Do Working Memory and Autistic Traits Predict Prosody Perception?

Grace Chen-Hsiu Kuo
Department of Foreign Languages and Literatures, National Taiwan University
graciakuo@ntu.edu.tw

ABSTRACT

This study examines the role of the *working memory* and *autistic traits* in predicting the prosodic prominence and the perceptual strength of the prosodic boundaries. Subjects completed the working memory questionnaire and the autism-spectrum quotient questionnaire before they participated in two perception tasks: (a) Rapid Prosody Transcription Task, where the subjects mark the prominent words and the prosodic boundaries, and (b) Boundary Detection Task, where the subjects predict the upcoming prosodic boundary sizes in English, Swedish, and Taiwanese. Previous studies showed that native speakers and L2 learners use different acoustic measures as predictors of their perception of prosodic prominence and phrasing. This study takes intrinsic individual differences into account and finds that they are not necessarily to be the influential factors for L2 learners' prosody perception.

Keywords: working memory, autistic traits, L2, prosody perception, prominence, prosodic boundary

1. INTRODUCTION

Research on prosodic structures demonstrate that *prosodic prominence* and *prosodic phrasing* are the two key issues in the production, comprehension and acquisition of language [20, 24, 27]. Many factors have been found to contribute to the assignment of prosodic structures such as the syntactic structure, the semantic relation, and the pragmatic consideration. To interpret an utterance correctly, listeners must understand its prosodic structure by means of the acoustic cues. For instance, the acoustic cues related to greater perceived *prominence* include salient f0 contour, greater intensity, increased segmental duration, and flatter spectral tilt; speech rate and phrasal length, on the other hand, would influence prosodic *phrasing* in terms of the occurrence and the strength of the prosodic boundaries [23, 29, 13, 14].

The acquisition of a second language is heavily influenced by the learners' first language and their second language experience [10, 19]. Learners must learn to perceive the fine prosodic differences and establish a new system of stress, rhythmic patterns and intonation when they learn their L2's prosody.

According to [8] and [28], L2 learners perceived syllable prominence like native listeners in most cases, but when there was a conflict between expected prominence and produced prominence, the native speakers appeared to be influenced by their expectations according to their top-down knowledge whereas the L2 learners relied more on acoustic cues than the native speakers. Furthermore, the native and the non-native listeners attach weights to different acoustic cues in the listening experiments [18, 30].

The acquisition of a new sound system is often influenced by the cognitive, social, and psychological factors. Among the factors, the intrinsic individual differences are the focus in this study.

The specific goals of this study are: (a) to evaluate L2 learners' perception of prosodic prominence and prosodic boundaries, and (b) to discuss the correlations between the perceptual results and the measures of cognitive ability and personality assessment, namely *working memory* and *autistic traits*.

2. INTRINSIC INDIVIDUAL DIFFERENCES

Among the influential factors that are relevant to second language acquisition [17], the ones that emphasize individual differences have gained growing interest in second language development. *Working memory* and *autistic traits* are of interest here in that (a) high correlation was found between language performance and individual differences in working memory capacity; (b) disordered prosody was found to be a feature of impaired communication, and general population with higher level of autistic traits tended to perceive prosody differently from those with lower level of autistic traits [1, 9].

2.1 Working Memory

Working memory capacity is essential in performing L2 language development, L2 learners' reading proficiency and listening comprehension performance [16, 7]. The working memory questionnaire [26] is a 30-item, self-administered assessment. The items are divided into three domains – short-term storage, attention and executive. Each item is rated on a 5-point Likert scale, ranging from 1 to 5. Higher scores correspond to more difficulties or more complaints.

2.2. Autistic Traits

Previous studies demonstrated that autistic traits among general population affect gaze-triggered attention. Higher level of autistic traits are associated with variation in a number of speech perception effects, such as the Ganong Effect, the McGurk Effect, phonotactic effect, perceptual compensation for coarticulation, and atypical lateralization. The autistic traits have also been implicated in variation in speech production, discourse processing and prosodic processing [22, 25, 31, 12, 21, 32].

