber, Muna Schönhuber, Janet Grijzenhout
talina.weber@uni-konstanz.de
University of Konstanz, Germany

Previous research has shown that in adult speech, statements in German and English are frequently accompanied by falling intonation and polar questions by a rise (e.g., Wochner, Schlegel, Dehé, & Braun, 2015 for German, Pruitt & Roelofsen 2013 for English). Regarding the acquisition of such prosodic patterns, it has been reported that English-learning 1-year-olds do not actively control sentence intonation and that pre-schoolers have difficulties with rising intonation (Snow, 2002, 2004). Instead of rising f0, 4-year-olds rely on final syllable duration to signal interrogativity, 7-year-olds use a combination of rising f0 and final syllable duration and only 11-year-olds are able to rely on f0 alone (Patel & Grigos, 2006). Yet, there are also studies suggesting good intonational control in 2- and 3-year-old German and Spanish mono- and bilinguals (Lleó & Rakow, 2011). In order to find out more about children's intonational realisation of polar questions in first language acquisition, we conducted a production experiment, similar to the procedure used by Crain and Nakayama (1987), eliciting polar questions and declarative statements from 12 German-learning 2.5- to 4-year-olds.

In the first phase of the experiment, children watched a hand puppet play, which set the scene for the succeeding elicitation phase. In this second phase, children examined one of the hand puppets with instruments of a doctor's bag and the experimenter encouraged the child to address the hand puppet with statements and polar questions. Prompts were presented in direct and indirect speech so that children either had to imitate or to rephrase the required utterances. Recordings were labelled for boundary tones and pitch accents, according to the GToBI annotation system proposed by Grice, Baumann, and Benzmüller (2005) as well as for f0 minima and maxima within the range from the final accented syllable to the right boundary tone. Pitch range was calculated in semitones (st).

Our preliminary results show that statements are predominantly marked by falling intonation, independent of age. In order to mark an utterance as a polar question, children between 3 and 4 years make use of rising intonation more reliably than children of the youngest age group (see fig. 1). This could mean that children younger than 3 years are still uncertain about which intonational pattern to apply. Alternatively, the inconsistent use of rises for polar questions might be due to the fact that, physiologically, the production of rises requires more effort than the production of falls, which emerge automatically as subglottal pressure decreases (see Lieberman, 1967). For the analysis of pitch range a linear-mixed-effects regression model revealed that age group had no effect at all, but there was a significant effect of intonation ($p = 0.03$), showing that the range of rising utterances (average 6.11st) was on average 1.52st larger than that of falls (average 4.59st; see fig. 2). Overall, our data support the conflicting findings from previous studies.

Our findings may be taken to reflect the conflicting findings of previous research. On the one hand, 2.5- to 3-year-olds do not use rising intonation for polar questions consistently and our data are in line with similar findings by, e.g., Snow (2002, 2004) and Patel and Grigos (2006), who suggest that young children may have difficulties with rising intonation. On the other hand, the analysis of accent range shows that intonational patterns can be mastered fairly early, similar to the findings by, e.g., Lleó and Rakow (2011). Thus, there is no problem of producing rises *per se*, but rather of producing them consistently. It remains to be shown which factors determine whether or not a polar question is produced with rising or with falling intonation.

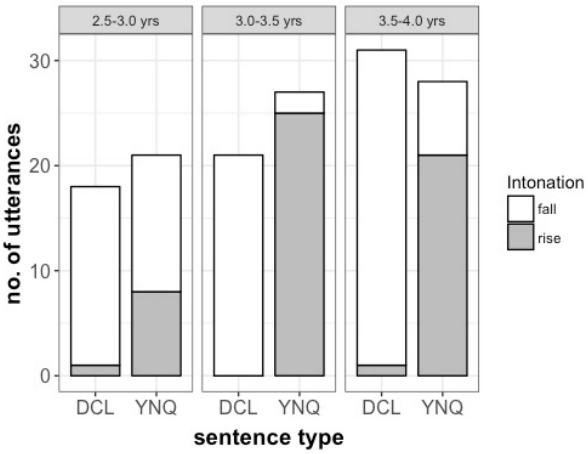


Fig. 1: *Realisation of intonation in statement (DCL) and polar question (YNQ) targets by age group.*

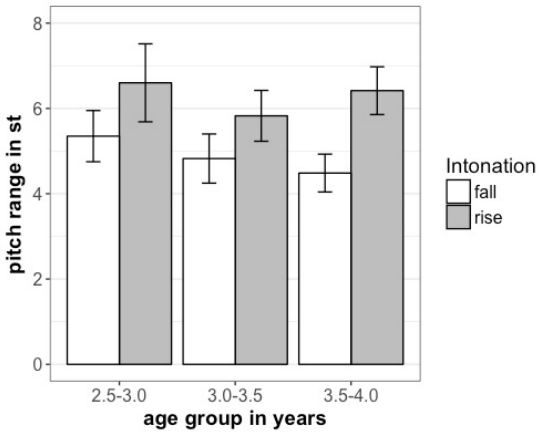


Fig. 2: *Pitch range in falling and rising utterances by age group (whiskers represent standard errors).*

References

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63*(3), 522-543.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. *Jun, Sun-Ah (ed.): Prosodic typology. The phonology of intonation and phrasing., Oxford University Press*, 55-83.

