Speech Prosody 2016
31 May - 3 June 2016, Boston, USA

Towards a typology of prominence perception: the role of duration

Adrian Leemann[1], *Marie-José Kolly*[2], *Yang Li*[1], *Ricky Chan*[1], *Geraldine Kwek*[1], *Anna Jespersen*[1]

[1]Phonetics Laboratory, DTAL, University of Cambridge
[2]Phonetics Laboratory, Department of Comparative Linguistics, University of Zurich
{al764|yl492|kwrc2|gsck2|abj28}@cam.ac.uk, marie-jose.kolly@uzh.ch

Abstract

Listeners of different languages have been reported to vary significantly in prominence perception tasks. We know very little, however, about which exact cues different listeners use in these tasks. In this study, we examined the role of duration in the perception of prominence in both typologically different and related languages. The stimuli consisted of the disyllabic logatome <baba> for which the durations of the first and second syllable were systematically manipulated. 80 listeners (8 varieties*10 listeners/variety) judged the relative prominence of the two syllables. We found that differences in the sensitivity to duration cues between varieties of the same language can be equal in magnitude to those found for typologically unrelated languages. Results are discussed in light of prosodic typology and speech perception.

Index Terms: prominence, perception, stress, linguistic typology, varieties of English, dialectology

1. Introduction

Articulating strings of sounds more prominently makes them more salient. Speakers place prominence on strings of sounds to mark expressive meaning in an utterance [1, 2]; at the same time, prominence placement is closely linked to lexical stress. Prominent syllables are rapid departures in loudness, duration, and/or pitch relative to adjacent syllables [3]. Typically, prominent syllables are louder, longer, and higher in f0 [4]. Not all languages mark prominence concurrently with these parameters: increased duration seems to be less of an acoustic correlate of stress for Japanese than it is for English, for example [5]. Research has shown that the relative contributions of these acoustic parameters in prominence production tend to be language-specific [6].

Acoustic correlates in prominence production can, but do not necessarily have to, translate to prominence perception. [7] found that a central parameter in prominence production for English is intensity; much less important is f0. However, the latter was reported to play a greater role in perception [ibid.]. Research further suggests a trading relationship between cues: [8] reports that listeners judge longer syllables with lower intensity equally prominent as shorter syllables with higher intensity. Space prevents a comprehensive review, but there are numerous studies which have investigated the complexity of prominence perception for individual languages or typologically similar languages. However, there has been little systematic study, using unified procedure and stimuli, of the cues that are used in prominence perception for speakers of typologically diverse languages.

The purpose of the experiment described in the present paper is to examine the perceptual reality of prominence in typologically similar and different languages. In this proof-of-concept study we examined the contribution of one cue only – duration – using the same disyllabic stimuli across all languages. Early examples of research in this area have shown, for instance, that Estonian listeners were more responsive to duration cues than English listeners [9]: this is potentially attributable to the greater role duration plays in the Estonian quantity system. This suggests that the prosodic structure of listeners' native languages significantly influences their perception of prominence. More recent studies found that English listeners tended to place equal weight to intensity, pitch, and duration cues – while Swedish listeners attached more weight to vocal effort [10]. In a pilot study on 6 languages, [11] reported duration to be a strong predictor in the perception of prominence, followed by f0 variation.

We selected 8 varieties (6 languages; 2 of which were represented by two different varieties) which have been reported to differ in prosodic structure, see Figure 1.

Figure 1: *Varieties selected for the present study.*

What follows is a brief account of research on prominence patterns and the use of durational cues for the varieties selected. For each variety, we propose predictions in terms of sensitivity to duration cues and bias towards perceiving the first or second syllable in a disyllabic word as more prominent. Predictions are based on our understanding of the literature.

- **Standard Southern British (SSBE)** and **Singapore English (SgE)**: SSBE has variable stress and shows quantity contrast – coinciding with quality contrasts – in its vowel system [12]. Despite variable stress, trochaic stress patterns are prevalent in disyllabic words, particularly in words of Germanic origin [13]. Although SgE, too, has been reported to have variable stress, it seems to be less predictable than for SSBE [14]. Additionally, SgE differs from SSBE and has been reported to bear more resemblance to Standard French in its rhythmic patterns as a rather syllable-timed language [15, 16, 17, 18]. Furthermore, [19] has reported little

10.21437/SpeechProsody.2016-91

variability in the duration of successive vowels. For SSBE, we thus predict high sensitivity to duration cues and a bias towards first syllable prominence perception. For SgE, we expect lower sensitivity to duration cues and no bias towards first or second syllable prominence perception.