The autism-spectrum quotient (AQ) questionnaire [2] is a 50-item, non-diagnostic, self-administered questionnaire designed to assess if adults of normal intelligence have symptoms typical of individuals in the autism spectrum. Each item is rated on a 4-point Likert scale, ranging from 1 to 4. The items are divided into five dimensions – social skills, attention to detail, attention switching abilities, imagination, and communication skills. The communication skills domain is more related to discourse and has been used to predict online interpretation of prosody. Thus, the AQ score in this research refers to the score obtained from this domain. Individuals with higher AQ score exhibit weaker sensitivity to prosodic prominence in a cross-modal priming task [11]. There was an interaction between the types of the identified pitch accent and autistic traits; more specifically, individuals with higher levels of autistic traits identified fewer pitch accented words. However, the interaction between autistic trait and prosodic boundary perception is not apparent [4].

3. METHODS

3.1. Subjects

89 college students (average age: 20.9; 29 male and 60 female) completed two questionnaires and participate in two perception tasks. They were all , who were all Mandarin native speakers, English L2 learners,.

3.2. Procedures

Each subject first completed the AQ and WM questionnaires and then was seated in front of a laptop screen to hear stimuli with Praat for the Rapid Prosody Transcription Task [29]. In this task, subjects were required to identify prominent words and locations of prosodic juncture in running speech. The “prominent” words are defined as “words that the speaker put emphasis on and highlighted for us listeners”. The materials were three segments of the political speech from “Weekly Address” recorded by Barack Obama. Two trained phoneticians have previously annotated the materials using the ToBI conventions for Mainstream American English [30] in which prominence (i.e., pitch accent), and prosodic boundaries (i.e., break sizes) were labelled.

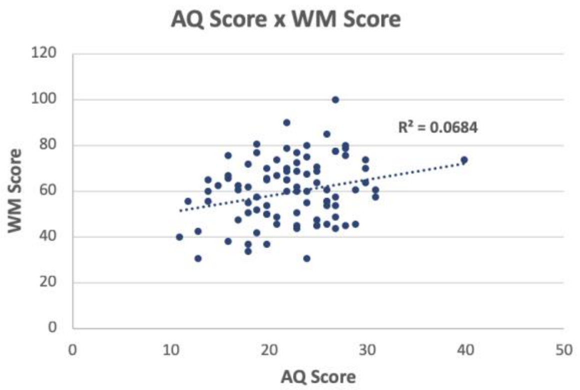
The second perception task (the Boundary Detection Task) was implemented on E-Prime 3.0. Subjects heard fragments of Taiwanese (native for some subjects), English (L2), and Swedish (foreign language) varying in length (2-second vs. one-syllable long), quality (low-pass filtered vs. unfiltered) and boundary type (word vs. phrase vs. sentence) via Sennheiser HD 380 pro headphone at a comfortable listening level. The unfiltered speech was extracted from a public speech or an interview (n=30 in each language). The low-pass filtered stimuli is adopted for Taiwanese especially to test whether listeners could perform well with segmental information removed. The subjects were asked to predict the upcoming boundary size by pressing the relevant buttons on the keyboard. Previous studies showed that listeners made better judgement when they heard fragments from a foreign language as opposed to fragments from a familiar language [5, 15]. There were 180 utterances in the stimuli in total (5 items x 3 Boundary Type x 3 Languages x 2 Lengths x 2 Quality). The correlation between the perceptual results and the intrinsic individual measures will be examined.

4. RESULTS & DISCUSSION

4.1. Questionnaires

Score results in Figure 1 showed that listeners’ sex, native language, and handedness did not lead to significant difference in AQ and WM scores. In addition, Shapiro-Wilk normality test showed that both AQ and WM scores are in normal distribution and are positively correlated with each other: subjects with greater level of autistic traits tend to have poorer working memory capacity.

