OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

https://www.castledown.com/journals/tltl/

Technology in Language Teaching & Learning, 5(1), 1–22 (2023)
https://doi.org/10.29140/tltl.v5n1.1045

The Relationship Between Interaction Moves During Text-Based SCMC and L2 Vocabulary Learning Efficiency

LEANDER S. HUGHES

Saitama University, Japan
hughes@mail.saitama-u.ac.jp

Abstract

Previous research comparing communicative to non-communicative computer-mediated tasks requiring output production found an advantage in L2 vocabulary learning efficiency in favor of the communicative task (Hughes, 2023). The present study analyzes the chat data from Hughes (2023) to determine what features of interaction may have contributed to the advantage in L2 learning efficiency. Notable features included the negotiation of meaning, paraphrasing keywords, and context-elaborative question formulation. Of these, elaborative question formulation was found to correlate most strongly with L2 learning efficiency. Furthermore, a significant difference in L2 learning efficiency was found between participants who had engaged in elaborative question formulation and those who had not in favor of those who had. Learners who had engaged in elaborative question formulation also demonstrated a marginally significant advantage in L2 learning efficiency over participants who completed the non-communicative task. The implications of these results are discussed from a cognitivist perspective.

Keywords: interaction, output, synchronous computer-mediated communication (SCMC), negotiation of meaning (NoM), negotiation for meaning (NfM), Involvement Load Hypothesis (ILH)

Introduction

While both theory (e.g., Gass & Mackey, 2020; Long, 1996; Loewen & Sato, 2018) and evidence (e.g., de la Fuente, 2002; Ellis & He, 1999; Keck et al., 2006; Tare et al., 2014) point to the efficacy of communicative interaction in promoting second language acquisition (SLA), there remains little direct evidence linking specific interaction moves in learner-learner communication to actual target language gains achieved by those learners. Previously, Hughes (2023) found that learners who undertook a communicative task made statistically equivalent vocabulary gains compared to those who did a non-communicative task that required output. However, the learners in the group that communicated produced significantly fewer words than those in the non-communicative group. A significant correlation was also found between output wordcount and vocabulary gains, suggesting that output production played a critical role in the learning process. These results indicate that the group that

Data Availability Statement: *All relevant data are in this paper.*

completed the communicative task learned more efficiently than the group that did not. The present study adopts a cognitivist perspective and reexamines the data from Hughes (2023) in order to investigate for a relationship between specific communicative behaviors ("interaction moves") and vocabulary learning efficiency.

The role of "involvement" in communicative and non-communicative language acquisition

It can be argued that, at its essence, to learn is to form or strengthen a cognitive association between two or more pieces of information (be they perceptual, emotional, or conceptual) in memory, such that a future encounter with one has an increased probability of activating the other (Ellis & Wulff, 2019; James, 1950). Of course, merely forming an association between the form of a word and its meaning as, for example, a brief L1 translation, provides only shallow knowledge of that word (Nation, 2001). To really "know" the word requires further encounters to establish and strengthen connections between not only that word's form and meaning, but also its usage, register, connotations, restrictions, and related target language words (Ellis, 2019; Nation, 2001; Nation, 2019). Cognitive processing to build and strengthen these connections has been referred to as "deep" or "generative" processing (Craik & Lockhart, 1972; Wittrock, 1974) and more recently as "involvement" (Laufer & Hulstijn, 2001). Laufer and Hulstijn's Involvement Load Hypothesis (ILH) posits that this processing is promoted by language learning tasks that encourage a sense of need to comprehend and use target language items, search or effort expended to find the meaning of those items, and evaluation which "entails a comparison of a given word with other words, a specific meaning of a word with its other meanings, or combining the words with other words" (p. 14). The ILH was developed primarily to elucidate vocabulary acquisition and has garnered recent support from meta-analyses by Liu and Reynolds (2022) and Yanagisawa and Webb (2022). The studies found that, in particular, the factors of evaluation as defined above and need to use the target language played a significant role in vocabulary acquisition. The remainder of this section discusses how the ILH, and particularly the concepts of need and evaluation, can be applied in conjunction with a wider cognitive-psycholinguistic conception of SLA to better understand how output production and communication may promote language acquisition.

Evaluation of target language items can be carried out in many ways and at different levels of cognitive load. A "lite" evaluative task might feature a fill-in-the-blank reading activity with keyword choices and their translations wherein learners compare the form-meaning choices with reference to the contexts around each blank. A number of studies have found evidence for the effectiveness of such tasks (e.g., Rott et al., 2002; Rabie-Ahmed & Mohamed, 2022; Yanagisawa et al., 2020).

Tasks involving output production should lead to deeper evaluation by engaging productive language faculties, which are hypothesized to be inactive during input-only tasks (Swain, 2005). Having to produce language may additionally increase learners' sense of need to comprehend and use the language (Laufer & Hulstijn, 2001), further promoting cognitive involvement. Studies on the effect of tasks requiring output production have found evidence for their effectiveness in promoting language acquisition both for tasks involving communication (Ellis & He, 1999; de la Fuente, 2002; Hughes, 2023; Keck et al., 2006; Tare et al, 2014) and non-communicative tasks (Hughes, 2023; Izumi, 2002; Joe, 1998; Keating, 2008; Kim, 2008; Leeser, 2008; Nguyen & Boers, 2019; Rott, 2004; Tare et al., 2014; Zaki & Ellis, 1999). Particularly relevant to this study is the finding by Nguyen and Boers (2019) that summarizing target input led to significant uptake in new vocabulary from that input. Summarizing can be considered a type of paraphrasing in which the same meaning is expressed in different words, a process requiring substantial evaluation on the part of the learner as well as the need to comprehend and use the target language.

Communication may add further opportunities for language learning through, for example, enabling the negotiation of (or for) meaning (NoM) (Gass & Mackey, 2020; Long, 1996; Loewen & Sato, 2018), as demonstrated in this exchange between a student (S) and teacher (T) working on an information gap task during a study by Ellis et al. (1994, p. 476):

> T: We have an apple. And I'd like you to put the apple in the sink.
> S: What is the sink?

T: Sink is a place to wash dishes. It's a hole where you wash dishes.

Here S encounters a word that they need to know in order to successfully complete the task ("sink"). S may also be experiencing a general sense of need to communicate effectively with T. S therefore initiates NoM through a clarification check (Gass & Mackey, 2020) to find out what the word "sink" means. This could be considered a kind of search or effort exerted to obtain the meaning of the target item. T responds by explaining what a sink is in the target language. To successfully complete the task, S will have to consider the meanings of the words spoken by T with reference to the task materials (a picture of a room) to figure out which object pictured is the sink, thereby engaging in a kind of evaluation with regard to the target item. Thus, through enabling opportunities for NoM, communication may promote language learning and has been supported by empirical research (e.g., de la Fuente, 2002, Ellis & He, 1999; Ellis et al., 1994; Keck et al., 2006; Mackey, 1999).

