Article

Effects of Instruction on Yes–No Responses to L2 Collocations

Junko Yamashita
Nagoya University
doi: http://dx.doi.org/10.7820/vli.v03.2.yamashita

Abstract

The lexical decision task (LDT), in which a participant makes dichotomous judgments on target letter strings, is an established method in psycholinguistic research to investigate the mental lexicon. With the expansion of research interests from single lexeme to collocations, second language (L2) researchers have started to use a similar judgment task at a phrasal level (referred to as a phrasal decision task or PDT in this paper). However, unlike the LDT, the PDT has not yet established a standard form of prompt, and variation has been observed in previous L2 studies. Hence, the purpose of this study was to examine effects of varying instructions on PDT performance. Three instructions (*acceptable/ commonly used/natural*) were tested with Japanese university students and native speakers of English, who were asked to make judgments on English word combinations. Examining responses to congruent (felicitous both in Japanese and English), incongruent (felicitous only in Japanese or English), and baseline items, the study identified some effects of instruction differences. However, these effects were not so strong as to obscure the expected cross-linguistic congruency effect. Therefore, this result has led to the conclusion that researchers have more freedom of instruction selection in the PDT, at least among the three examined in this study and to the extent that the congruency effect was measured by accuracy scores.

1 Introduction

The lexical decision task (LDT) is an established method in psycholinguistic research to gauge various research questions in order to understand the construct of the mental lexicon. In its most standard format, a participant looks at a letter string on a computer screen and makes a dichotomous judgment regarding whether it is a real word or not. Yes-items in this task (items that require an affirmative response) consist of real words, whereas No-items (those that require a negative response) comprise pseudowords (artificially made letter strings that usually adhere to spelling-sound correspondence rules). Both reaction times and error rates are measured because they provide different types of information. In brief, the error rate is an index of the degree of knowledge representation and the reaction time is an index of the efficiency of processing.

With the advancement of research from single lexeme to multiword units, second language (L2) researchers have started to use a similar judgment task at a

phrasal level (Jiang & Nekrasova, 2007; Wolter & Gyllstad, 2013; Yamashita & Jiang, 2010). This task is called the phrasal decision task (PDT) in this paper. In the PDT, Yes-items are felicitous (real, existing, and legitimate) phrases (e.g., *as soon as, take a bath*) and No-items are syntactically and/or semantically infelicitous word combinations (e.g., *is sun the*, *hear blood*). Unlike the LDT that has a standard form of prompt (*Is the letter string a real word or not?*), variations have been observed in the instruction for the PDT or a similar task in L2 research practice. For example, Laufer (2003) embedded collocations in a sentence (e.g., *I closed the TV.*) and asked whether sentences were *correct*. Jiang and Nekrasova (2007) examined formulaic sequences (e.g., *as a result*) and inquired whether word sequences were *grammatical*. Yamashita and Jiang (2010) and Wolter and Gyllstad (2013) both researched collocations, but the former asked whether word combinations were *acceptable* and the latter asked whether they were *commonly used*. This variation may be due to this task's relatively new history in the L2 research, but instructional differences may have impacts on judgmental performances. The primary purpose of this paper is to examine effects of different instructions on L2 PDT.

To proceed, it is imperative that the target structure be clarified because there are distinct types of multiword units. This paper focuses on collocations: lexical patterns that are determined by conventions within the language rather than by a set of rigid rules. Countering idioms, collocations are looser lexical patterns in that each constituent word can co-occur naturally with many other words (e.g., *heavy stone, heavy rain, fine rain*). However, the combination is not totally free and the language imposes restrictions that are sometimes arbitrary and difficult to explain (e.g., *high mountain* vs. *tall mountain*). Thus, collocations are normally acquired through exposure to the target language rather than by applying memorized rules. In addition to the importance of input, past studies indicated the lasting influence of first language (L1) on the acquisition of L2 collocations (Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010). In Yamashita and Jiang's study, for example, lower proficiency L2 learners were slower and made more errors in their responses to incongruent collocations (those that do not have L1 equivalents) than to congruent collocations (those that have equivalents in L1); higher proficiency L2 learners, on the other hand, also made more errors in incongruent collocations but they improved their reaction times and no measurable difference was observed between congruent and incongruent collocations. Thus, the L1–L2 congruency affects the acquisition of L2 collocations. This paper calls this L1 influence a congruency effect.