- **Valais (VS) and Schaffhausen (SH) Swiss German:** There is little literature on the perception and production correlates of stress, particularly duration, in Swiss German (SwG). However**,** SH and VS SwG show quantity contrasts for nearly all vowels [20, 21], and the dialects exhibit similar stress patterns to Standard German, where word stress tends to be assigned to the lexical root, which is often placed on the first syllable [20, 22]. VS Swiss German has been reported to attract stress on first syllables to a greater extent [23]. For both dialects, we expect high sensitivity to duration cues and a bias towards first syllable prominence perception.

- Two other European languages with seemingly different prominence patterns were chosen: **Standard Danish** has variable stress and, as in SSBE, its location is partially predictable from phonological and morphological context [24]. Vowel length is distinctive, but stød – a distinctive voice quality – may lead to shorter duration in phonologically long vowels [25, 26]. On the other hand, **French** has predictable stress on the final syllables of accentual phrases [27, 28], and listeners have been described as stress 'deaf' in experimental tasks that included high phonetic variability and memory load, and thus required listeners to have an abstract phonological representation of stress [29, 2]. Yet, French listeners have been reported to be sensitive to the acoustic correlates of stress – including duration – in cognitively less demanding tasks [29, 2]. Furthermore, vowel length is not distinctive in French. We expect high sensitivity to duration cues but no bias to first or second syllable prominence perception for Danish. For French, we predict lower sensitivity but a bias towards second syllable prominence perception.

- Two tone languages were chosen to add to prosodic typological diversity. While **Hong Kong Cantonese** does not have lexical stress, it contrasts six lexical tones in which intrinsic duration differences have been observed [30, 31, 32]. The three level tones also have shorter allotones which appear only in stop-final syllable [33]. Duration has also been found to be a correlate of both the production and perception of prosodic focus [34, 35]. **Fuzhou** employs five tones which are characterized by pitch height and movement [36, 37] while duration and voice quality serve as secondary cues [38]. Fuzhou has a complex tone sandhi system: in a disyllabic word, the first syllable undergoes sandhi, which reduces its duration [39]. Fuzhou can thus be seen as having predictable final stress. For Cantonese we hypothesize high sensitivity to duration cues and an unspecified bias towards the first or second syllable; for Fuzhou, too, we expect high sensitivity but a clear bias towards the second syllable.

2. Method

2.1. Material

Stimuli were adopted from Kohler's study on prominence perception in German [1]. He extracted the first <ba> from a natural production of <baba> trochees by a male speaker whose native language is German. Kohler duplicated the syllable, monotonized the f0 of the vowels and linearly lengthened stop and vowel durations. Stop duration was 63 ms, vowel duration (incl. burst) 189 ms. Based on this control stimulus he created 5 stimulus series (pitch, intensity, duration, pitch+intensity, pitch+duration). For the present paper, we only report listeners' responses to the duration series. For the duration series, the following changes to the first and second syllable durations were applied, see Table 1:

Stimulus	1st syll.	2nd syll.
d13	+30%	-30%
d12	+20%	-20%
d11	+10%	-10%
d00	no change	no change
d21	-10%	+10%
d22	-20%	+20%
d23	-30%	+30%

Table 1: *Stimulus set*.

The first digit in d?? refers to increasing the duration of the first syllable (i.e. 1?) or the second syllable (i.e. 2?). The second digit refers to the magnitude of the increase (?1, ?2, or ?3). Thus, d13 shows a duration increase in the first syllable by 30% (complementarily, duration in the second syllable was decreased by 30%). <baba> is not a real word in any of the languages tested in the present study; all segments appear in a legal phonotactic combination for every language.