Communicative interaction additionally enables the negotiation of form (NoF), in which learners' awareness of problems with a particular form (grammatical, lexical, or otherwise) is raised through feedback or correction and language related episodes (LREs) (Ellis, 2017; Schmidt, 1990; Swain & Lapkin, 1998). NoF also involves evaluation as the form in question is compared to other possible forms with reference to the context in addition to a sense of need to resolve the problem with the form. It may also involve searching to obtain the correct form. Research has found some empirical support for NoF in promoting language acquisition (e.g., Loewen, 2005; Watanabe & Swain, 2007; Williams, 2001; Yilmaz, 2016).

To summarize, tasks requiring output production should induce greater involvement and consequent language acquisition than input-only tasks. Likewise, tasks requiring communication should offer greater opportunities for acquisition than non-communicative tasks through enabling communicative interaction moves here defined as behaviors that are either made possible or more probable by the presence of an interlocutor and involve the attempt to actively convey meaning to achieve a particular communicative aim. This aim could be to make a prior message from the interlocutor more comprehensible, as in the case of NoM, or to explore a linguistic issue (NoF), or simply express the meaning of the content being discussed (e.g., paraphrasing). Such interaction moves may further encourage cognitive involvement while potentially focusing the learner's attention on relevant yet-to-be-acquired linguistic items, increasing the likelihood of subsequent acquisition.

The benefits of communicative compared to non-communicative output production

A number of empirical studies have investigated input-only versus communicative tasks involving negotiated input (e.g., Ellis et al., 1994; Mackey, 1999) as well as communicative tasks involving only negotiated input versus those involving both negotiated input and negotiated output (e.g., Ellis & He, 1999; de la Fuente, 2002). However, few studies have attempted to discern the benefits of interaction by comparing communicative tasks to non-communicative tasks that require a similar amount and type of output production while controlling for input. Therefore, it remains unclear whether the apparent benefits of interaction discovered by previous studies were due to communication or simply to the additional output produced during that communication. To the author's knowledge, only two studies so far have experimentally compared communicative tasks to non-communicative tasks requiring output production while controlling for input and attempting to elicit an equal amount of output from the groups compared (Hughes, 2023; Tare et al. 2014). In both studies, the interactive tasks were completed through text-based synchronous computer-mediated communication (SCMC) while the individual tasks were also completed online via computer (see Hughes, 2022 for a review of research on SLA through text-based SCMC).

Tare et al. (2014) conducted a quasi-experimental study with 45 intermediate learners of Russian at a language training institute in the U.S. The study involved two intact classes, one which was assigned interactive tasks and the other non-interactive tasks that were estimated to require a similar amount of output production. The different versions of the tasks included the same input and were designed to take approximately 20 minutes to complete. The communicative versions of the tasks involved information, opinion, and reasoning gaps, while the individual study versions were designed such that their focus and aims were as similar as possible to the interactive tasks despite not involving commu-

nication. Participants completed three tasks per week outside of class for six weeks. Pre-post test results showed a significant difference in vocabulary gains in favor of the group that communicated.

A follow-up study by Golonka et al. (2017) analyzing the chat data from Tare et al. (2014) attempted to uncover what interaction moves may have contributed to the difference in vocabulary gains favoring communication. It found that NoM occurred rarely: only 21 times over the approximate total time-on-task of six hours. Furthermore, seven of the 25 participants in the group that communicated did not engage in any NoM at all. Golonka et al. (2017) hypothesized that other interaction moves which occurred more frequently may have played a role in the gains. Specifically, they highlighted self- and partner-correction as well as the maintenance of positive affect as potential candidates. However, there were only ten instances of lexical correction and no explanation is given about how the maintenance of positive affect could have accounted for the vocabulary gains (although, we may speculate based on the ILH that this interaction move may have helped increase learners' perceived need to comprehend and use the target language). The researchers also mention using partners as a resource to understand vocabulary and usage as an interaction move that may have contributed to language development, but only 45 such instances occurred, of which the majority (21) concerned understanding task requirements. Note that 55 instances of correction were for spelling, but these were unlikely to have played a role in the gains because the vocabulary pre- and post-tests did not require learners to spell the words (Tare et al., 2014). In short, no specific interaction move seems to have accounted for the significant vocabulary gains in favor of the group that communicated. There was one clear difference, however, between the group that communicated and the group that did not, and that was the amount of output produced during the task, which was significantly greater for the group that communicated (Tare et al., 2014). This raises the question of whether it was simply the fact that the learners in the communicative group produced more output that led to their advantage in gains rather than some intrinsic property of communication itself.

A further issue that Tare et al. (2014) shared with prior studies comparing communicative to non-communicative tasks (e.g. Ellis et al., 1994) is that learners doing the communicative task would have been expected to feel a greater need to use the target language than learners who engaged in the non-communicative task. Based on the ILH, we may postulate that such a difference in perceived need alone, rather than a behavior that was intrinsic to the communicative process (e.g., NoM or NoF), may have led to the difference in gains in favor of the group that communicated. Hughes (2023) therefore conducted a similar study to Tare et al. (2014), but with a computer-mediated non-communicative task designed to raise learners' perceived need to use the target language to a level that was more comparable to that of the perceived need induced by the communicative task (see the Materials section for further details).

In the experiment by Hughes (2023), both the communicative and non-communicative groups did text-based information gap tasks requiring the paraphrasing of key vocabulary and their contexts. The information gap questions were as similar as possible given the fundamental differences between the tasks. Despite the relative brevity of the main tasks in the experiment of 12 minutes, pre- and delayed post-test results showed that both groups achieved significant gains in knowledge of form and meaning of target content. Moreover, there was a significant one-tailed rank correlation between the amount of output produced (as measured by words typed) during the task and delayed gains in productive knowledge of lexical meaning for both the communicative and non-communicative study groups, $r(16) = .46$, $p = .028$ and $r(14) = .53$, $p = .018$ respectively (p. 124). These significant correlations suggest that output played a central role in the gains. However, there was no significant difference in absolute gains between groups. This lack of a difference was considered supportive of the hypothesis that the apparent advantage of communicative tasks found in previous studies may have been at least partially due to a difference in the perceived need to use the target language rather than to the presence or absence of communicative interaction moves such as NoM. On the other hand, there was a statistically significant difference between groups in the amount of output produced during the task, $U = 61.00$, $p < .001$ (p. 124). Specifically, participants who did the communicative task typed significantly fewer words than participants who undertook the non-communicative task. Thus, while absolute vocabulary gains were statistically equivalent, learners in the group that communicated appear to have achieved those gains more efficiently. The present study seeks to investigate what interaction moves, if any, may have contributed to this apparent advantage in vocabulary learning

efficiency.

The present study

Similar to the follow-up study on Tare et al. (2014) by Golonka et al. (2017), the present study investigates the chat logs from Hughes (2023) in order to identify interaction moves that may have contributed to the advantage in vocabulary learning efficiency found in favor of communication. Specifically, it investigates the following research questions:

1. What interaction moves associated with language learning were engaged in by the group that completed the communicative task?
2. What differences in interaction moves are identifiable between more and less efficient language learners within the group that completed the communicative task?
3. Can any interaction moves identified through Questions 1 and 2 explain the difference in learning efficiency found in favor of the group that completed the communicative task (the chat group) versus the group that completed the non-communicative task (the note group)?