Lieberman, P. (1967). *Intonation, perception, and language* (Vol. 38): M.I.T. Pr.

Lleó, C., & Rakow, M. (2011). Intonation targets of yes/no questions by Spanish and German monolingual and bilingual children. In E. Rinke & T. Kupisch (Eds.), *The development of grammar. Language acquisition and diachronic change. In honour of Jürgen Meisel* (pp. 263-286). Amsterdam/Philadelphia John Benjamins.

Patel, R., & Grigos, M. I. (2006). Acoustic characterization of the question–statement contrast in 4, 7 and 11 year-old children. *Speech Communication, 48*(10), 1308-1318.

Pruitt, K., & Roelofsen, F. (2013). The interpretation of prosody in disjunctive questions. *Linguistic Inquiry 44*, 632-650.

Snow, D. (2002). Intonation in the monosyllabic utterances of 1-year-olds. *Infant Behavior and Development, 24*(4), 393-407.

Snow, D. (2004). Falling intonation in the one- and two-syllable utterances of infants and preschoolers. *Journal of Phonetics, 32*(3), 373-393.

Wochner, D., Schlegel, J., Dehé, N., & Braun, B. (2015). The prosodic marking of rhetorical questions in German. *Speech Prosody*, [no page numbers].

Tracking the Perceptual Effects of Backchannel Behaviour in Asperger Syndrome and Second Language Speech

Simon Wehrle, Timo B. Roettger & Martine Grice
IfL Phonetik, Universität zu Köln

When compared to native neurotypicals, learners of a second language (L2) and speakers diagnosed with Asperger Syndrome (AS) share a number of prosodic characteristics such as a slower speech rate, a different distribution of disfluencies [1,2,3], different accentuation patterns (presence and type of pitch accents), as well as extremes of pitch range (either wider or narrower than controls) [4,5].

For L2 speakers, there is also evidence of problems with backchannel (BC) behaviour in conversation [6,7,8]. We are currently investigating backchannels in a parallel corpus of Map Task dialogues with pairs of Vietnamese L2 speakers and pairs of speakers with a diagnosis of AS. Analysis is focussed on the frequency, timing and intonation contours of BC productions.

Dialogues containing these productions will then be employed in a perception experiment designed to test the effects of backchannel behaviour on factors such as the perceived naturalness and friendliness of speakers. As opposed to the perception study in [9], which employed Likert scales to rate individual BC utterances post-hoc, our aim is to dynamically track attributed listener involvement during the unfolding of whole stretches of discourse.

As evaluation of discourse information of this kind spans a large temporal window and therefore makes classic real-time assessment difficult, we will be adapting the “meandering mice” paradigm, an innovative methodology borrowed from cognitive science. In [10] this way of tracking a participant’s mouse movements over longer periods is used to reveal the time course of decision making in a cognitive task which necessitates the inhibition of intuitive responses (the Cognitive Reflection Test). Mouse cursor trajectories are linked to processes of reasoning, thereby revealing, in real-time, the path of decision-making that leads to participants’ final answers in the test.

This easy-to-implement, flexible methodological paradigm can be adapted to our study of the perception of BC behaviour for instance by telling participants to listen to excerpts from the Map Task dialogues and rate the level of interest they perceive to be shown by the instruction follower/listener through reference to certain pre-defined regions on a computer display. Simply ensuring a participant’s hand rests on a computer mouse during the entire course of the experiment should provide us with a contextually valid, continuous measurement of attributed listener involvement. Moreover, this approach affords a unique way to judge the effects of a lack of – or even a complete absence of – BCs produced in a communicative situation, an important factor which is not amenable to experimental investigation at all with methodologies used in previous studies on similar topics.

The overarching aim of the project that forms the basis of this study is a large-scale comparative analysis of the communicative characteristics of L2 learners and those of AS speakers as expressed through prosody. Notwithstanding the obvious differences, the motivation for this novel approach lies in a number of intriguing similarities between the two groups, such as problems in social interaction and interpersonal communication [11,12,13].

Whilst AS speakers often find the behaviour and the utterances of neurotypical speakers bewildering, this situation is familiar to many neurotypicals themselves when living in a non-native culture and speaking a non-native language. In the latter case, even individuals with a reasonable knowledge of the culture and a proficiency in the language sometimes experience situations in which communication breaks down or misunderstandings occur. As prosody and subtle conversational signals such as BCs play an integral role in determining the interpretation and the eventual outcome of any such communicative situation [9,14,15], we

believe that the study presented here will be an important contribution to a deeper understanding of these issues.

References

Shriberg, L.D., Paul, R., McSweeny, J.L., Klin, A., Cohen, D.J. & Volkmar, F.R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research, 44*(5), 1097-1115.

Munro, M.J. & Derwing, T.M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: the role of speaking rate. *Studies in Second Language Acquisition, 23*(4), 451-468.

MacFarlane, H., Gorman, K., Ingham, R., Hill, A. P., Papadakis, K., Kiss, G., & van Santen, J. (2017). Quantitative analysis of disfluency in children with autism spectrum disorder or language impairment. *PloS one, 12*(3).

McCann, J. & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders, 38*(4), 325-350.