2.2. Subjects

We tested 80 listeners – 10 per language. On average, listeners were 30.5 years old (SD=13.9). 43 males (54%) and 37 females (46%) participated in the experiment, see Table 2.

Variety	Gender	Age
Cantonese	5f, 5m	M=23.3, SD=3.2
Danish	6f, 4m	M=30.2, SD=7.5
SgE	5f, 5m	M=34.8, SD=3.7
SSBE	5f, 5m	M=23.8, SD=6.6
French	5f, 5m	M=31.9, SD=4.7
Fuzhou	6f, 4m	M=62.2, SD=2.0
Swiss German (SH)	0f, 10m	M=18.0, SD=0.0
Swiss German (VS)	5f, 5m	M=20.2, SD=1.4

Table 2: *Distribution of subjects' age and gender*.

Cantonese listeners spoke English as an L2. Danish listeners were speakers of Sealand Standard Danish; SgE listeners were ethnically Chinese, who completed basic education in Singapore's bilingual education system. English was reported as the dominant language and Mandarin Chinese as the next most frequently used language. British English subjects were speakers of SSBE; French listeners spoke Parisian French. Fuzhou listeners were monolingual with minimal command of Mandarin. SwG listeners were monolingual speakers of the respective dialects. None of the subjects reported problems with hearing.

2.3 Procedure and data analysis

In a Praat [40] interface, subjects heard each stimulus once through high-quality headphones in a quiet room. The order of the series and of the stimuli in the series were randomized separately for each subject. Each stimulus occurred 5 times in randomized order in a series. Following the presentation of each stimulus, subjects decided which syllable they perceived as stronger by clicking the corresponding button on the experiment interface shown in Figure 2.

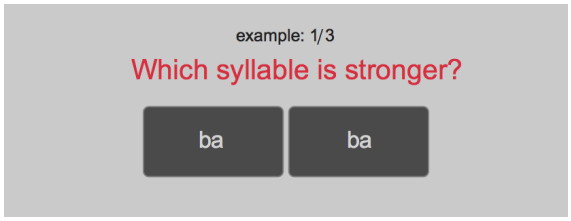

Figure 2: Experiment interface.

The data was analyzed using R [41]. If not indicated otherwise, we used generalized linear models (GLM) for the analysis. An effect of *item,* for example, was tested by model comparison between a full model (e.g. model_full<-glm(formula = response ~ item + age + gender, family = binomial(), data=French) and a reduced model excluding *item*.

3. Results

3.1. Effect of *variety* (left/right bias)

We tested for an effect of *variety* by examining the listeners' responses to baseline, i.e. non-manipulated, stimuli (d00). Figure 3 shows the relative proportions of responses by *variety*: first (red) and second (turquoise) syllable responses in ascending order of prominence perceived on the first syllable (dashed line=50%).

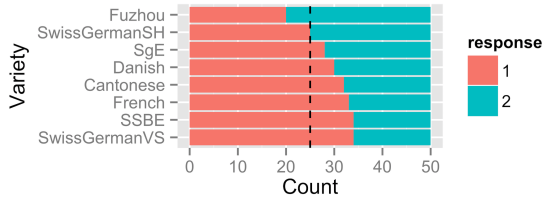

Figure 3: Responses *by* variety.

SwG (VS) and SSBE showed the most responses of prominence perceived on the first syllable (68%); Fuzhou listeners perceived most syllables as bearing second syllable prominence (60%). The GLM that included *age* and *gender* as factors did not reveal a significant effect of *variety*. A less conservative model, which excludes *age* and *gender* – factors that did not show significant effects to begin with – revealed a significant effect of *variety* (X^2(7)=-14, p=.048*).