It is hoped that this investigation will reveal insights into the benefits of interaction that may have been less apparent in prior studies which were not able to attain as high a level of control as Hughes (2023) for differences in input, output amount, and need to engage the target language.

Method

Participants

Students enrolled in a first-year (n = 26) and second-year (n = 22) intermediate-level general English course at a public university in Japan participated in this study (mean age: 19.30). Participants were randomly divided within class into a chat group who undertook a communicative task and a note group who completed a similar but non-communicative task. The total combined number of participants in each group was n = 24 for the chat group and n = 23 for the note group. Informed consent was obtained from participants before conducting the experiment and task performance had no bearing on participants' grades in the courses.

Materials

Participants used an online application called Text Detective (TD) which automatically generates and administers both communicative and non-communicative information gap tasks from virtually any text (https://su-apps.org/td). TD was created by the researcher for the purpose of familiarizing learners with key vocabulary and their contexts within course texts. Learning of the vocabulary was promoted through tasks that required paraphrasing of those vocabulary and/or their contexts. In addition, TD featured an onclick dictionary function to assist learners in establishing initial form-meaning associations for unknown words. The communicative versions of the information gap task generated by TD were designed based on the task characteristics posited by Pica et al. (1993) to promote language acquisition, including the establishment of favorable conditions for NoM. As participants comprised two classes with differing curriculums, a different target text was employed for each. The tasks generated for this experiment involved four keywords and their contexts or key sections (1 to 3 sentences each) within each target text and focused on two of those keywords (the target words) and their respective key sections (target sections). See Table A1 in the appendices to view the target sections and associated target words which were the correct answers to the questions about those sections as well as the keyword distractors (the three keywords that were not the correct answer) used for each target text.

Two versions of TD were compared in the experiment from which this study's data was obtained: the communicative Chat version and the non-communicative Note version. Learners who used TD Chat worked in pairs to answer two questions generated by TD. For each question, one participant played the role of the detective, while the other played the witness, switching roles upon completion. Both partners were shown the target text during questions and could click on unfamiliar words to view their Japanese translations and associated images. However, the learners saw different parts of that

text highlighted: the detective saw a target word and three other keywords (the distractors) highlighted, whereas the witness saw one key section (one of the two target sections) highlighted (Figure 1).

The detective's job was to find out which of the four highlighted keywords was contained within the witness's target section through interacting with the witness via text chat. To ensure engagement with meaning, TD redacted certain words from learners' messages including keywords and other content words from key sections as well as overly similar word combinations, non-English words, and words commonly used to indicate text sections by location. These were replaced with blanks ("___") when the message was displayed in the chat window (Note: In this paper, redacted words in chat excerpts are indicated with an underline). The detective had two chances to choose the correct keyword and both players were awarded points when the detective answered correctly as well as when either player used a synonym for a keyword or content word from a key section. There was a 6-minute time limit per question and a total maximum time allowed of 12 minutes.

TD Note involved the same target content, dictionary function, and redaction and reward system as TD Chat, but was played individually rather than in pairs. TD Note first showed players three key sections—two of which were target sections—highlighted in the text, one by one. For each section, the player was asked to take a note which was recorded in the chat window. Afterward, TD showed the player a question below the target text window which featured a blurred-out key section and the note the player had taken for that section as well as the four keywords. The question asked which of the four keywords was contained in the blurred-out section for which they had taken the note (Figure 2).

As with TD Chat, the target text was displayed throughout the task. After a maximum of two answer attempts, another question appeared for which the second target word was the correct response. The maximum time allowed was 12 minutes, equal to the total maximum time allowed for the TD Chat task.

Before the experiment, both groups completed an input-only warm-up task through TD, which featured the same target content and dictionary function but did not involve output production. The pur-

Figure 1 *Left: The detective's view with one of the four highlighted keywords (the target word "Engineering") scrolled into view; Right: The witness's view with one of the key sections highlighted and scrolled into view*

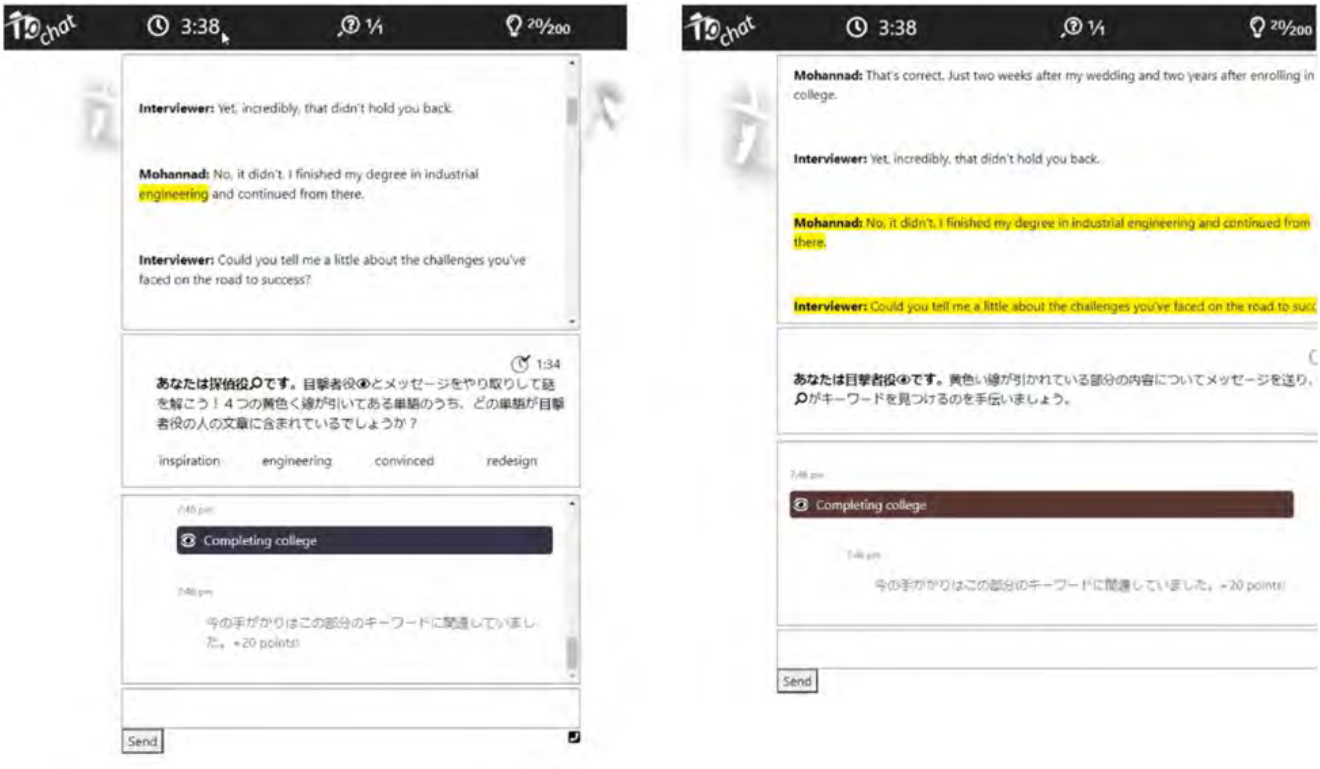

Figure 2 *TD Note question featuring the learner's note ("completed college") for the blurred-out key section and four keyword choices below it*