Mennen, I., Phonological and phonetic influences in non-native intonation. (2004). In J. Trouvain & U. Gut (Eds.), *Non-native Prosody: Phonetic Description and Teaching Practice* (pp. 53-76). Berlin: Mouton de Gruyter.

Cutrone, P. (2011). Politeness and face theory: Implications for the backchannel style of Japanese L1/L2 speakers. *Language Studies Working Papers, 3*, 51-57.

Li, H.Z. (2006). Backchannel responses as misleading feedback in intercultural discourse. *Journal of Intercultural Communication Research, 35*(2), 99-116.

Shelley, L., & Gonzalez, F. (2013). Back Channeling: Function of Back Channeling and L1 Effects on Back Channeling in L2. *Linguistic Portfolios, 2*(1), 97-108.

Ha, K.P., Ebner, S. & Grice, M. (2016). Speech prosody and possible misunderstandings in intercultural talk - A study of listener behaviour in Vietnamese and German dialogues. *Proc. Speech Prosody 8*, Boston, 801-805.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition, 150*, 109-118.

Vogeley, K. (2012). *Anders sein: Asperger-Syndrom und Hochfunktionaler Autismus im Erwachsenenalter*. Weinheim: Beltz.

Grandin, T. & Panek, R. (2013). *The autistic brain: Thinking across the spectrum*. Boston: Houghton Mifflin Harcourt.

Baron-Cohen, S. (2008). *Autism and Asperger Syndrome*. Oxford: Oxford University Press.

Gudykunst, W.B. & Mody, B. (2002). *Handbook of International and Intercultural Communication*, 2nd ed. Thousand Oaks: Sage.

Hirschfeld, U., Lange, F. & Stock, E. (Eds.) (2016). *Phonetische und rhetorische Aspekte der interkulturellen Kommunikation*. Berlin: Frank & Timme.

Pitch and syntactic disambiguation in English and German: An Eye-tracking Study

Jenny Yu[1,2], Heather Kember[1,2], Robert Mailhammer[1], and Anne Cutler[1,2]
[1]*The MARCS Institute, Western Sydney University, Australia*
[2]*The ARC Centre of Excellence for the Dynamics of Language, Australia*

English and German have similar prosodic structure, but may use pitch differently to disambiguate syntactic structures. To distinguish verb-phrase (VP) attachments (e.g. *The man visited the zoo* ***with his niece***) from noun-phrase (NP) attachments (e.g. *The man visited* ***the zoo with the tigers***), for example, both language groups tend to lengthen the direct object e.g. '*zoo*' in VP-attachments, but the verb e.g. '*visited*' in NP-attachments; however, German speakers further mark the boundary in the VP-attachment case with a pitch rise on the direct object (Snedeker & Trueswell, 2003; O'Brien et al., 2014). In a visual world experiment, we examine whether this production difference in pitch cues for temporarily structurally ambiguous sentences affects the time-course of disambiguation by listeners across languages. We predict that if the additional pitch rise cue is significant in helping to distinguish the two types of attachment in German, then native listeners should disambiguate at an earlier time point in German than in English.

We created sentences in English and German that had a temporarily ambiguous structure until the disambiguating clause. For each experimental sentence, four variations were created. Two variations of each sentence ended with a clause attachment that referred to the verb (VP attachments) and two variations included a clause attachment that modified the noun (NP attachments) (see (1) for an example). Sentences were equivalent in meaning and construction across languages, and nouns in clause attachments were also controlled for frequency and length between and within languages. For each experimental item, a scene was created comprising four pictures organised randomly in four quadrants. Two pictures depicted NP-related objects (e.g. tiger, lion) and two pictures depicted VP-related objects (e.g., niece, nephew).

Native English and German speakers listened to recorded sentences in their L1 and their eye movements were tracked as they inspected a scene. To maintain participants' attention, they were instructed to click on the picture most representative of each sentence. Experimental sentences were divided so that participants are presented with only one sentence from each pairing, and equal numbers of NP-attachment and VP-attachment sentences. In addition, 32 filler sentences were created with a mix of different sentence structures. These were semi-randomly interspersed with experimental items so that participants listened to 48 items in total. For filler sentences, the visual display showed pictures related to nouns or verbs in the sentence. The influence of prosodic cues on disambiguation was measured by comparing the proportion of anticipatory eye movements directed to VP vs. NP pictures over the course of sentences.

Preliminary results indicate that both English and German listeners prefer looking at NP related objects over VP related objects before the arrival of disambiguating information, regardless of attachment type (refer to *Figure. 2*). However, English participants show this preference prior to the verb, while German participants indicate NP object preference throughout the entire course of the utterance. In addition, eye movement patterns show that German listeners shift towards looking at VP objects earlier in VP attachment sentences than English listeners. Therefore, German listeners are disambiguating faster than English speakers, but it is not clear whether they are using information such as the article for the noun, or whether it is prosodic. Further stimuli analyses will investigate the influence of syntax and prosodic cues on the time of disambiguation.

