The Effect of Explicit Training on Comprehension of English Focus-to-Prosody Mapping by Indonesian Learners of English

Aprilia Sasmar Putri[1], Haoyan Ge[2], Aelish Hart[1], Virginia Yip[2] and Aoju Chen[1]

[1]Utrecht University, [2]The Open University of Hongkong

april.sasmar@ymail.com, hyge@cuhk.edu.hk, aelish.m.hart@gmail.com, vcymatthews@cuhk.edu.hk, aoju.chen@uu.nl

ABSTRACT

Past work has shown that even advanced L2 learners of English are unable to process the focus-to-prosody mapping in English in online comprehension if focus is prosodically encoded differently in their L1. Adopting a pre-test – explicit training – post-test paradigm, we investigated whether online prosodic comprehension is learnable to learners with such an L1. We focused on intermediate-advanced Indonesian learners of English because prosody is not used at all to mark focus in varieties of Indonesian. Participants' comprehension was assessed via a Reaction-Time based make-sense judgement task and their sensitivity to prominence was assessed via an acoustic perception task in pre- and post-tests. Between the tests, participants received one-hour explicit training on the use of prosody in English contrastive focus. Results show that participants went from not responding to mismatches between prosody and focus conditions to showing native-like responses after training, suggesting learnability of online prosodic comprehension.

Keywords: explicit training, focus-to-prosody mapping, Indonesian learners of English

1. INTRODUCTION

1.1. Focus and prosody

Focus is a key information structural category and typically refers to new information in a sentence [1]. Languages differ in the use of prosody in the realisation and interpretation of focus. In many languages, focus is associated with prosodic prominence, which can be achieved in different ways across languages. In languages like English and Dutch, prosodic prominence is achieved by placing a pitch accent on the focal word or part of the focal constitute. In Cantonese, a tonal language, duration is used to a much larger degree than pitch in focus marking [2-4]. Some languages do not use prosody for focus-marking purposes. For example, [5]

found that there is no prosodic focus marking at all in Ambonese Malay, and suggested that this may also be the case with other varieties of Indonesian, in line with earlier observations [6-10].

1.2. Comprehension of focus-to-prosody mapping in L2

Usually, prosody is not explicitly taught in English classrooms. At most, limited attention is paid to question intonation in English. It thus seems that prosody has to be acquired implicitly by learners of English. In this process, the similarities and differences between learners' L1 and English can cause rather different learning outcomes.

Ge et al. [4, 11] investigated the effect of L1 knowledge on comprehension of the focus-to-prosody mapping in L2 English by advanced Dutch and Cantonese learners of English, compared to native controls. Using a reaction time (RT)-paradigm [12], they found that the Dutch learners and the native controls were slower in comprehension when the focus-to-prosody mapping was inappropriate than when it was appropriate, whereas the comprehension of Cantonese learners was not affected by the appropriateness of the focus-to-prosody mapping. In a subsequent acoustic perception experiment, it was found that a failure to perceive prosodic prominence in the Cantonese learners cannot be attributed to their insensitivity to the focus-to-prosody mapping in English [13]. In fact, the Cantonese learners of English were as accurate as the native controls in the perceptual accuracy and even outperformed the native controls in the speed at which they detected the most prominent-sounding word in a sentence [13]. Together, these studies suggest that the differences in the focus-to-prosody mapping between L1 and L2 appear to hinder L2 learners to process this mapping on online sentence comprehension in the absence of explicit knowledge of this mapping in L2.

1.3. Past training studies on L2 prosody

Previous studies have shown that L2 learners *can* successfully learn nativelike use of prosody through specific training. For example, [14] used a three-component training programme (a mini-lecture on theoretical knowledge, listening and imitating with the help of visual displays of pitch contours; speaking and direct feedback) to teach Dutch learners of English to express sarcasm in English. They found clear effects of explicit training on Dutch learners' ability to use prosody to convey sarcasm in English. The same results were also obtained in several training studies concerning prosodic focus-marking. For example, [15] combined audio examples with visual displays of pitch contours in the training to aid participant's learning and perception of contrastive focus, implicational verb focus, and indirect insults marked as compliments and found an increase in correct interpretations of all these three functions of intonation after training.

1.4. The current study

As discussed in section 1.3, training can significantly improve L2 learners' offline interpretation and production of certain uses of prosody in English. A question arises as to whether explicit training in focus-to-prosody mapping in English can improve L2 learners' online comprehension. We have addressed this issue by examining the effect of short explicit training on the comprehension of focus-to-prosody mapping in English sentences with "only" by Indonesian learners of English. We focused on sentences with "only", following [4],[11] and [13]. Indonesian learners of English are particularly suitable to test the effect of explicit training because prosody is not used to realise focus in varieties of Indonesian. It has been suggested that the absence of prosodic focus marking in Indonesian can affect Indonesian listeners' understanding of focus-to-prosody mapping in English [5-10] and they may fail to recognise the structural connections between accentuation and focus in West Germanic languages [5]. The training concerns the use of prosody in marking contrastive focus in English as in sentences with 'only' (e.g. *The boys only painted the BOAT, not the door* vs. *The boys only PAINTED the boat, not repairing the boat*.)