pose of the warm-up task was to provide the learners with some initial exposure to the target text, key sections, and keywords as viewed through TD as well as to refamiliarize learners with the program (which they had used before, but not recently). During the warm-up task, learners were shown three of the four key sections (including the two target sections) highlighted in the text one at a time and asked to study each. Afterward, TD scrolled to a target section with its target word and six surrounding words blurred out and asked players which of the four keywords was within the blurred-out section. After the player answered, one final question was displayed in the same format but on the second target section. The maximum time allowed was 4 minutes.

Measures

Interaction moves. The coding of interaction moves was carried out by the author and another researcher who first coded all of the chat logs individually and then resolved the differences through discussion until 100% agreement was reached. The following describes the types of interaction moves for which coders initially investigated and Table 1 below provides examples of each from the actual chat data (except in the case of NoF and partner-correction which did not occur at all during learner interaction).

Paraphrasing of target content was defined as any attempt to express the meaning of a given section of the target text that was not a verbatim copy or obvious attempt at a verbatim copy of that section. Therefore, "this is the memo to inform about the party" and "Let everyone details" were both considered attempts to paraphrase Target Section 1.1 (See Table A1), despite their reusing some words from the section. However, "it is memo, not e-mail" and "first sentence" were not considered paraphrasing as they did not express the meaning of the section referred to but rather indicated the loca-

tion of that section. Regarding the accuracy of paraphrasing, if the message was accurate enough for the coders to identify the section to which it likely referred, then it was considered paraphrasing. Otherwise, it was not.

Paraphrasing a target word was operationalized as an attempt to express the meaning of a target word through the use of another word or words. It was decided to focus on the paraphrasing of only the two target words as opposed to all four keywords, because preliminary coding of the chat logs revealed only two instances in which a non-target keyword was paraphrased, implying the unlikelihood of chat participants' output having led to gains in knowledge of these words.

Negotiation was coded based on the guidelines of Silver and Huynh (2010) which, in accordance with Pica et al. (1996), Van den Branden (1997) and Lyster and Ranta (1997), distinguishes between *NoM, NoF*, and *negotiation of content (NoC).* When coding for NoM, logs were first checked for indications that a prior message was not understood through identifying signals of possible attempts to initiate negotiation, including comprehension checks and requests for clarification and confirmation (Gass & Mackey, 2020). Next, coders attempted to locate the prior message that may have been the "trigger" for negotiation. If the trigger was judged as a potential source of a lack of understanding, the exchange was coded as NoM. Otherwise, if the trigger involved a grammatical, lexical, or spelling error, which was corrected or otherwise commented on by the initiator of negotiation, it was considered NoF, and if not, it was coded as NoC. Note that it was sometimes difficult to distinguish these types of negotiation from each other. For instance, the NoM example in Table 1 may upon first glance appear to be NoF. However, since the correct form "liked" was not mentioned and the initiator of negotiation (I) chose the correct answer immediately upon receiving confirmation that the responder (R) was talking about the past, the aim of negotiation appears to have been to figure out which part of the target text R was referring to rather than to discuss the linguistic issue of how to form the past tense correctly (see the example of NoF for comparison). Likewise, the example of NoC might first appear to be NoM. However, R's addition of "since he was little" substantially clarifies the fact that R is referring to the key section about the inventor's childhood. I's message "about dolls or robot and so on?" was therefore interpreted as an attempt to elaborate further on the content of the discussion rather than to address a breakdown in communication. The exchange was thus coded as NoC rather than NoM.

Correction was identified as any instance where a learner supplied a correction to a linguistic error that they had made (*self-correction*) or their partner had made (*partner-correction*).

Maintenance of positive affect was defined as any instance in which a learner expressed praise or encouragement with regard to their partner, themselves, the task, or their cooperative effort.

Base target word proficiency, target word gains, and L2 learning efficiency. Hughes (2023) measured L2 development with a modified C-test administered through TD before and immediately after the experimental task and then once again unannounced one week later. The test was created from the target texts used during the task and consisted of completion items which required learners to complete the last half of five words from each of four key sections (including each keyword). The test also included a translation item for each of the four keywords requiring learners to fill in its meaning in Japanese. This latter item type measured gains in lexical knowledge of meaning and delayed gains measured by these items were found to be significantly correlated with the amount of output learners produced during the task, suggesting that interaction moves may have contributed to the gains. Therefore, the current study focuses only on these translation items and delayed gains made on them. Furthermore, upon coding the chat logs for interaction moves, it became apparent that all cases in which students paraphrased a keyword except two were for target words. Therefore, only the delayed gains for the two target word translation items are considered in this study and are henceforth referred to simply as *target word gains*. Likewise, *base target word proficiency* is operationalized as the pre-test scores for these two target word items. Each of these items was rated by a native Japanese speaker who was highly proficient at English. The items were rated on a scale of 0 (incorrect) to 2 points (correct) with partially correct answers receiving 1 point. The total maximum score for target word translation items was thus 4 points. L2 vocabulary learning efficiency (henceforth *L2 learning efficiency*) was operationalized as target word gains divided by learners'

Table 1 *Types and examples of interaction moves (in dialogic exchanges, the initiator of an interaction move is indicated as "I" and the responder as "R")*

Type	Example	Note
Paraphrasing of target content	he completed his study and didn't stop	referring to "I finished my degree in industrial engineering and continued from there." (transcribed from Scanlan & Vargo, 2015, p. 183)
Paraphrasing a target word	science type of major	referring to "engineering"
Negotiation of (or for) meaning (NoM)	R: He like inventing I: Past? R: Past *I chooses the correct answer*	referring to "I started inventing in my childhood. I liked to redesign my toys, and to try and repair household items. (transcribed from Scanlan & Vargo, 2015, p. 183)
Negotiation of content (NoC)	R: he likes to make up something from nothing R: since he was little I: about doll or robot and so on? R: yeah like that.	referring to "I started inventing in my childhood. I liked to redesign my toys, and to try and repair household items. (transcribed from Scanlan & Vargo, 2015, p. 183)
Negotiation of form (NoF)*	R: He like remaking toys.** I: Likes or liked? R: liked	referring to "I started inventing in my childhood. I liked to redesign my toys, and to try and repair household items. (transcribed from Scanlan & Vargo, 2015, p. 183)
Self-correction	there is few boxes witch you can put on the bags so you have to pack your baggage which	correcting the spelling of "which"
Partner-correction*	R: He like remaking toys. I: You mean he "liked" remaking toys. R: Yes, he liked that.	referring to "I started inventing in my childhood. I liked to redesign my toys, and to try and repair household items. (transcribed from Scanlan & Vargo, 2015, p. 183)
Maintenance of positive affect	nice, ok	sent just after their partner provided information that was necessary to correctly answer the question

* The examples for these interaction moves were created by the researcher as they did not occur during learner interaction (all other examples are from the actual chat data collected during the experiment).
** Underlined words were redacted by TD and replaced with blanks ("___") in the TD chat window.

output wordcounts during the TD task.