(1) Example of experimental stimuli in English and German:

	English	**German**
NP	The man visited the zoo *with the new tigers.*	Der Mann besuchte den Zoo *mit den neuen Tigern.*
	The man visited the zoo *with the new lions.*	Der Mann besuchte den Zoo *mit den neuen Löwen.*
VP	The man visited the zoo *with his young niece*.	Der Mann besuchte den Zoo *mit seiner Nichte.*
	The man visited the zoo *with his young nephew.*	Der Mann besuchte den Zoo *mit seinem Neffe.*

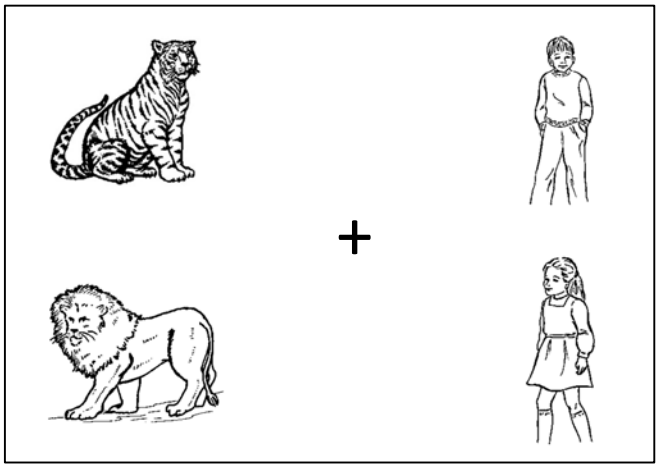


Figure 1. Example of a visual scene presented to participants for the sentence beginning "The man visited the zoo with…" with NP-related objects *tiger* and *lion*, and VP-related objects *nephew* and *niece*.

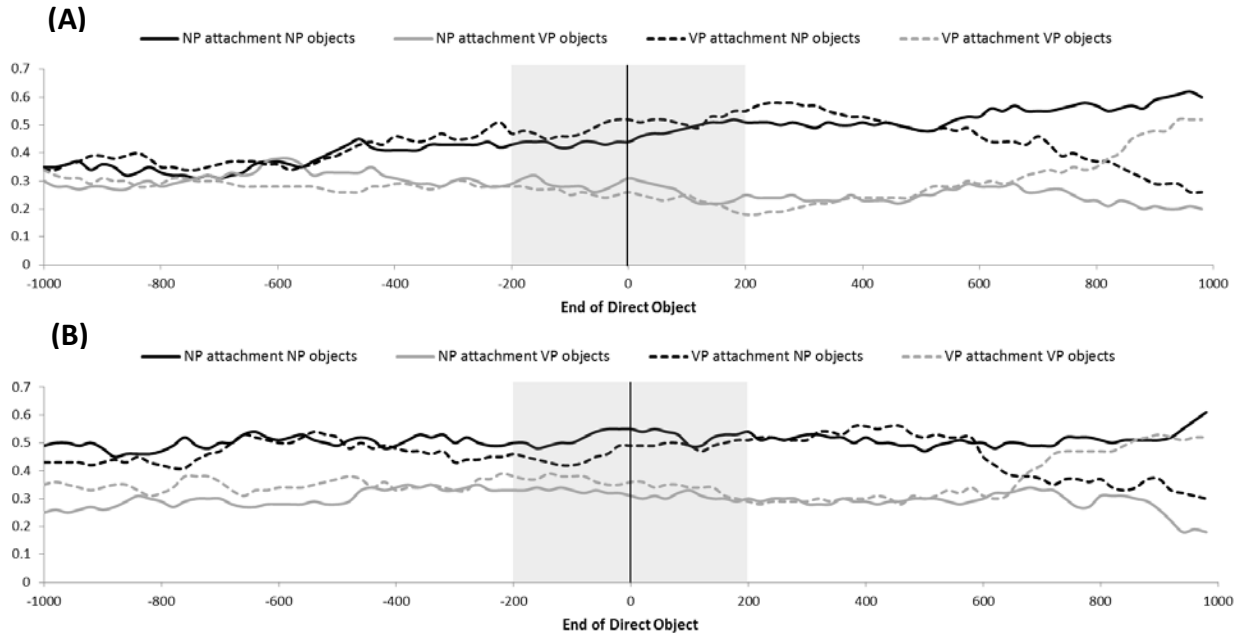


Figure 2. Average proportion of looks to noun-phrase (NP) and verb-phrase (VP) objects aligned to -1000 ms prior to the end of the direct object to 1000 ms post (preliminary data from 18 English participants (A) and 15 German participants (B), from a planned total of 24 participants per group)

References

O'Brien, M. G., Jackson, C. N., & Gardner, C. E. (2014). Cross-linguistic differences in prosodic cues to syntactic disambiguation in German and English. *Applied Psycholinguistics, 35*(1), 27-70. doi: 10.1017/S0142716412000252.

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48, 103-130.

Processing Tonal Contrast between native Chinese and English listeners

Sin Kwan Yuen & Yuki Asano
sin-kwan.yuen@student.uni-tuebingen.de; yuki.asano@es.uni-tuebingen.de
University of Tübingen

Mandarin Chinese is a tone language where a change of tone also changes the lexical meaning, whereas English is a non-tonal language since pitch movement in English lexemes does not play a role in alternation of lexical meaning (Yip, 2002). However, English still possesses tonal features such as lexical stress and post-lexical or paralinguistic intonation (e.g., Ladd, 2008; Cutler, 1986; Cutler, Dahan and Donselaar, 1997). The present study aims to investigate how well native English listeners could discriminate rising and falling Chinese tonal contrasts (namely Tone 2 and Tone 4 respectively) by exploiting their tonal familiarity in English despite the absence of lexical distinction based on a tonal contrast.