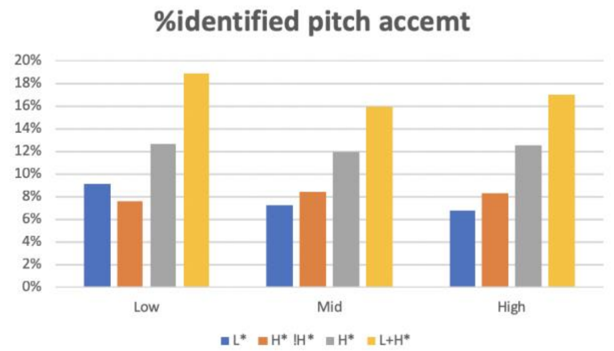
Figure 1: The correlation between AQ Score and WM Score



4.2. Rapid Prosody Transcription Task

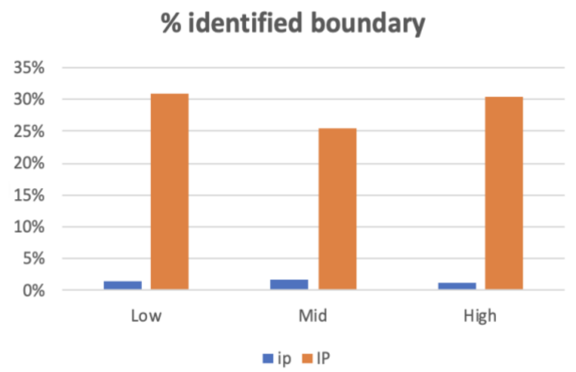
Subjects are divided into three groups based on their performance in the AQ score (language ability specific; not the overall score) and WM score – higher score means more autistic-like and poorer working memory capacity. Statistical results showed that significant difference was found across different pitch accents (only L*, H* !H*, H*, and L+H* are taken into consideration in this study) (as illustrated in Figure 2). Neither AQ score nor WM score were good predictors of L2 learners' prominence perception.

Figure 2: Percentage of words identified as prominent by L2 learners, as a function of ToBI pitch accent types. The learners were grouped by their performance in AQ and WM scores.



Statistical results showed that only "boundary types" result in significant difference; questionnaire scores can not help predict the perception of prosodic boundaries.

Figure 3: Percentage of prosodic boundaries identified by L2 learners, as function of ToBI boundary type. The learners were grouped by their performance in AQ and WM scores.

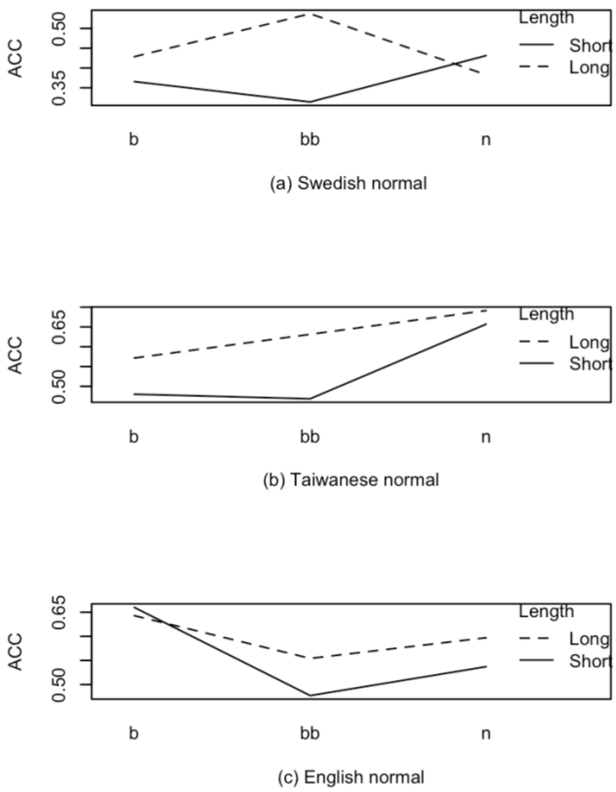


Previous studies showed that L2 learners and naïve listeners performed as well or even better than native speakers when it came to prosody perception task. However, the correctness in this study on identifying prominence and boundaries are 11.3% and 15.2% respectively (cf. native listeners' 32% and 25% in [6, 3]). In other words, working memory and autistic traits can't predict the L2 prosody perception.