For the final stage of the analysis, L2 learning efficiency was compared across communicative (chat) and non-communicative (note) groups and thus had to be calculated for the latter. This was done in the same way as for the chat group, except that the output wordcount for note participants was adjusted to omit lines they had typed about the third key section, which only note participants had been prompted to paraphrase.

Procedure

In the two months prior to the study, learners were given practice tasks in each version of TD as well

as tests in the same format as the ones used in the study in order to familiarize them with the system. The target texts used for this study had been assigned as homework the previous week, but the assignments had focused on general comprehension rather than on keywords and their contexts. During the experiment from which the data was obtained, learners first took the pre-test (4 minutes), then did the warm-up task in TD, after which they did the main task using either TD Chat or TD Note for a maximum of 12 minutes. The post-tests, which were identical to the pre-test were administered once immediately after the main task and then again unannounced one week later.

Results

1. What interaction moves associated with language learning were engaged in by the group that completed the communicative task?

Table 2 displays the interaction moves that were identified by the coders in the TD chat data. Note that NoF and partner-correction did not occur at all, while self-correction occurred only twice (once for spelling and once for punctuation) and maintenance of positive affect just once ("nice, ok"). The infrequency of these moves implies that they were not relevant to the gains and learning efficiency experienced by the chat group. The table displays the overall counts for the other interaction moves identified and how many students enacted each move as well as the breakdown of counts for target-word proficient learners defined as those who scored in the 3rd percentile or higher on the pre-test target word translation items (3 or more points out of 4) versus non-target-word-proficient learners who scored below the 3rd percentile (less than 3 out of 4 points).

This breakdown allows us to observe that paraphrasing target content and target words occurred for the same or similar percentages of proficient versus non-proficient learners, whereas the percentages were weighted in favor of proficient compared to non-proficient learners for NoC (40% versus 21%) and especially NoM (80% versus 5%). Notably, only one of the 19 non-proficient learners initiated NoM. It is important to consider whether or not learners who engaged in a given interaction move were already target-word proficient before the task began because only those who were non-target-word proficient had the potential to gain further knowledge of those words as measured by the test (while target-word-proficient learners had already achieved maximum or near-maximum scores on the pre-test and therefore had no further "room to grow"). For this reason, only interaction moves enacted by learners who were non-target-word proficient at the outset of the task could be expected to have contributed to the gains.

2. What differences in interaction moves are identifiable between more and less efficient language learners within the group that completed the communicative task?

Recall that L2 learning efficiency was measured by dividing delayed gains in the knowledge of target word meanings (target word gains) by the number of words typed during the task (learners' output). More efficient learners achieved greater gains in fewer words than less efficient learners. To identify potential differences in interaction between more and less efficient learners, a qualitative analysis was conducted involving four non-target-word proficient chat participants whose output wordcounts happened to be equal within two words but whose gains in target word knowledge differed (Table 3) and their communicative behaviors compared (Table 4). Note that all four participants received a pre-test score of 0 for both target words, meaning that they all had the potential to achieve a 4-point gain.

As shown, all four learners gained full points on the delayed post-test item for the target word they encountered as the witness, but only two students (S1 and S2) received points for the target word they encountered as the detective. In other words, if an interactive behavior was the cause of the difference in learning efficiency, that behavior should have emerged while the learners were in the role of the detective. Therefore, the chat logs for the four learners while they were playing the detective were analyzed and compared (see Table A2 for the complete logs). Table 4 shows interaction moves observed for some learners but not others.

Note that as the detectives saw only keyword choices highlighted in the text, they did not have any direct knowledge of where the target section shown (through highlighting) to the witness began and

Table 2 *Interaction moves observed in the chat data*

	Paraphrasing target content	Paraphrasing a target word	NoM	NoC
Total Instances	57	23	6	6
Pairs that engaged in the interaction move	12 (100%)	10 (83%)	4 (33%)	5 (42%)
Learners who initiated the interaction move	24 (100%)	14 (58%)	5 (21%)	6 (25%)
Target-word-proficient learners who initiated the interaction move	5 (100%)	3 (60%)	4 (80%)	2 (40%)
Non-target-word-proficient learners who initiated the interaction move	19 (100%)	11 (58%)	1 (5%)	4 (21%)

Table 3 *Output, target word gains, and L2 learning efficiency per student*

Student	Output word-count	Target word gain while in the role of:		Target word gain (max = 4)	L2 learning efficiency
		Witness (max = 2)	Detective (max = 2)		
S1	26	2	2	4	0.15
S2	26	2	2	4	0.15
S3	25	2	0	2	0.08
S4	27	2	0	2	0.07

Table 4 *Interaction moves and target word gains while playing the detective*

Student	Move enacted (1 = yes, 0 = no)					Target word gain (max = 2)
	Paraphrased a target word	Paraphrased target word context	Initiated NoC	Responded to NoC	Asked a question	
S1	0	1	0	0	1	2
S2	0	1	1	1	1	2
S3	0	0	0	0	0	0
S4	1	0	0	0	1	0

ended. Therefore, the detectives' paraphrasing of content near but not necessarily including the target word is referred to as paraphrasing the "target word context" as opposed to the "target section."

The only interaction move observed for the efficient learners that did not occur among other learners was paraphrasing the target word context. However, every instance of target context paraphrasing in

Table 5 *EQF with other interaction moves for comparison*

	EQF	Paraphrasing a target word	NoM	NoC
Total Instances	11	23	6	6
Pairs that engaged in interaction move	8 (67%)	10 (83%)	4 (33%)	5 (42%)
Learners who initiated interaction move	11 (46%)	14 (58%)	5 (21%)	6 (25%)
Target-word-proficient learners who initiated interaction move	5 (100%)	3 (60%)	4 (80%)	2 (40%)
Non-target-word-proficient learners who initiated interaction move	5 (28%)	11 (58%)	1 (5%)	4 (21%)

part or in whole initiated by a detective was done so through the formulation of a question (as a confirmation check).

Interestingly, neither efficient learner paraphrased the target word, whereas one of the other learners (S4) did, paraphrasing "lockers" through the confirmation check, "circle box?" This attempt at paraphrasing was puzzling at first until the coders recalled that clubs at university in Japan are often referred to as "circles." Notice that this confirmation check by S4 does not include any reference to the context in which the target word appeared (see Target Section 1.2 in Table A1). On the other hand, the efficient learners formulated questions which were context-elaborative in that they paraphrased the content surrounding or nearby the target word (See Table A2, Lines 3 and 25). Thus, an additional interaction move that may have contributed to learning efficiency was obtained: (context-) elaborative question formulation (EQF). Next, the chat logs were re-examined to identify all instances of EQF. Table 5 displays these along with previously identified interaction moves for comparison. Note that every occurrence of EQF was produced by learners in the role of the detective.