Twelve native English subjects from different English speaking countries and twelve native Mandarin Chinese subjects from Mainland China participated in a speeded same-different judgement task with their age range between 18 – 40 years old, and they had no hearing impairment. In addition, all English listeners had no prior knowledge in Mandarin Chinese. Twenty minimal pairs of monosyllabic Chinese word and twenty minimal pairs of monosyllabic Chinese pseudoword that were either *same* or *different* with respect to their pitch contour (T2-T2, T4-T4 or T2-T4, T4-T2) were used, e.g. *ai* -T2 (= cancer) vs. *ai* -T4 (= love), and *fiao* -T2 vs. *fiao* -T4 (= pseudowords). The stimuli were recorded by a native female speaker of standard Mandarin Chinese. For each category, only one token was used and each token was arranged in four conditions. In order to prevent participants from relying on acoustic correlates of the stimuli to distinguish the contrast, a longer inter-stimulus interval (2000 ms) was used (Asano, to appear). One of the six randomized lists was assigned to each subject. Mean accuracy rates and mean reaction times were calculated in each tonal contrast condition separately for Chinese and English listeners.

The result showed that native English subjects had 100% accuracy rate in the same condition both for the Chinese word and pseudoword conditions, but their accuracy rate decreased in the *different* condition (98.33% in T2- T4 pair; 99.36% in T4-T2 pair, which statistically did not differ from each other, $p > .99$, and between T4-T4 vs. T2-T4 and T2-T2 vs. T2-T4 conditions, both $p < .05$). On the other hand, native Mandarin Chinese subjects maintained their higher accuracy rate both in *same* and *different* conditions (100% in T4-T4 pair; 99.79% in T2-T2 pair; 100% in T2- T4 pair; 99.79% in T4-T2 pair, they did not differ from each other, $p > .99$). Moreover, all four conditions performed by Chinese subjects differed significantly with English subjects in *different* T2-T4 condition ($p < .05$), but not when compared to *different* T4-T2 condition ($p > .99$). Chinese subjects showed longer reaction times in the pseudoword condition than in the word condition ($p < .05$), whereas English listeners did not show this difference between the conditions ($p > .99$).

English native listeners generally showed high accuracy rates indicating that their knowledge of the lexical stress and post-lexical characteristics in their mother tongue language aid them to discriminate nonnative lexical tonal contrasts. However, their discrimination ability decreased in *different* conditions, particularly in T2-T4 pair, showing that English listeners placed more importance on pitch height than pitch direction (Chandrasekaran et al., 2007), and the height of F0 onset for the first tone has to be higher than that of the second tone within a tone pair during tonal discrimination (Lu et al., 2015). In addition, the longer reaction times found in the pseudoword condition than in the word condition by Chinese subjects illustrated that there was a lexical facilitation effect in the word condition to accomplish the discrimination task.

References

Asano, Y. (to appear). Discriminating non-native segmental length contrasts under increased task demands, *Language and Speech*.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. Brain Research, 1128, 148-156.

Cutler, A. (1986). Phonological structure in speech recognition. *Phonology Yearbook, 3*, 161-178.

Cutler, A., Dahan, D., & Donselaar, W. van (1997). Prosody in the comprehension of spoken language: A literature review. Language and Speech, 40(2), 141-201.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: University Press.

Lu, S., Wayland, R., & Kaan, E. (2015). Effects of production training and perception training on lexical tone perception - A behavorial and ERP study. Brain Research, 1624, 28-44.

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

Identifying Interrogative Tunes in Tianjin Mandarin

Cong Zhang
University of Oxford

Introduction: Alternating pitch accents and boundary tones are common methods of making an interrogative tune in non-tonal languages (e.g. English, Bengali, Greek, etc.); however, in Tianjin Mandarin, a tonal dialect of northern Chinese (see Figure 1 for tonal inventory), making a syntactically unmarked yes-no question (YNQ) is much restricted by the pitch contours and registers of lexical tones, since both lexical tones and intonational tunes are intrinsically linked with pitch modulation. Zhang (2016) discovered that the differences between syntactically unmarked YNQ and statement in Tianjin Mandarin are **register** change and an extra **floating H% boundary tone** for intonational YNQ: YNQ is always higher in register than statements; a floating H% boundary tone is added at the right boundary in YNQ, which facilitates the rise in rising tones and deters the falling in falling tones. The differences do not evidently display in surface phonetic pitch contour. With the restriction from lexical tones, can native listeners still perceive the YNQ tune well? The current study investigates i) whether listeners could distinguish YNQ from statements by merely using the subtle cues of register change and the H% floating boundary tone; and ii) how the cues are processed during perception.

Methods: 28 native Tianjin Mandarin speakers (15 male and 13 female) took part in the experiment. They were instructed to perform a forced-choice identification task by pressing either the 'Q' button or 'S' button on the handsets as quickly as possible when hearing an utterance. Accuracy and reaction time were recorded and statistically analysed with multi-factor ANOVA Test, with sentence TYPE (Question, Statement), lexical TONE (L, H, LH, HL), and stimuli speaker GENDER (male, female) as independent variables, and accuracy rate or reaction time as the dependent variable.