4.3. Boundary Detection Task

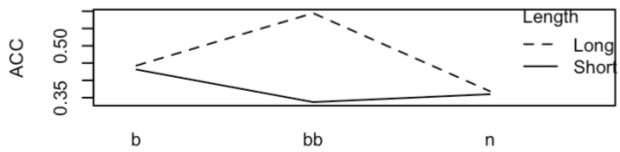
Listeners' correctness were entered into a series of two-way repeated measures ANOVA. For Swedish normal stimuli, significant effects were found in "Boundary Type" ($F(2, 138)=0.679$, $p < .05$), "Length" ($F(1, 69)=17.84$, $p < .05$), and their interaction ($F(2, 138)=17.97$, $p < .05$). For the Taiwanese normal stimuli, significant effects were found in "Boundary Type" ($F(2, 138)=8.302$, $p < .05$) and "Length" ($F(1, 69)=25.44$, $p < .05$). For English normal stimuli, significant effects were found in "Boundary Type" ($F(2, 138)=6.121$, $p < .05$) and "Length" ($F(1, 69)=5.321$, $p < .05$). Native speakers as well as L2 learners and foreign listeners are all able to tease three boundary types apart. The figures below are the interaction plots of Correctness (ACC), Boundary Type, and Length across the three languages.

Figure 4: The interaction plots of <u>unfiltered /normal</u> stimuli: b, bb, and n stand for "phrase boundary", "sentence boundary", and "word boundary" respectively.

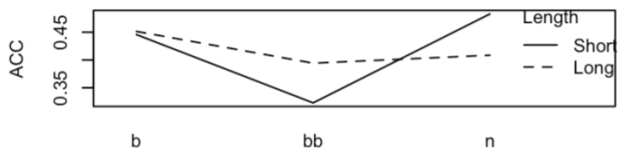


As shown in the figures, the subjects chose more correct answers when the fragments were longer, whether the language was their native, L2 or a foreign language. In addition, the accuracy of detecting the "word boundary" is relatively high across three languages. Interestingly, the detection of the "phrase boundary" in the L2 stimuli (English) seems to be more accurate than the detection of the "sentence boundary" which was unexpected.

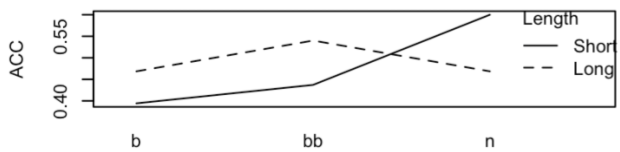
Figure 5: The interaction plots of filtered stimuli: b, bb, and n stand for "phrase boundary", "sentence boundary", and "word boundary" respectively.



(a) Swedish filtered



(b) Taiwanese filtered



(c) English filtered

Similarly, for the filtered stimuli, the accuracy of boundary detection was higher when the presentation of the stimuli was longer in all three languages, and it seems that listeners also had the tendency to choose "word boundary" as the answer when the fragments were short.

Moreover, significant correlations were found between AQ scores and the detection correctness in both normal and filtered speech, regardless of the language difference. Similar results were found in the correlations between the WM score and the detection correctness. These results suggest that intrinsic individual difference plays an important role in predicting the upcoming boundary size in an utterance, whether the listeners were familiar with the language or not. In order to understand which acoustic measures were helping the listeners make the judgment, further correlation analyses between the scores and the acoustic measures should be examined.

5. CONCLUSION

In this study, we find that *working memory* and *autistic traits* do not appear to be useful in Rapid Prosody Transcription Task, whereas they seem to play an important role in the Boundary Detection Task. The relation between the perceptual results (both *prominence* and *phrasing*) and the other domains in the individual difference assessment (e.g., *attention*-related ones) will be examined in future work.

6. REFERENCES

[1] Baddeley, A. 2003. Working memory and language: an overview. *Journal of Communication Disorders* 36, 189-208.

[2] Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. 2001. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5-17.

[3] Beckman, M. E., Ayers, G. M. 1994. Guidelines for ToBI Labelling.

[4] Bishop, J., Kuo, G. 2016. Do "autistic-like" personality traits predict prosody perception? *LabPhon 15 Satellite Workshop.*

[5] Carlson, R., Hirscheberg, J., Swerts, M. 2005. Cues to upcoming Swedish prosodic boundaries: subjective judgements studies and acoustic correlates. *Speech Communication* 46, 326-333.

[6] Cole, J., Mo, Y.-S., Baek, S.-D. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25, 1141-1177.

[7] Cziraky, L. 2008. Working memory and English as a second language listening comprehension tests: a latent variable approach. Thesis. The University of Hong Kong.