As shown, the total instances of EQF (11) were nearly equal to the instances of NoM and NoC combined (12). Similar to NoM and NoC, however, EQF occurred disproportionately among target-word-proficient learners compared to non-proficient learners (100% versus 28%). Also, there was an overlap between EQF and NoC, with all initiations of NoC except one ("Is it in the beginning?") also coded as EQF. Some EQF instances would have been considered NoC initiations except that they were not produced in response to a prior message by a partner (such as "Is it about amount of lockers?" on Line 3 in Table A2). Three of the five NoM initiations were also marked as EQF.

3. Can any interaction moves identified through Questions 1 and 2 explain the difference in learning efficiency found in favor of the group that completed the communicative task (the chat group) versus the group that completed the non-communicative task (the note group)?

One-tailed rank correlation analyses ($\alpha < .05$) were conducted between whether or not learners enacted a given interaction move (1 = enacted, 0 = did not enact) and their target word gains and L2 learning efficiency. Rank—as opposed to Pearson—correlations were employed due to the non-normality of the data distributions as indicated by Shapiro-Wilk tests ($p < .05$). Correlations with base target word proficiency and output wordcount were also examined. Base target word proficiency was examined to investigate whether it was predictive of certain interaction moves as was suggested by the results pertaining to RQ1 above, which showed that a larger percentage of target word proficient learners initiated NoM and NoC compared to non-target word proficient learners. Table 6 displays the results. Note that, because one of the efficient learners (S2) had responded to NoC, implying the

Table 6 *One-tailed rank correlations between interaction moves and base proficiency, output, target word gains, and L2 learning efficiency*

	Base target word proficiency (*n* = 24)		**Output word count (*n* = 24)**		**Target word gain (*n* = 18)**		**L2 learning efficiency (*n* = 18)**	
Interaction move	*r*	*p*	*r*	p	*r*	*p*	*r*	*p*
Paraphrasing a target word	.16	.223	.14	.265	.30	.110	.21	.205
Initiating NoM	.48	.009	.12	.282	.12	.315	.26	.149
Initiating NoC	.05	.405	.40	.027	.54	.011	.51	.016
Responding to NoC	-.36	.044	.07	.378	.21	.199	.15	.283
EQF	.22	.152	.60	.001	.69	.001	.54	.010

move's potentially positive effect on learning efficiency, it is also included in the table. Also note that whereas, the correlations with base proficiency and output included all chat participants (n = 24), correlations for target word gains and learning efficiency included only the non-target-word proficient learners (n = 18) who had "room to grow" with regard to their target word knowledge.

As shown, EQF and the initiation of NoC correlated significantly with L2 learning efficiency, while neither was significantly correlated with base proficiency. Both moves were also significantly correlated with output wordcount. However, even after gains were divided by output wordcounts (to obtain L2 learning efficiency), the correlations with EQF and initiating NoC remain significant. On the other hand, initiating NoM was significantly correlated with base proficiency but not with any other factor. Meanwhile responding to NoC did not correlate significantly with any factor except base proficiency (but negatively). Finally, paraphrasing the target word was not significantly correlated with gains or efficiency. Other interaction moves in addition to those shown in the table, including responding to NoM and responding to EQF, were examined for relationships with gains and efficiency, but no significant correlations emerged.

EQF correlated most strongly with L2 learning efficiency and was engaged in by more non-proficient participants than NoC (five versus four). Therefore, the learning efficiency of non-proficient participants in the chat condition who engaged in EQF was compared to the efficiency of those in the same condition who did not engage in EQF as well as to the efficiency of non-proficient participants who worked individually in the note condition. Table 7 displays the results.

As shown, a significant difference in L2 learning efficiency emerged in favor of chat participants who engaged in EQF compared to those in the same condition who did not, while a marginally significant difference emerged in favor of chat participants who engaged in EQF compared to note participants. Meanwhile, there was no significant difference in efficiency between note participants and chat participants who did not engage in EQF.

Table 7 *Comparison of L2 learning efficiency*

Groups compared	Median (IQR) of former group	Median (IQR) of latter group	*U**	*p**	*r**
Chat with EQF (*n* = 5) vs. Chat WITHOUT EQF (*n* = 13)	0.12 (0.03)	0.00 (0.08)	10.00	.025	.53
Chat with EQF (*n* = 5) vs. Note (*n* = 17)	0.12 (0.03)	0.05 (0.07)	19.00	.064	.40
Chat WITHOUT EQF (n = 13) vs. Note (*n* = 17)	0.00 (0.08)	0.05 (0.07)	96.50	.552	.11

Discussion

The results provide insights into what interaction moves were and were not likely responsible for the advantage in efficiency demonstrated by learners who communicated compared to learners who did the non-communicative task. Foremost among the interaction moves that were unlikely to have led to the advantage was NoM: Whereas one-fourth of the chat participants (6) engaged in NoM, only one of those participants had been non-target-word proficient at the outset of the task, meaning that only one of the learners who enacted the move had the potential to gain further knowledge of the target words through that move. It was thus unsurprising that no significant correlations were found between initiating or responding to NoM and target word gains or learning efficiency. Likewise, other interaction moves posited as supporting language acquisition, including self- and partner-correction as well as the maintenance of positive affect occurred rarely or not at all, yet significant L2 gains were made nonetheless.

The most likely explanation for the chat group's advantage in learning efficiency was identified through the comparison of efficient and non-efficient learners who produced similar amounts of output. The comparison revealed that the more efficient learners engaged in EQF while in the role of the detective, whereas less efficient learners did not. To investigate whether this finding might apply to the entire group, chat logs were recoded for instances of EQF and statistical analyses conducted. Results showed that EQF not only significantly correlated with L2 learning efficiency, but chat participants who engaged in EQF demonstrated significantly higher efficiency compared to those who did not. Furthermore, there was a marginally significant difference in learning efficiency between chat participants who engaged in EQF and note participants in favor of the former group. These results suggest that engaging in EQF may have been the main factor contributing to the chat group's advantage in L2 learning efficiency.

The above findings raise the question of why paraphrasing target words failed to have a significant positive relationship with target word gains and learning efficiency. Paraphrasing target words in output, though explicit, is not the only indication a learner can provide in output of generative or evaluative processing of a target word. Rather paraphrasing nearby context may also indicate processing of that word as understanding the word may have been necessary for that context to be paraphrased.

In addition, as shown in the case of participant S4, simply paraphrasing target words without reference to the context in which they appeared did not seem to result in gains in knowledge of those words. This may have been due to the measures requiring completion and translation of the target words within context. To do this successfully, developing a cognitive association between the mean-

ing of that context and the target word was likely important. Of course, learners in the role of detective did not technically have to produce any output to make this connection, and many simply waited for their partner to paraphrase the target section well enough for them to locate it and the target word within it. However, it seems that producing output in which the connection was made, strengthened it. This conjecture aligns both with the ILH and the Output Hypothesis (Swain, 2005) and helps to explain why EQF appears to have led to an advantage in learning efficiency.