Results: Table 1 shows the accuracy rate of the identification task. The lexical tones are presented by average accuracy rate (Figure 2). When results from two lexical tones are not statistically significant, the relation is represented with a "≈".

The crucial cue for perceiving statements is the initial L tone at the left boundary. Lexical Tone 1 and Lexical Tone 3 (L tone and LH tone) both start with a L, so they both achieve the highest accuracy. When the initial tone does not help with identification, such as in Lexical Tone 4 and Lexical Tone 2 (HL and H tone), the ending H tone interferes with the identification. Lexical Tone 2 (H) ends with a H tone that interferes with the identification, therefore the accuracy is extremely low. As for Lexical Tone 4 (HL), although it does not have a facilitating L initial tone, it does not have an interfering H tone at the end either. T4(HL) therefore is easier to identify than T2(H).

Conversely, during the perception of YNQ, the pitch height of the right boundary is the first determining factor. Then, contrary to that of the statements, the initial L interferes with the identification. Lexical Tone 2 (H) ends with a H tone and does not have any interfering L tone. It, therefore, achieves the highest accuracy. T1, a L tone which starts with a L and ends with a L, has the lowest on the contrary. T3 is higher than T4 but not significantly, which indicates the effect of both missing the facilitative ending H and having the interfering initial L both burdens the identification.

The reaction time results (Figure 3) are in accordance with the accuracy results: the more accurate, the faster the responses.

Conclusions: The results show that it is not easy to identify tune types in a tonal language, especially when the segmental information is limited. During the process of perceiving tunes, our brain processes the pitch height of both ends of an utterance but subconsciously uses strategies with different directionalities. The low accuracy of this task also provides a

potential reason for the existence of syntactically marked yes-no questions, which is more frequently used to ask a question without prior context.

Table 1: Results of average accuracy rate of identification task

Sentence type	Accuracy
Statement	T3(LH) ≈ T1(L) > T4(HL) > T2(H)
Yes-No Question	T2(H) > T3(LH) ≈ T4(HL) > T1(L)

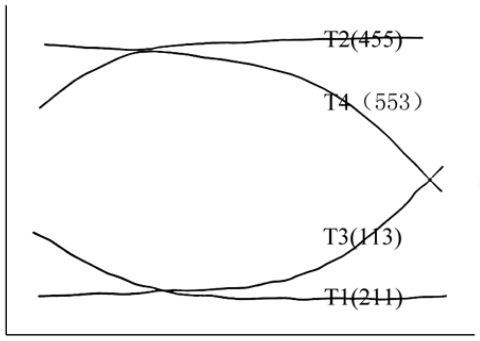


Figure 1: Tianjin Mandarin lexical tones with tone numbers in brackets.
T1: L (211); T2: H (455); T3: LH(113); T4: HL (553)

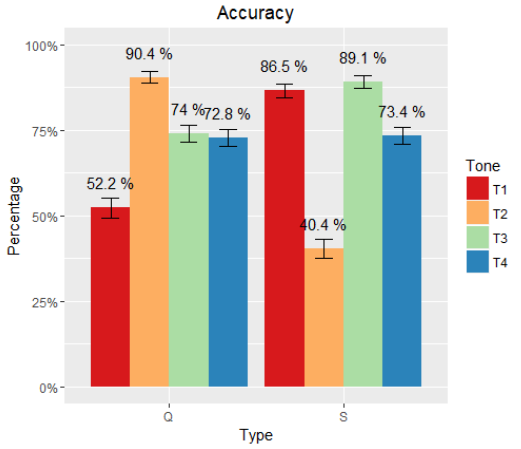


Figure 2: Accuracy for identification of Yes-No questions and statements

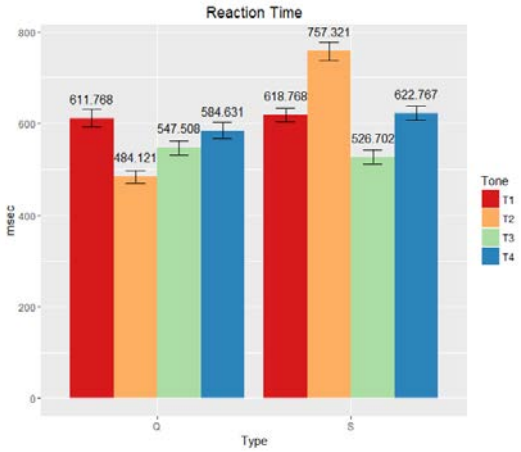


Figure 3: Reaction Time for identification of Yes-No questions and statements

References

Zhang, C. 2016. Tones and Tunes in Tianjin Mandarin. *LAGB 2016 (Linguistics Association of Great Britain Annual Meeting 2016)*, 6-9 September, 2016. University of York, UK.