Our three specific research questions (RQs) are as follows:

RQ1: How do Indonesian learners of English with different levels of proficiency in English process the focus-to-prosody mapping before training?

RQ2: If they fail to do so, is it because they cannot perceive prosodic prominence in the first place?

RQ3: Do Indonesian learners of English comprehend the focus-to-prosody interface in a more native-like way after training, compared to their performance before training?

Based on previous studies, we hypothesise that the comprehension of Indonesian learners of English will not be affected by the appropriateness of the focus-to-prosody mapping in English before training. Moreover, we hypothesise that Indonesian learners of English can detect the accentuation in English spoken sentences. Regarding the training effect, the successes reported in earlier training studies on the production and offline interpretation suggest that training may be beneficial to online comprehension. On the other hand, the success of training in those studies may not be generalisable to online comprehension, because online comprehension requires not only having the knowledge of the focus-to-prosody mapping in English but also the ability to automatically operating on that knowledge. We thus treat the third question as an exploratory question and suggest two possible scenarios. In one scenario, a short training focusing on establishing knowledge of prosodic focus marking in English is already sufficient to change patterns in online comprehension; in the other scenario, such a training may prove to be insufficient to alter patterns in online comprehension.

2. METHOD

This study was approved by the Ethical Assessment Committee Linguistics (ETCL) at Utrecht University and conducted following the guidelines of the ETCL.

2.1. Participants

Twenty-two Indonesian learners of English ($N = 22$, *Mean* age = 25.98, *SD* age = 2.32) participated in this study. They were post-graduate students at Utrecht University at the time of testing and are native speakers of Bahasa Indonesia. Prior to the experiment, they took the English version of the LexTALE test [16], which has been shown to be a reliable indicator for one's overall proficiency in English. The participants' LexTALE scores (*Mean*=75.31, *SD*=9.42) suggested that they had intermediate-advanced level of English proficiency.

2.2. The pre- and post-tests

The participants performed two tasks in both testing phases: a 'make-sense' task and an acoustic perception task. These tasks were adopted from [4] and [11]. We used the same stimuli and carried out the tasks using the same procedure, although the experiments were executed via the software ZEP [17], instead of E-prime.

2.2.1. Material and stimuli

The make-sense task. The make-sense task was conducted to address the first and third RQs. In this task, the participants listened to question-answer dialogues and judged whether the answer make sense by pressing a button of a button-box. The main measurement was the RTs, measured from the end of the answer sentence. If the participants could process the focus-to-prosody mapping, they would have longer RTs when the mapping was inappropriate than when the mapping was appropriate.

In the experimental stimuli, two factors were varied: the focus-to-prosody mapping (hereafter prosody) in the answer sentences (2 levels: contextually appropriate vs. contextually inappropriate prosody) and focus position (2 levels: focus on the object vs. focus on the verb). Combining the two factors led to four experimental conditions: object-focus with appropriate prosody, verb-focus with appropriate prosody, object-focus with inappropriate prosody, verb-focus with inappropriate prosody. Each condition was implemented on 40 question-answer dialogues. That is, each question-answer dialogue appeared in each experimental condition, as illustrated in (1), where accented words in in the bold letters.

(1)

    Experimental stimuli

    Story: The fox has some honey and ice cream. She was going to lick and freeze both of them. Then she changed her mind.

    a) object-focus with appropriate prosody
        i) Question : I wonder what the fox is licking.
        ii) Answer   : The fox is *only* licking the **HONEY**
    b) verb-focus with appropriate prosody
        i) Question : I wonder what the fox is doing with the honey.
        ii) Answer   : The fox is *only* **LICKING** the honey
    c) object-focus with inappropriate prosody
        i) Question : I wonder what the fox is licking.
        ii) Answer   : The fox is *only* **LICKING** the honey
    d) verb-focus with inappropriate prosody
        i) Question : I wonder what the fox is doing with the honey.
        ii) Answer   : The fox is *only* licking the **HONEY**

Moreover, 160 dialogues were included as fillers to distract the participants from the true purpose of the experiment. For fillers, following [4], the incorrect answer sentences were constructed half of time by either a semantic error or a pronunciation error. More specifically, the semantic errors in the answer sentences were built by referring to different VPs or NPs (e.g., mentioning *washing* when the question has said *examining*, or mentioning *violin* when the question had

said *window*). Meanwhile, the pronunciation errors were intentionally presented by mispronouncing the NPs in the answer sentence (e.g., pronouncing *blowtorch* as *blowtouch* or *bread* as *blead*). The other half of the fillers were error free. The prosody was also varied in both the fillers containing errors and the error-free fillers in the same way as in the experimental stimuli, half with appropriate prosody and half with inappropriate prosody.