Regarding whether or not it was necessary to formulate the paraphrasing of target context as a question, it seems likely that it was not. This is because, whereas many learners gained knowledge of the target word they encountered while playing the role of the witness, none had engaged in EQF while in that role. However, this raises the question of what made communication during the task different from the non-communicative study that the note group engaged in. We may speculate that the detective's access to the keyword choices during output production allowed the edge in learning efficiency. Recall that learners in the note group saw only key sections during output production, meaning that they could not be sure which word in those sections was a keyword. On the other hand, detectives in the chat group could see all the keyword choices while producing output, presumably encouraging stronger cognitive associations between the context and the target word during paraphrasing.

We might further speculate, then, that had TD Note identified which word in each key section was the keyword during output production, then the note group would have matched the chat group in learning efficiency. However, the case of S4 suggests that simply paraphrasing a target word without reference to its context was not effective for gaining and retaining knowledge of its meaning. If TD Note were to have identified the keywords during output production, learners may have simply resorted to this kind of decontextualized keyword paraphrasing. Moreover, that task would still have differed from the one achieved through true communication in that an unknown element still existed for the detective in the chat task—namely, the boundaries of the target section as seen by the witness. In the chat task, the detective, seeking to efficiently confirm whether or not the section contained a target word, was likely motivated to express the most comprehensible part of that section to ensure quick and accurate confirmation. However, since the boundaries of the target section were unknown, the detective could not be sure how far away from the keyword the section extended and therefore was also under pressure to paraphrase parts that were closer to the keyword more than parts that were farther away. This tension between the desire to paraphrase the easiest parts, but also the parts closest to the keyword, arguably promoted more in-depth evaluation of the target word and its context than would have been achievable if the detective had known the boundaries of the target section. Thus, the nature of communication itself, in which something unknown is made known through an exchange of meaning, arguably contributed to the efficiency of the communicative task.

Conclusion

Similar to Golonka et al. (2017), this study found that NoM occurred only rarely during communication. Also, the majority of participants who engaged in NoM were already target-word proficient at the outset of the task, whereas only one non-target word proficient learner engaged in NoM. Furthermore, responding to NoM did not correlate significantly with L2 gains or learning efficiency. These findings suggest that NoM may be less important for language learning than previous studies have implied.

Meanwhile, paraphrasing target vocabulary was also found not to significantly correlate with target word gains and learning efficiency. Rather, paraphrasing the context around or nearby target words via EQF emerged as having the strongest relationship with these variables. It is hypothesized that this was because the meanings of the target vocabulary were likely to have been processed in connection with those contexts during the paraphrasing even if those meanings were not explicitly given in the output. This paraphrasing approach may have been more effective for participants in the communicative group because, in the role of the detective, they could see the keyword choices which included the target vocabulary during output production while at the same time being unsure of the boundaries of the target section as seen by the witness. The tension between the known keyword choices and unknown contextual boundaries may have created an optimal level of involvement during the paraphrasing process that could not have been achieved in a non-commutative task.

This study had several limitations including the brevity of the experimental tasks and small sample size, as well as the arguably unnatural style of interaction due to TD's content mediation, and further research is necessary before any firm conclusions can be reached. Still, it holds some important implications with regard to the role of interaction in promoting language development. First, it suggests that patterns of interaction between learners, such as NoM may not be as important as processing input through the production of output in order to resolve gaps in knowledge. Second, it provides evidence that a vocabulary item need not be directly used or paraphrased in output for the production of that output to contribute to the development of knowledge of that item. Rather output paraphrasing the immediate context around or nearby an item seems effective, perhaps because learners process the item while producing that output even if they do not mention or paraphrase it directly. Overall, this study suggests that to better understand how communication leads to language development, it is important to investigate the relationship between the content of output and the input about which the communication is taking place.

Acknowledgements

This study was made possible in part by the Ministry of Education, Culture, Sports, Science, and Technology of Japan under Grant 21K00673. Also, the author would like to thank Corey Fegan for his invaluable assistance.

References

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition, 24*(1), 81–112. https://doi.org/10.1017/ S0272263102001043

Ellis, N. C. (2019). Essentials of a theory of language cognition. *Modern Language Journal, 103*(s1), 39–60. https://doi.org/10.1111/modl.12532

Ellis, N. C., & Wulff, S. (2019). Cognitive approaches to L2 acquisition. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 41–61). Cambridge University Press.

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications* (pp. 3–18). Routledge.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*(2), 285–301. https://doi.org/10.1017/ S0272263199002077

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*(3), 449–491. https://doi.org/10.1111/j.1467-1770.1994.tb01114.x

Gass, S. M., & Mackey, A. (2020). Input, interaction, and output in L2 acquisition. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (3rd ed., pp. 192–222). Routledge.

Golonka, E. M., Tare, M., & Bonilla, C. (2017). Peer interaction in text chat: Qualitative analysis of chat transcripts. *Language Learning & Technology, 21*(2), 157–178. https://dx.doi.org/10125/44616

Hughes, L. S. (2022). Text-based SCMC for SLA: A narrative review. *Technology in Language Teaching & Learning, 4*(1), 1–17. https://doi.org/10.29140/tltl.v4n1.712

Hughes, L. S. (2023). Effects of dynamically computer-mediated communication and "pseudo" communication on L2 learning. *CALL-EJ, 24*(1), 105–136.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, *24*(4), 541–577. https://doi.org/10.1017/S0272263102004023

James, W. (1950). *The principles of psychology, Volume 1*. Dover.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19*(3), 357–377. https://doi.org/10.1093/applin/19.3.357

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365–386. https://doi.org/10.1177/1362168808089922

Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Benjamins.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language learning, 58*(2), 285–325. https://doi.org/10.1111/j.1467-9922.2008.00442.x

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1–26. https://doi.org/10.1093/applin/22.1.1

Leeser, M. J. (2008). Pushed Output, Noticing, and Development of Past Tense Morphology in Content-Based Instruction. *Canadian Modern Language Review, 65*(2), 195–220. doi:10.3138/cmlr.65.2.195

Liu, S., & Reynolds, B. L. (2022). Empirical support for the involvement load hypothesis (ILH): A systematic review. *Behavioral Sciences, 12*, 354. https://doi.org/10.3390/bs12100354

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27*(3), 361–386. https://doi.org/10.1017/S0272263105050163

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. https://doi.org/10.1017/S0261444818000125

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bathia (Eds.), *Handbook of research on second language acquisition* (pp. 413–468). Academic Press.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66. https://doi.org/10.1017/S0272263197001034

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*(4), 557–587. https://doi.org/10.1017/S0272263199004027

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2019). The different aspects of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 15–29). Routledge.

Nguyen, C. D., & Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quarterly, 53*(1), 5–29. https://doi.org/10.1002/tesq.441

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice. Vol 1.* (pp. 9–34). Multilingual Matters.