Some Lunch Recommendations

In the Old Town of Tübingen

Schmälzle (€) **Kirchgasse 10**
Swabian lunch specialities takeaway (eat e.g., on the stairs of Stiftskirche)

Collegium (€€) **Lange Gasse 8**
Affordable and cosy restaurant with one 2-course lunch option and a few options à la carte (e.g., tarte flambée)

Esszimmer (€) **Am Lustnauer Tor 4**
International, sustainable, healthy food

Kalender (€) **Gartenstraße 1**
One of Tübingen's best typical Döner Kebap takeaway (few seats available)

Some Lunch Recommendations

In the Old Town of Tübingen

Neckarmüller (€€) **Gartenstraße 4**
Tübingen's most famous beer-garden serving Swabian beer-garden food

Bäckerei Gehr (€) **Am Lustnauer Tor 5**
Bakery - Sandwiches to go

Kichererbse (€) **Metzgergasse 2**
Vegetarian falafel takeaway (few seats available)

Vegi (€) **Kornhausstr. 1**
Vegetarian falafel restaurant and takeaway (nice seating area inside and outside)

ProPro Workshop 2017

Thursday, August 31, 2017

Lunch break, 13:00 - 14:30

Quick Bites

1) Schmälzle (Kirchgasse 10)

2) Kichererbse (Metzgergasse 2)

3) Vegi (Kornhausstraße 1)

4) Bäckerei Gehr (Am Lustnauer Tor 5)

5) Kalender (Gartenstraße 1)

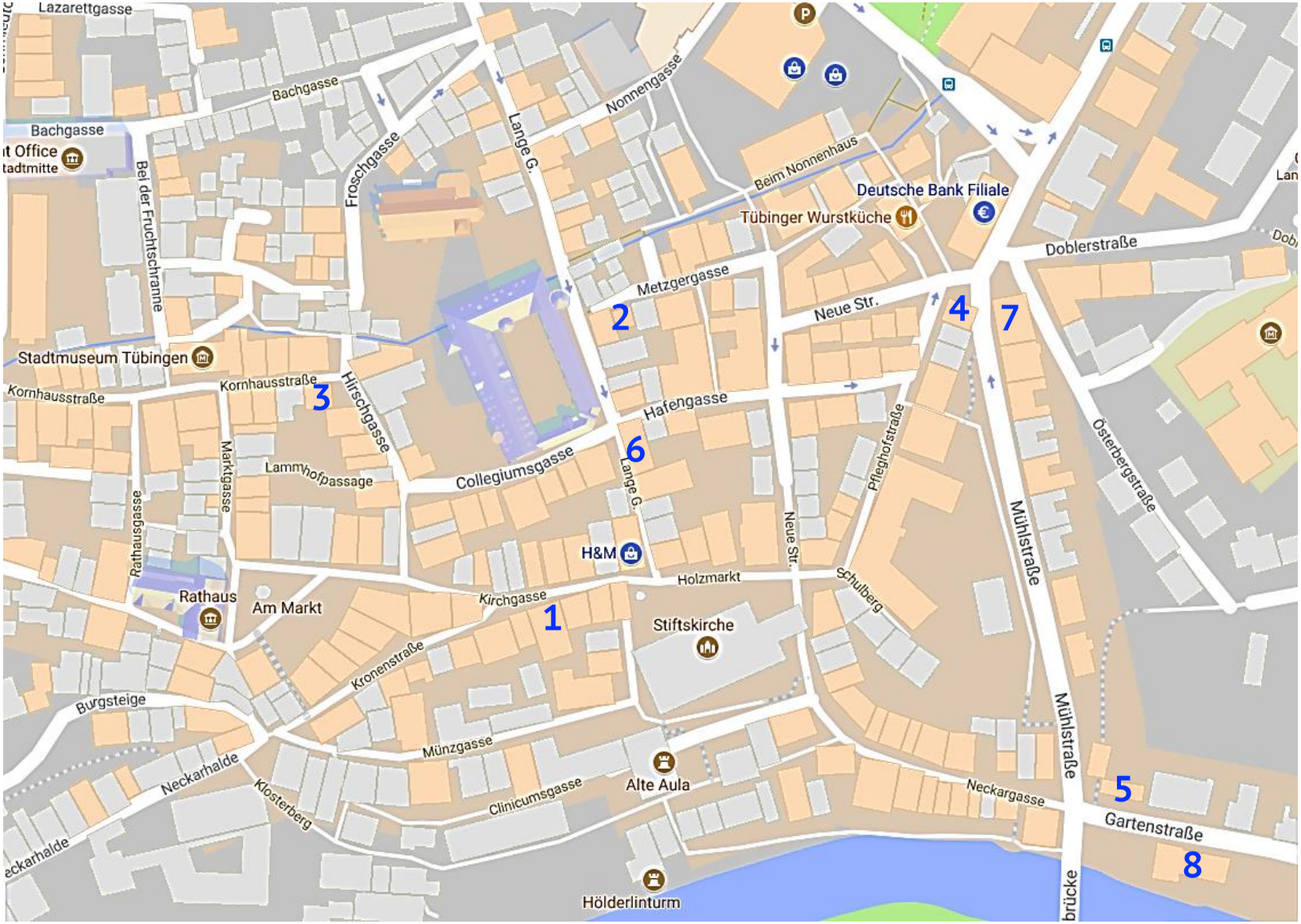
Student-friendly Restaurants

6) Collegium (Lange Gasse 8)

7) Esszimmer (Am Lustnauer Tor 4)