[8] Eriksson, A., Grabe, E., Traunmüller, H. 2002. Perception of syllable prominence by listeners with and without competence in the tested language. *Proceedings of Speech Prosody* 2002, Aix-en-Provence, 275-278.

[9] Ferreira, F., Karimi, H. 2015. Prosody, performance, and cognitive skill: evidence from individual differences. *Explicit and Implicit Prosody in Sentence Processing*. Studies in Theoretical Psycholinguistics 46, 119-132.

[10] Hopp, H., Schmid, N. S. 2013. Perceived foreign accent in first language attrition and second language acquisition: the impact of age of acquisition and bilingualism.

[11] Jun, S., Bishop, J. 2015. Priming implicit prosody: prosodic boundaries and individual differences. *Language and Speech*, 58(4), 459-473.

[12] Kingston, J., Rich, S., Shen, A., Sered, S. 2015. Is perception personal? *Proceedings of the 18th International Congress of Phonetic Sciences.*

[13] Krivokapic, J. 2010. Speech planning and prosodic phrase length. *Speech Planning and Dynamics*, 157-190.

[14] Krivokapic, J., Mooshammer, C., Tiede, M. 2013. Exploring prosodic boundaries: gradiency and categoricity of prosodic boundaries in articulation. In *Proceedings of meetings on Acoustics*, 19.

[15] Kuo, G. 2013. Perception and Acoustic Correlates of the Taiwanese Tone Sandhi Group. Ph. D. Dissertation. UCLA.

[16] Leikin, M., Share, D., Schwarts, M. 2005. Difficulties in L2 Hebrew reading in Russian-speaking second graders. *Reading and Writing* 18, 455-472.

[17] Moyer, A. 2013. *Foreign accent: the phenomenon of non-native speech.* Cambridge University Press.

[18] Pintér, G., Mizuguchi, S., Tateishi, J. 2014. Perceptino of prosodic prominence and boundaries by L1 and L2 speakers of English. *Interspeech* 2014.

[19] Saito, J., Trofimovich, R., Isaacs, T. 2016. Second language speech production: investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels.

[20] Shattuck-Hufnagel, S., Turk, A. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25, 193-247.

[21] Skewes, J., Jegindø, E. Gebauer, L. 2015. Perceptual inference and autistic traits. *Autism*, 19(3), 301-307.

[22] Stewart, M. E., Griffiths, T. D., Grube, M. 2015. Autistic traits and enhanced perceptual representation of pitch and time. *Journal of Autism and Developmental Disorders*.

[23] Terken, J., Hermes, D. 2000. The perception of prosodic prominence. *Prosody: theory and experiment*, 89-127. Kluwar Academic Publishers.

[24] Trofimovich, P., Baker, W. 2007. Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *SSLA* 28, 1-30.

[25] Ujiie, Y., Asai, T., Wakabayashi, A. 2015. The relationship between level of autistic traits and local bias in the context of the McGurk effect. *Frontiers in Psychology* 6,:891.

[26] Vallat-Azouvi, C., Pradat-Diehl, P. Azouvi, P. 2012. The working memory questionnaire: a scale to assess everyday life problems related to deficits of working memory in brain injured patients. *Neuropsychological Rehabilitation*, 1-16.

[27] Wagner, M., Watson, D. 2010. Experimental and theoretical advances in prosody: a review. *Language and Cognitive Process* 25, 905-945.

[28] Wagner, M. 2005. *Prosody and recursion*. Ph.D. Dissertation. MIT.

[29] Winters, S., O'Brien, M. G. 2013. Perceived accentedness and intelligibility: the relative contributions of f0 and duration. *Speech Communication* 55, 486-507.

[30] You, H.-J. 2012. Determining prominence and prosodic boundaries in Korean by non-expert rapid prosody transcription. *Speech prosody* 2012.

[31] Yu, A. 2010. Perceptual compensation is correlated with individuals' "autistic" traits: implications for models of sound change. *PLoS One* 5(8), e11950.

[32] Yu, A., Abrego-Collier, C., Sonderegger, M. 2013. Phonetic imitation from an individual difference perspective: subjective attitude, personality and "autistic" traits. *PLoS One* 8.