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly*, *30*(1), 59–84. https://doi.org/10.2307/3587607

Rabie-Ahmed, A., & Mohamed, A. (2022). Collaborative and individual vocabulary learning in the Arabic classroom: The role of engagement and task demands. *Foreign Language Annals*, 1–19. https://doi.org/10.1111/flan.12636

Rott, S. (2004). A comparison of output interventions and un-enhanced reading conditions on vocabulary acquisition and text comprehension. *Canadian Modern Language Review, 61*(2), 169–202. https://doi.org/10.3138/cmlr.61.2.169

Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research, 6*(3), 183–222. https://doi.org/10.1191/1362168802lr108o

Scanlon, J., & Vargo, M. (2015). *Q: Skills for success: Listening and speaking.* Oxford University Press.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11* (2), 129–158. https://doi.org/10.1093/applin/11.2.129

Silver, R. E. & Huynh, T. C. D. (2010). *Manual for coding negotiation for meaning in classroom interaction*. Centre for Research in Pedagogy and Practice, National Institute of Education (Singapore). https://hdl.handle.net/10497/17728

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320–337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Tare, M., Golonka, E. M., Vatz, K., Bonilla, C. L., Crooks, C. & Strong, R. (2014). Effects of interactive chat versus independent writing on L2 learning. *Language Learning & Technology, 18*(3), 208–227. https://dx.doi.org/10125/44391

Trew, G. (2009). *Tactics for TOEIC Listening and Reading Test.* Oxford University Press.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, *47*(4), 589–636. https://doi.org/10.1111/0023-8333.00023

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, *11*(2), 121–142. https://doi.org/10.1177/1362168806070745

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, *29*(3), 325–340. https://doi.org/10.1016/S0346-251X(01)00022-7

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, *11*(2), 87–95. https://doi.org/10.1080/00461527409529129

Yanagisawa, A., & Webb, S. (2022). Involvement load hypothesis plus: Creating an improved predictive model of incidental vocabulary learning. *Studies in Second Language Acquisition, 44*(5), 1279-1308. https://doi.org/10.1017/S027226312200033X

Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis. *Studies in Second Language Acquisition, 42*(2), 411-438. https://doi.org/10.1017/S0272263119000688

Yilmaz, Y. (2016). The role of exposure condition in the effectiveness of explicit correction. *Studies in Second Language Acquisition, 38*(1), 65–96. https://doi.org/10.1017/S0272263115000212

Zaki, H., & Ellis, R. (1999). Learning vocabulary through interacting with a written text. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 153–169). John Benjamins.

Appendices

Table A1 *Texts, target sections and words, and distractors*

Text	Target Section ID	Target Section	Target Word	Distractors
A (Trew, 2009, p. 159)	1.1	MEMORANDUM To: All members of the sales department Re: Year-end party From: Annette Derringer Date: November 26 This is just a quick note to let you all know the arrangements for next week's year-end party.	arrange-ments	lockers, attend, request
	1.2	Also, the number of lockers available is limited, so guests should try to keep belongings to a minimum. Thanks in advance, Annette	lockers	arrangements, attend, request
B (transcribed from the listening activity in Scanlon & Vargo, 2015, on p. 183)	2.1	Mohannad: No, it didn't. I finished my degree in industrial engineering and continued from there. Interviewer: Could you tell me a little about the challenges you've faced on the road to success?	engineering	redesign, inspiration, convinced
	2.2	Interviewer: So inventing is something that you were always interested in? Mohannad: Yes. I started inventing in my childhood. I liked to redesign my toys, and to try and repair household items.	redesign	engineering, inspiration, convinced

Table A2 *Dialogues and player actions of students (S1-4) with similar output wordcounts but different learning outcomes and their partners (P1-4) while the students played the role of detective*

Line ID	Dialogue and Player Actions	Time
	S1	
1	*P1 is shown Target Section 1.2 and S1 the corresponding keyword choices.*	(2:23 pm)
2	*S1 scrolls to section containing the new target word ("lockers").*	
3	**S1:** is it about amount of lockers*	
4	**S1:** ?	(2:24 pm)
5	**P1:** so storage is few	
6	**S1:** ok	
7	*S1 chooses the target word, correctly answering the question.*	
	S2	
8	*P2 is shown the target section 2.1 and S2 the corresponding keyword choices.*	(11:24 am)
9	*S2 scrolls twice, allowing the viewing of sections containing two different keywords.*	
10	*P2 looks up the word "continued."*	(11:25 am)
11	*S2 scrolls to a section containing another keyword.*	
12	**P2:** he still continue after end of degree	
13	*S2 scrolls to a section of text containing the remaining of the four keyword choices ("inspiration"), and then to the section containing the target word ("engineering").*	(11:26 am)
14	*S2 scrolls back to the section of text containing "inspiration" and then back to the section containing the target word.*	
15	**P2:** interviewer ask the difficult to best carrer	(11:27 am)
16	*S2 scrolls back to the section of text containing "inspiration."*	

17	*P2 scrolls to a section of text that does not contain any keywords.*	
18	*S2 scrolls to the section of text containing the distractor keyword "convinced."*	
19	*P2 scrolls back to the section of text that contains the target word.*	
20	**P2:** at <u>first</u> he <u>denys</u>	(11:28 am)
21	*P2 looks up the word "degree"*	
22	*S2 scrolls to the section containing the target word.*	
23	*P2 looks up "industrial" and "engineering."*	
24	**P2:** he graduate the science type of major	(11:29 am)
25	**S2:** like he continued after the end of his study in school / *["/" is assumed to be a mistyped "?"]*	
26	*Players are awarded bonus points for S2's use of synonyms for content words.*	
27	**P2:** <u>yes</u>	
28	*S2 chooses the target word, correctly answering the question.*	(11:30 am)
	S3	
29	*P3 is shown Target Section 1.1 and S3 the corresponding keyword choices.*	(2:21 pm)
30	*S3 scrolls to a section of text containing two keyword distractors.*	
31	**P3:** Let everyone details	(2:22 pm)
32	*S3 scrolls to the section of text containing the target word ("arrangements").*	
33	*S3 looks up the word "arrangements."*	
34	**P3:** schedule	
35	**P3:** know preparation	(2:23 pm)
36	**S3:** ok	
37	*S3 chooses the target word, answering the question correctly.*	

	S4	
38	*P4 is shown Target Section 1.2 and S4 the corresponding keyword choices.*	(2:27 pm)
39	*S4 currently has the section of text containing keyword distractors "attend" and "request" in view.*	
40	**S4:** participate?	
41	**S4:** need?	
42	**S4:** circle box?	(2:28 pm)
43	*S4 scrolls to section of text containing the target word "lockers" and keyword distractor "arrangements."*	
44	*S4 looks up "arrangements" for the third time (looked it up twice while playing the witness).*	
45	**S4:** decided	
46	**P4:** sircle box	
47	*S4 chooses the target word, answering the question correctly.*	

*Underlined words were redacted by TD