3.2. Effect of *item* by *variety* (sensitivity)

We tested for an effect of *item* for every variety. Figure 4 shows the responses crossed for *items* and *variety*. Varieties are ordered by the proportions found in 3.1. For Fuzhou, the proportions of first and second syllable perceived prominence were most different between d13 (68% first, 32% second) and d22 (16%, 84%). Overall differences in proportions were significant (X^2(6)=-65, p<.0001*). SwG SH showed a stepwise progression from first to second syllable perceived prominence (X^2(6)=-65, p<.0001*). For SgE and Danish, responses cluster around 50%, i.e. no preference for first or second syllable, for virtually all stimuli. The models did not show a significant effect of *item*. Cantonese listeners showed a stepwise increase in perceiving the syllables as more prominent with an increase in duration (X^2(6)=-38, p<.0001*). French listeners exhibited a similar trend as SgE listeners, in that overall they seemed less sensitive to duration cues. Still, the French perceived d13 items (72%, 28%) differently from d23 items (44%, 56%) (X^2(6)=-21, p=.002*). On the whole, the model did not reveal a significant effect of *item*. SSBE listeners behaved differently from their SgE counterparts: they demonstrated the greatest sensitivity in d13 (92%, 8%) and revealed a drop from d21 (50%, 50%) to d22 items (20%, 80%) (X^2(6)=-120 p<.0001*). Swiss German VS listeners, too, showed a stepwise change from first to second syllable prominence perception when duration changed from the first to the second syllable (X^2(6)=-46 p<.0001*).

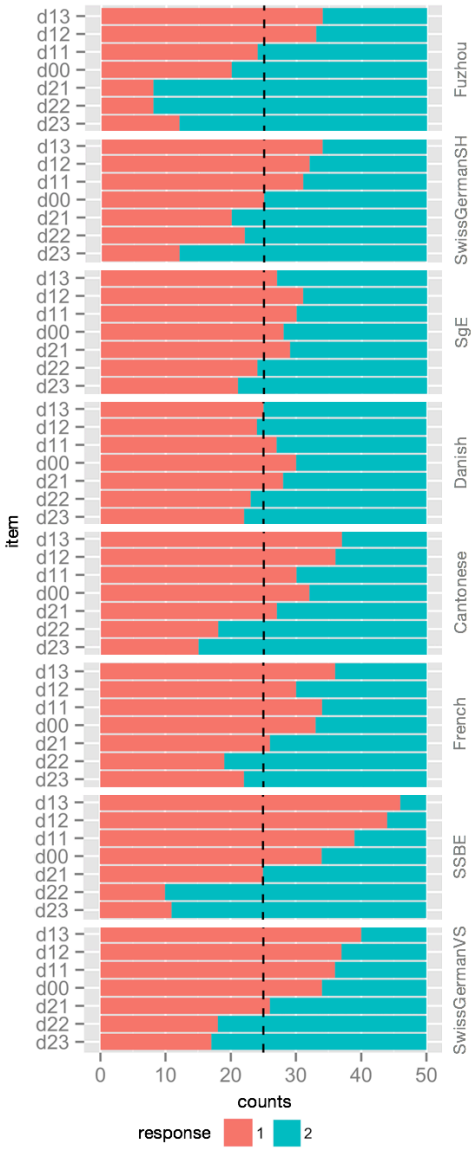

Figure 4: *Responses crossed for* item *and* variety.

4. Discussion & Conclusions

4.1. Effect of *variety* (left/right bias)

We predicted that the varieties tested would differ in their biases towards perceiving the first or second syllable as more prominent. Examining responses of the baseline stimuli only (d00), we found a marginally significant effect of *variety*. The small sample size within a language and the relatively large number of languages tested, which required correction for multiple comparisons, might have jointly raised the p-value.

Expectations for the individual languages stated in the introduction were partially met. Results for Fuzhou and SwG strongly corroborate our predictions. Predictable stress on the final syllable for Fuzhou seems to have translated into a bias towards perceiving the second syllable as more prominent. As for SwG, the strong first syllable prominence perception bias for the VS variety may be grounded in the variety's nearly ubiquitous first syllable stress in speech production, a relic from Old German stress rules. [1], too, found a strong bias towards first syllable prominence perception in Standard German (75%). As predicted, we found SSBE listeners to also have a distinct bias towards perceiving the first syllable as more prominent. A similar case seems to be found for the Danish listeners. For the French listeners, we expected a bias towards second syllable prominence perception, as stress usually falls on the final syllable; results showed, however, a trend to first syllable prominence perception in the baseline condition. According to [42], disyllabic French words may carry *emphatic* stress on the first syllable. Perhaps, the isolated logatomes used in the present study primed French listeners to perceive the stimuli as carrying *emphatic* stress, which may explain the high number of responses on first syllables. This hypothesis will have to be explored further. The lack of a bias for SgE may have to do with SgE speakers' tendency to place equal prominence on both syllables in disyllabic content words [14]. Cantonese listeners' lack of a bias may be explained in part by the absence of lexical stress and reduced vowels in Hong Kong Cantonese [31, 33].