Traditional

8)Neckarmüller (Gartenstraße 4)



Further Recommendations

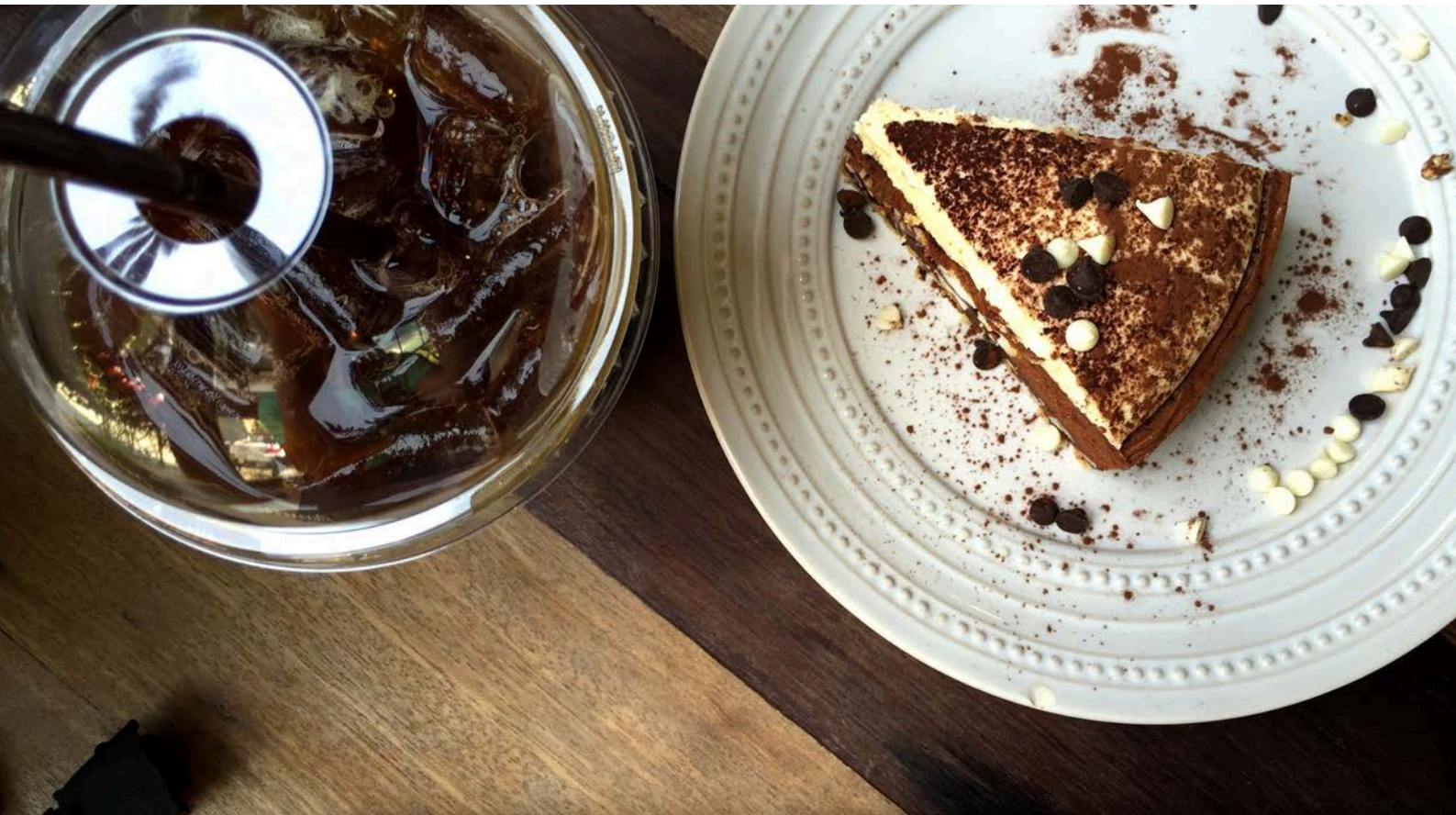
Around Tübingen

Asia-Imbiss Wok-In (€) — **Wilhelmstr. 2**
Chinese food

Saints & Scholars (€€) — **Wilhelmstr. 44**
Irish Pub

Willi (€) — **Wilhelmstr. 3/1**
Coffee shop

Manufaktur (€) — **Vor dem Haagtor 1/2**
Pizza



Important Telephone Numbers

- Police: 110
- Fire/ Ambulance: 112
- Deutsche Bahn (German Railways): +49 (0) 180 6 99 66 33
- Train schedule information: +49 (0) 800 1 50 70 90
- TüBus (local bus company): +49 (0) 7071 157 157
- Airport Stuttgart: +49 (0) 711 948 0
- Taxi: +49 (0) 7071 920 555 +49 (0) 7071 14 38 591

Pharmacies
You can find pharmacies all over Tübingen. Most of them are listed here: http://goo.gl/rF8xeZ If you click on the name of the pharmacy, a map will show you its location. Emergency Pharmacies are open during the night, bank holidays and on Sundays (Notdienstapotheken).

Hospital
The university hospital is across from campus. It contains all important institutes. If you need only a doctor, you can call the German Red Cross: +49 (0) 7071 7000 0