To investigate the effects of instruction on the L2 PDT, the current study drew on the congruency effect as a benchmark. Namely, if the PDT has successfully demonstrated the congruency effect, it is interpreted that the instruction has validly served its purpose and detected the impact of collocation congruency on the acquisition of L2 collocations. In so doing, the present study has expanded the congruency effect to infelicitous L2 word combinations that are felicitous in their translation in L1. For example, *sweet judgment* (*amai handan*) and *wide mind* (*hiroi kokoro*) represent such word combinations in the case of L1-Japanese and L2-English. When such expressions are compared with baseline word combinations, it is expected that the former would be more incorrectly identified as felicitous L2 expressions as a result of L1 influence. Although this is a likely scenario on the basis of the documentation of L2 learners' production of unnatural L2 collocations

that could be traced back to their L1 (e.g., Laufer & Waldman, 2011; Nesselhauf, 2005), the congruency effect has not been identified with the PDT by using this type of item. Hence, it was of additional interest to see the congruency effect on infelicitous L2 word combinations. Unlike mainstream psycholinguistic research, however, this study only measured accuracies. Therefore, the congruency effect is limited to accuracy measures in this paper.

The following research question guided the study: Does the form of the instruction affect the congruency effect? Three formulations of instruction were tested (*acceptable, commonly used,* and *natural*) because these were all plausible forms of instruction to investigate collocational knowledge with the PDT.

2 Method

2.1 Participants

Participants were 114 L1-Japanese university students learning English as a foreign language (EFL) and 24 native speakers of English (NS). The NS group consisted of undergraduates, postgraduates, and teachers of English at universities, who resided in Japan at the time of testing. Their self-evaluation of Japanese proficiency on a scale from 0 (no knowledge) to 10 (native speaker level) averaged 5.7 (SD = 2.1).

2.2 Instruments

Linguistic stimuli were adjective–noun combinations, which were either adopted from past studies (Yamashita & Jiang, 2010; Wolter & Yamashita, 2014) or created for this study. To ensure that all words would be known to the participants, the majority of constituent words were within the most frequent 3000-word level in the JACET List of 8000 Basic Words (the standard high school level in the Japanese education system). There were four categories of items with 24 stimuli in each: JE (felicitous expressions both in English and Japanese), E-only (felicitous in English but infelicitous in Japanese when they are literally translated), J-only (infelicitous in English but felicitous in Japanese when they are literally translated), and baseline (infelicitous either in English or in Japanese). Table 1 shows sample items in each category.

Items were controlled for five features: lengths, adjective frequencies, noun frequencies, collocation frequencies, and MI scores (values of the last four features were obtained from the Corpus of Contemporary American English). The entire set of 96 items was divided into three lists, each of which was combined with one of the instructions. There were no significant differences in the five item features across the lists, either. The order of instruction was counterbalanced, and items were randomized within a list. The following instructions were used. They were in English for the NS and in Japanese for the EFL learners.

Do you think the following 32 expressions are acceptable in English?
Do you think the following 32 expressions are commonly used in English?
Do you think the following 32 expressions are natural in English?

Table 1. Sample Items

Category	Sample items
E-only	bad debt, brief moment, careful attention, early bird, flat rate, slow learner, wet weather
JE	true love, cold tea, deep sleep, front door, economic growth, tragic death, high salary
J-only	dark society, frank question, light content, quiet war, straight heart, weak rain, wide mind
Baseline	abstract fruits, medical key, male table, formal ice, nervous papers, wild pen, calm eggs

The EFL learners additionally provided their confidence rating on each item and answered a posttest open-ended questionnaire on how they made judgments according to different instructions. Due to space limitations, however, only some of the questionnaire answers are discussed in this paper.