The most relevant finding that emerges from the cross-linguistic analyses presented in 3.1 is that differences between varieties of the same language (e.g. SgE vs. SSBE or SwG SH vs. SwG VS) can be similar in magnitude to those found for typologically unrelated languages (e.g. Fuzhou vs. VS SwG). This is particularly true for the differences reported for the two SwG dialects. This finding is likely to be related to differences in stress rules mentioned earlier. For English, which is spoken widely, substantial variability of prosodic structure can be expected. SgE has been reported to demonstrate less durational variability in vocalic intervals [19]. This lack in durational variability may explain the listeners' lack of a bias towards perceiving the first or second syllable as more prominent in the baseline stimulus.

4.2. Effect of *item* by *variety* (sensitivity)

We predicted that listeners of all languages tested would be sensitive to duration cues, except for SgE and French. Results seem to meet these expectations only in part. The lack of sensitivity for SgE may have to do with a less salient vowel length distinction and less frequent vowel reductions in content words [14] than SSBE. Moreover, SgE speakers tend to make durational contrasts between vowels more on the phrasal than on the word-level, as manifested in phrase-final lengthening, for example [43]. French listeners seemed to be more sensitive to duration contrasts than expected. [29] reported French listeners to demonstrate difficulties in ABX stress discrimination tasks that presupposed an abstract phonological representation of stress on the part of the listeners. Yet, French listeners were able to discriminate non-words that differed only in position of stress in AX discrimination tasks, in which – compared to ABX tasks – memory load was less strenuous. When durational cues of stress were absent, 'stress-deafness' increased [2]. Fuzhou and Cantonese listeners' sensitivity to duration cues in the perception of prominence patterns may be explained by the languages' use of temporal information in tone production [31, 38]. Additionally, duration is used in the production and perception of prosodic focus in Cantonese [34, 35]. In the same way, it could be argued that listeners of both SwG dialects and SSBE were sensitive to duration cues as all three varieties feature quantity contrasts in vowels [12, 20, 21]. Contrary to expectations, this study found little sensitivity to duration cues for the Danish listeners. A potential explanation may be found in the interaction between vowel length and stød: stød may occur with phonologically long vowels and has the effect of shortening them. Studies of words in isolation have shown that long stød vowels are 25% shorter than long vowels without stød [26]. This could mean that vowels without manipulations may sound extra prominent to the Danish listeners, as they can be interpreted as having both – phonological length and stød.

These findings help us better understand how the prosodic structure of listeners' native languages affects their perception of prominence. The existence of different levels of sensitivity to one component of the prominence cues, e.g. duration, might explain cross-linguistic stress perception difficulties in second language learners. In the domain of f0, for example, English speakers find it difficult to perceive tonal contrasts of Mandarin [44]. Similarly, French listeners are known to have difficulties with perceiving Spanish lexical stress, despite being sensitive to the acoustic correlates of stress to some extent [29]. Several questions remain unanswered at present. The role of pre-final and phrase-final lengthening in the context of the present experiment needs to be examined further. It is likely that SSBE speakers, whose native language features phrase-final lengthening [45], may expect a longer second syllable to start with, and require larger durational differences in favor of the second syllable before perceiving it as prominent, while the same may not be expected of Danish listeners, for example [46]. There is abundant room for examining, in further depth, the contribution and interaction of duration, pitch, and intensity cues in the perception of prominence for listeners of typologically different languages. It is likely that changes in f0 will be more predictive for the perception of prominence than duration or intensity cues [7, 1]. At the same time, it is conceivable that the relative contribution of pitch, duration, and intensity cues in prominence perception is a salient characteristic for between-language differences [6].