The acoustic task. The acoustic task was conducted to address the second RQ, testing the perception of prosodic prominence in English sentences. In this task, the participants listened to declarative sentences (e.g. *The monkey is only eating the pancake*.), similar to the answer sentences used in the make-sense task, and judged which word sounded most prominent by choosing from a list of five options (e.g. *monkey*, *is*, *only*, *eating*, *pancake*).

In the experimental stimuli, only one factor was varied, namely position of accentuation (2 levels: accentuation on the verb vs. accentuation on the object). Each experimental condition were implemented on 40 experimental dialogues In addition, twelve fillers for each list were included, with two lists in total. The fillers were similar to the experimental stimuli, but only the subject noun was accentuated.

2.2.2. Procedure

A Latin Square design were employed to distribute the stimuli over different lists (four lists for the 'make-sense' task and two lists for the acoustic perception task). Each list was randomised and then randomly assigned to the participants. In doing so, the participants heard each questions-answer dialogue once in different conditions in the 'make-sense task' and each sentence once in the acoustic perception task. Each participant received the same list for the tasks in both pre- and post-tests. The participants performed the make-sense task first and then the acoustic task in a one-hour long test session in both the pre-test and the post-test. It took about an hour to finish the two tasks.

The participants were tested individually in a sound attenuated booth at the Linguistics Laboratory at Utrecht University. Prior to the test session, the participants were informed orally and in written about the procedure and asked for their consent. In both tasks, the sound was played over BeyerDynamic DT770 headphones. The only difference was that in the make-sense task the participants used the button box while in the acoustic task they used the computer keyboard to give their responses.

3

2.3. The training

Between the pre- and post-test, the participants received explicit training. The training included a pre-recorded 12-minute mini-lecture on prosodic focus marking in general and the use of prosody in marking contrastive focus in English in particular, and three exercises. The training session took place in a quiet room at the Linguistics laboratory at Utrecht University at least three hours after the pre-test and between 24 and 78 hours before the post-test. It was held in groups of three or four participants and lasted for about an hour.

The procedure of the training session is as follow: First, after a brief welcome from the experimenter, the participants were introduced to the main goal of the training session (i.e. obtaining a better understanding of English prosody). Then, they were given the mini-lecture delivered by a native speaker of English via a video clip. The participants were asked to pay attention to the video and not to take any notes during the lecture. After that, the participants worked on three exercises, one about the perception of correct or incorrect use of accentuation in marking contrastive focus, the second one on using accentuation to mark contrastive focus in speaking, and the third one on using information on accentuation to predict how a sentence may end (e.g. Max just finished moving house. Finally, he can unpack everything! He is sorting and emptying the bags and boxes. Oh wait, no. He's only emptying the BAGS…… A. not emptying the BOXES/ B. not SORTING the bags.) After each exercise, the participants checked their responses and corrected the errors in their responses with the help of the experimenter, who was trained to deliver the training properly.

3. ANALYSIS AND RESULTS

3.1. Online comprehension

The RT data from the 'YES' trials in the 'make-sense' task from both pre- and post-test sessions were analysed using a generalized linear mixed-effect model in R with the *lme4* package [18, 19]. The RTs did to have a normal distribution (W=0.83912, p< 0.001) and were log-transformed to reduce the non-normality of residuals. The predictor variables included SESSION (pre-test, post-test), PROSODY (contextually appropriate, contextually inappropriate), FOCUS (object-focus, verb-focus), and PROFICIENCY (the LexTALE scores); PARTICIPANTS was entered as a random factor.

Starting from an 'empty' model containing only the random factor, we added the main effects of the fixed factors, the two-way interactions between each two fixed factors, and the three-way interaction between all of them to the model in a stepwise fashion, building seven additional models. The best-fit model emerging from the

comparisons between models contained the main effects of SESSION ($\chi^2(1)=124.97$, $p<2.2e-16$), PROSODY ($\chi^2(1)=4.575$, $p<0.05$) and their interaction ($\chi^2(1)=14.506$, $p<0.0001$). Subsequent analysis showed that PROSODY only had an effect on comprehension in the post-test ($\chi^2(1)=4.9$, $p=0.027$): longer RT in the inappropriate prosody condition than in the appropriate prosody condition, as illustrated in Figure 1.