2.3 Procedure

The test was given collectively to the EFL learners in their English classes and individually to the NS.

3 Results

Descriptive statistics of test results is summarized in Table 2. After square root and log transformations were applied to the EFL and NS data, respectively, to improve distribution where necessary, 3×2 (instruction $\times$ item type) repeated-measures ANOVAs were run with two response types (Yes/No) and two participant groups separately. When the sphericity was violated, Greenhouse-Geiser correction was adopted.

For the real English collocations (Yes-items), ceiling effects were evident in NS's responses (all $> 90\%$), which lead, unsurprisingly, to insignificant effects of any independent variables. The EFL learners, in contrast, displayed lower and more varied accuracies across conditions ($50 \sim 81\%$), and three significant effects were identified: a main effect of instruction, $F(2, 214) = 89.49$, $\eta_p^2 = .44$, $p < .001$, a main effect of item type, $F(1, 113) = 375.67$, $\eta_p^2 = .77$, $p < .001$, and the interaction between instruction and item type, $F(2, 226) = 83.41$, $\eta_p^2 = .43$, $p < .001$. According to Bonferroni-adjusted post hoc multiple comparisons, the EFL learners were consistently more accurate on the JE than E-only items, demonstrating the expected congruency effect. Moreover, among the three instructions under the E-only condition, they were more accurate with the instruction containing *acceptable* (66%) than with the other two (52%, 50%).

With regard to illegitimate phrases (No-items), NS's accuracy decreased in the J-only items ($53 \sim 76\%$), which resulted in a main effect of instruction,

Table 2. Descriptive Statistics of Accuracy

		EFL			NS		
	Item type	Mean	SD	% Mean	Mean	SD	% Mean
Yes-items							
Acceptable	E-only	5.27	1.83	.66	7.63	0.71	.95
	JE	6.01	2.02	.75	7.71	1.04	.96
Commonly used	E-only	4.18	1.56	.52	7.75	0.44	.97
	JE	6.39	1.41	.80	7.75	0.53	.97
Natural	E-only	4.00	1.85	.50	7.33	1.61	.92
	JE	6.47	1.35	.81	7.50	1.14	.94
No-items							
Acceptable	J-only	4.44	1.76	.55	5.17	2.51	.65
	Baseline	6.39	1.44	.80	7.00	1.69	.88
Commonly used	J-only	4.66	1.86	.58	6.04	1.68	.76
	Baseline	6.57	1.47	.82	7.54	0.93	.94
Natural	J-only	4.70	1.55	.59	4.21	1.84	.53
	Baseline	6.64	1.48	.83	7.42	1.47	.93

Note: The maximum achievable score is 8.

$F(2, 46) = 5.85$, $\eta_p^2 = .20$, $p = .005$, a main effect of item type, $F(1, 23) = 295.21$, $\eta_p^2 = .93$, $p < .001$, and the interaction between the two, $F(2, 46) = 5.27$, $\eta_p^2 = .19$, $p = .009$. Post hoc analysis indicated that the NS were consistently more accurate on the baseline (88 ∼ 94%) than J-only items, suggesting the impact of Japanese collocations. In the J-only condition, they were significantly more accurate with the instruction containing *commonly used* (76%) than with that using *natural* (53%). Their accuracy with the instruction including *acceptable* (65%) was located between these, and manifested no significant difference from either of them. As for the EFL learners, only a main effect of item type was significant, $F(1, 113) = 532.09$, $\eta_p^2 = .83$, $p < .001$, indicating that they responded more correctly to baseline items (80 ∼ 83%) than J-only items (55 ∼ 59%) under all instructions.