5. Acknowledgements

The authors thank Francis Nolan (Cambridge) for helpful comments to previous versions of this manuscript. We are grateful to Klaus Kohler (Kiel) for providing us with stimuli created for a previous experiment and allowing us to use them in the present study. Adrian Leemann thanks the *Swiss National Science Foundation* for funding this research, grant Nr. P300P1_151210, http://p3.snf.ch/project-151210.

6. References

[1] K. J. Kohler, "The perception of prominence patterns," *Phonetica*, vol. 65, no. 4, pp. 257–269, 2008.

[2] S. Peperkamp, I. Vendelin, and E. Dupoux, "Perception of predictable stress: A cross-linguistic investigation," *Journal of Phonetics*, vol. 38, pp. 422–430, 2010.

[3] D. L. Bolinger, "A theory of pitch accent in English," *Word Journal of the International Linguistic Association*, vol. 14, no. 2–3, pp. 1–149, 1958.

[4] I. Lehiste, *Suprasegmentals*. Cambridge, MA: MIT Press, 1970.

[5] H. Fujisaki, K. Hirose, and M. Sugito, "Comparison of acoustic features of word accent in English and Japanese". *Journal of the Acoustical Society of Japan*, vol. 7, no. 1, pp. 57–63, 1986.

[6] J. Vaissière, "Language-independent prosodic features." in A. Cutler and D. R. Ladd (Eds.), *Prosody: Models and Measurement*. New York: Springer, pp. 53–66, 1983.

[7] G. Kochanski, E. Grabe, J. Coleman, and B. Rosner, "Loudness predicts prominence: Fundamental frequency lends little," *Journal of the Acoustical Society of America*, vol. 118, no. 2, pp. 1038–1054, 2005.

[8] D. B. Fry, "Duration and intensity as physical correlates of linguistic stress," *Journal of the Acoustical Society of America*, vol. 27, pp. 765–768, 1955.

[9] I. Lehiste and R. A. Fox, "Perception of prominence by Estonian and English listeners," *Language and Speech*, vol. 35, no. 4, pp. 419–434, 1992.

[10] A. Eriksson, E. Grabe, and H. Traunmüller, "Perception of syllable prominence by listeners with and without competence in the tested language," *Proceedings of Speech Prosody 2002, Aix-en-Provence*, 2002.

[11] A. Eriksson, "Syllable prominence: an experimental study," *Lingue e linguaggio*, vol. 14, no. 1, pp. 43-60, 2015.

[12] P. Roach, "British English: Received Pronunciation," *Journal of the International Phonetic Association*, vol. 34, no. 2, pp. 239–245, 2004.

[13] U. Domahs, I. Plag, and R. Carroll, "Word stress assignment in German, English and Dutch: quantity-sensitivity and extrametricality revisited," *The Journal of Comparative Germanic Linguistics*, vol. 17, no. 1, pp. 59-96, 2014.

[14] D. Deterding, "The Intonation of Singapore English," *Journal of the International Phonetic Association*, vol. 24 no. 2, pp. 61–72, 1994.

[15] A. Brown, "The staccato effect in the pronunciation of English in Malaysia and Singapore," in J. Foley (Ed.), *New Englishes: The case of Singapore*. Singapore: Singapore University Press, pp. 115–128, 1988.

[16] J. T. Platt and H. Weber, *English in Singapore and Malaysia: Status, features and functions*. Kuala Lumpur: Oxford University Press, 1980.

[17] M. W. J. Tay, "The phonology of educated Singapore English," *English Worldwide*, vol. 3, no. 2, pp. 135–145, 1982.

[18] R. K. Tongue, *The English of Singapore and Malaysia*. Singapore: Eastern Universities Press, 1974.

[19] E. L. Low, E. Grabe, and F. Nolan, "Quantitative characterisations of speech rhythm: syllable-timing in Singapore English," *Language and Speech*, vol. 43, pp. 377–401, 2000.

[20] J. Fleischer and S. Schmid, "Zurich German," *Journal of the International Phonetic Association*, vol. 36, no.2, pp. 243–253, 2006.

[21] G. Wanner, *Die Mundarten des Kantons Schaffhausen*. Frauenfeld: Steiner, 1941.

[22] A. Weber, *Zürichdeutsche Grammatik*. 3rd ed. Zürich: Rohr, 1987.