Figure 1. Mean log-RT in each prosody condition in pre- and post-tests.
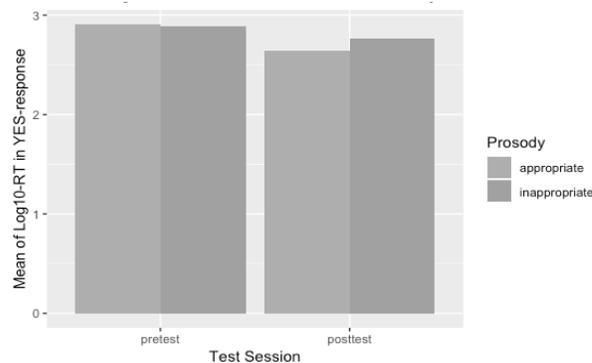
3.2. Prominence detection

The data from the acoustic perception task showed that the participants demonstrated a high accuracy rate regardless of position of accentuation and test-session (accent on the object: *Mean*=82.3%, *SD*=0.383; accent on the verb: *Mean*=95.9%, *SD*=0.199). This result suggested that Indonesian learners are sensitive to prosodic prominence in English.

4. DISCUSSION AND CONCLUSION

To conclude, we have found that the comprehension of the Indonesian learners of English were affected by the focus-to-prosody mapping only after the training, supporting our hypotheses. Their insensitivity to the focus-to-prosody mapping in the pre-test was not caused by a failure to perceive prosodic prominence. These results indicate that the short explicit training session was sufficient to alter how Indonesian learners responded to the focus-to-prosody mapping in online comprehension in L2 English.

Our study has provided first evidence on the learnability of native-like prosodic processing patterns. Future research is needed to find out how long the effect of training can retained and whether it is generalisable to other aspects of prosodic processing.

5. REFERENCES

[1] Halliday, M. A. (1967). Notes on transitivity and theme in English: Part 2. *Journal of linguistics*, *3*(2), 199-244.

[2] Chao, Y. R. (1947). *Cantonese primer*. Cambridge, MA: Harvard University Press.

[3] Wu, W. L., & Xu, Y. (2010). Prosodic focus in Hong Kong Cantonese without post-focus compression. In *Speech Prosody 2010-Fifth International Conference*.

[4] Ge, H.**,** Chen, A., Yip, V. (2018) L1 Effects on L2 comprehension of focus-to-prosody mapping: A comparison between Cantonese and Dutch learners of English. *Proceedings of the 9th International Conference on Speech Prosody 2018*, 187-191.

[5] Maskikit-Essed, R., & Gussenhoven, C. (2016). No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology*, *33*(2), 353-389.

[6] Odé, C. (1994). On the perception of prominence in Indonesian. In *Experimental studies of Indonesian prosody*. RUL. Vakgr. Talen en Culturen van Zuidoost-Azië en Oceanië.

[7] Van Zanten, E., & van Heuven, V. J. (1998). Word stress in Indonesian: Its communicative relevance. *Bijdragen tot de Taal-, Land-en Volkenkunde*, *154*(1), 129-149.

[8] van Zanten, E., & Goedemans, R. (2009). Prominence in Indonesian Stress, phrases, and boundaries. *Wacana*, *11*(2), 197-225.

[9] Goedemans, R., & Zanten, E. V. (2007). Stress and accent in Indonesian. *LOT Occasional series*, *9*, 35-62.

[10] Van Zanten, E., Goedemans, R., & Pacilly, J. (2003). The status of word stress in Indonesian. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, 151-178.

[11] Ge, H., Chen, A., Mulders, I. & Yip, V. (in prep). Processing prosodic focal information in second language: Evidence from eye movement in the Visual World Paradigm.

[12] Birch, S., & Clifton, C. E. (1995). Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech,* 38, 365-391.

[13] Ge, H. (2018). *The Processing of Focus in Cantonese-speaking Second Language Learners of English.* Unpublished PhD Dissertation. The Chinese University of Hong Kong.

[14] Smorenburg, B. J. L. (2015). *The effect of explicit training on the prosodic production of L2 sarcasm by Dutch learners of English* (Bachelor's thesis).

[15] Pincus, N. (2014). Teaching Intonation With Explicit Instruction and Audio-visual Techniques. INTED2013 Proceedings. 7th International Technology, Education and Development Conference, Valencia, Spain. 3428-435. Iated Digital Library. Retrieved from http://www.academia.edu/.

[16] Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. Behavior Research Methods, 44, 325-343.

[17] Veenker, T.J.G. (2018). The Zep Experiment Control Application (Version 6.1) [Computer software]. Beexy Behavioral Experiment Software. Retrieved from http://www.beexy.org/zep/.

[18] R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org.

[19] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.