4 Discussion

Throughout this study, the EFL learners displayed the congruency effect by both wrongly rejecting the E-only items and wrongly accepting the J-only items. The NS were, unsurprisingly, highly accurate in accepting the real English collocations, but interestingly they were also wrongly accepting J-only items more than the baseline, suggesting an effect of Japanese collocations. Although the result from J-only items aligned with the expectation for the EFL group, why did the NS say yes to non-existing English word combinations? A most likely reason is the influence from their L2 to L1 (Laufer, 2003) because all NS in this study had been exposed to the Japanese language. A piece of informal evidence is an NS participant who correctly suspected that some items were direct translations from Japanese, which suggests that, at least for this person, Japanese collocations were activated during the task in English. However, the story may not be as simple as this. In Wolter and Yamashita (2014), NS who were not exposed to the Japanese

language demonstrated the Japanese influence by reacting to J-only items more quickly than the baseline. Although the item type and the task were distinct between the studies, it may be the case that lexical patterns that exist in one language may make more sense than artificially made non-sense word combinations. This question, however, is beyond the scope of this paper and should be further explored in the future.

Instruction differences affected the magnitude of accuracy on two occasions. First, the EFL learners replied more affirmatively (thus correctly) to the E-only items with the instruction containing *acceptable* than that containing either *commonly used* or *natural*. Precise reasons are not known but the questionnaire responses aligned with this trend. The number of students who replied that their judgments did not differ across different instruction conditions was 23 (all three instructions), 2 (*acceptable* and *commonly used*), 2 (*acceptable* and *natural*), and 26 (*commonly used* and *natural*). These figures suggest that *commonly used* and *natural* tended to instigate similar judgment processes among the students. Second, the NS responded more affirmatively (thus wrongly) to the J-only items with the instruction using *natural* than that including *commonly used*. Again precise reasons are not known, but this might be because the NS had been daily exposed to English lexical patterns that the Japanese tend to mistakenly produce and may have developed such tolerance as to perceive them to be *natural*, but they may still somehow feel more strongly that such word combinations are not quite *commonly used* by native speakers.

Despite the differences caused by varying instructions, the congruency effect stayed intact in the EFL group. This result suggests that researchers could choose more freely from among these instructions when their purpose is to examine the congruency effect on the acquisition of L2 collocations, and lend support to the previous studies that used different instructions with the PDT but identified the congruency effect (Yamashita & Jiang, 2010; Wolter & Gyllstad, 2013).

However, this discussion needs to be received with caution because instruction differences showed some effects on the performance, depending on the type of items and participants. These minor differences within the realm of this study may have serious effects depending on the purposes of future studies. Moreover, the current study only investigated accuracy scores, thus it has no implication for reaction times. A similar study examining reaction times is recommended.

5 Conclusion

The purpose of this study was to examine the effects of different instructions used with the PDT, a relatively new task that has appeared in L2 psycholinguistic research. With the growing interest in the acquisition of collocations, this task may become one of the standard tasks in the field. This study indicated that, despite some differences, all three tested instructions captured the congruency effect when it was measured by accuracy scores. Hence, at least for this type of data, researchers have freedom in instruction selection for examining the congruency effect. In addition, this study demonstrated the congruency effect not only in Yes-items as did previous studies, but also in No-items, thus it has expanded the scope of linguistic items where the congruency effect makes its appearance. Evidence of the

congruency effect in infelicitous L2 word combinations is limited and further scrutiny in this area may help deepen our understanding of the acquisition of L2 lexical patterns.

References

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, *91*, 433–445. doi:10.1111/j.1540-4781.2007.00589.x

Laufer, B. (2003). The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 19–31). Clevedon: Multilingual Matters.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61, 647–672. doi:10.1111/j.1467-9922.2010.00621.x

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, the Netherlands: John Benjamins.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32, 430–449. doi:10.1093/applin/amr011

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, 35, 451–482. doi:10.1017/S0272263113000107

Wolter, B., & Yamashita, J. (2014). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*. doi:10.1017/S0142716414000113

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647–668.