[23] E. Wipf, *Die Mundart von Visperterminen im Valais*. Frauenfeld: Huber, 1910.

[24] N. Grønnum, "Danish", *Journal of the International Phonetic Association*, vol. 28, pp 99–105, 1998.

[25] N. Grønnum, M. Vazquez-Larruscaín, and H. Basbøll, "Danish Stød: Laryngealization or tone", *Phonetica*, vol. 70, pp. 66–92, 2013.

[26] E. Fischer-Jørgensen, "Phonetic analysis of the stød in standard Danish," *Phonetica*, vol. 46, pp. 1-59, 1989.

[27] M. Grammont, Traité pratique de pronunciation française. Paris: Delagrave, 1934.

[28] P. Delattre, "Accent de mot et accent de groupe," *The French Review* vol. 13, no. 2, pp. 141–146, 1939.

[29] E. Dupoux, S. Peperkamp, and N. Sebastián-Gallés, "A destressing 'deafness' in French?," *Journal of Memory and Language* vol. 36, pp. 406–421, 2001.

[30] Y. Fok-Chan, *A perceptual study of tones in Cantonese*. Hong Kong: Hong Kong University Press, 1974.

[31] P. Rose, "Between and within-speaker variation in the fundamental frequency of Cantonese citation tones," in P. J. Davis and N. Fletcher (Eds.), *Vocal Fold Physiology – Controlling Complexity and Chaos*. San Diego: Singular Publishing Group, 1996

[32] E. Khouw and V. Ciocca, "Perceptual correlates of Cantonese tones," *Journal of Phonetics*, vol. 35, pp. 104–117, 2007.

[33] R. S. Bauer, S. Robert, and KP. K. Benedict, *Modern Cantonese phonology*. Berlin: Mouton de Gruyter, 1997.

[34] R. S. Bauer, K. H. Cheung, P. M. Cheung, and L. Ng, "Acoustic correlates of focus-stress in Hong Kong Cantonese," *Papers from the 11th Annual Meeting of the Southeast Asian Linguistics Society 2001, Mahidol, Bangkok, Arizona*, pp. 29–49, 2004.

[35] W. L. Wu and Y. Xu, "Prosodic focus in Hong Kong Cantonese without post-focus compression," *Proceedings of Speech Prosody 2010, Chicago, 2010*.

[36] C. Donohue, *Fuzhou tonal acoustics and tonology*. Munich: Lincom Europa, 2014.

[37] Y. Li, "Tone sandhi and tonal coarticulation in Fuzhou min". *Proceedings of ICPhS, 10.–14.08.2015, Glasgow, 2015*.

[38] C. Donohue, "The significance of 'secondary cues' for tonal identification in Fuzhou," Proceedings of the 17th ICPhS, Hong Kong, pp. 607–610, 2011.

[39] M. S. Wright, *A metrical approach to tone sandhi in Chinese dialects*. Doctoral dissertation, University of Massachusetts, Amherst, MA, 1983.

[40] P. Boersma and D. Weenink. *Praat: doing phonetics by computer*. Version 6.0.04. www.praat.org, 2015.

[41] R Core Team, *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Version 3.0.0, http://www.R-project.org, 2013.

[42] A. Di Cristo, "Intonation in French," in D. Hirst, A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages*. Cambridge: CUP, pp. 195-218, 1998.

[43] E. L. Low, "Is lexical stress placement different in Singapore English and British English?," in A. Brown, D. Deterding, and E. L. Low (Eds.), *The English Language in Singapore: Research on pronunciation*. Singapore: Singapore Association for Applied Linguistics, pp. 22–34, 2000.

[44] T. L. Gottfried and T. L. Suiter, "Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones," *Journal of Phonetics*, vol. 25, no. 2, pp. 207–231, 1997.

[45] D. H. Klatt, "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," *The Journal of the Acoustical Society of America*, vol. 59, no. 5, pp. 1208–1221, 1976.

[46] N. Grønnum, "Intonation in Danish", in D. Hirst and A. Di Christo (Eds.), *Intonation systems: A Survey of Twenty Languages*, pp. 131–151, 1998.