



Effects of Video-Based Interaction on the Development of Second Language Listening Comprehension Ability: A Longitudinal Study

Kazuya Saito & Yuka Akiyama

Abstract

In our precursor research (Saito & Akiyama, 2017 in *Language Learning*), we reported that one academic semester of video-based L2 interaction activity was facilitative of various dimensions of the Japanese learners' spontaneous *production* ability development (e.g., comprehensibility, fluency and vocabulary). In this paper, we aimed to revisit the dataset to examine the effects of long-term interaction on the development of L2 *comprehension* ability. Consistent with the interactionist account of L2 comprehension ability development, the results showed that longitudinal interaction enhanced Japanese learners' comprehension (measured via a general listening proficiency test), as it provided opportunities for comprehensible input and output (measured via video-coding analyses) during meaning-oriented discourse.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Effects of Video-Based Interaction on the Development of Second Language Listening Comprehension Ability: A Longitudinal Study

One of the most extensively researched topics in the field of second language acquisition (SLA) concerns the relationship between interaction, negotiation for meaning and second language (L2) development. According to the interaction hypothesis (e.g., Long, 1996), much learning happens precisely when non-native speakers (NNSs) engage in meaningful conversations with native speakers (NSs), and encounter communication breakdowns for language-related reasons. To retrieve impaired meaning, NSs must make an intuitive and/or a conscious effort to facilitate NNSs' comprehension by way of negotiation strategies—clarification requests, confirmation checks and repetition—in addition to recasting NNSs' erroneous production (i.e., comprehensible input). As a result of negotiation for meaning, NNSs are induced to notice and understand the gap between their own interlanguage systems and the incoming input, and then produce more targetlike forms (i.e., comprehensible output).

To date, many researchers have claimed that opportunities to negotiate meaning through interaction facilitate comprehension (e.g., Ellis, Tanaka, & Yamazaki, 1994). In light of an L2 listening research perspective, negotiation for meaning can provide acquisitionally-rich contexts for the development of bottom-up processing (drawing on phonological, temporal, lexical and grammatical information) as well as top-down processing (connecting the linguistic knowledge with world knowledge) (Rost, 2011). In the face of communication breakdowns, NNSs can receive a great deal of comprehensible input thanks to NSs' use of negotiation strategies. According to observational studies (e.g., Long, 1983), such interactionally modified speech typically contains linguistic characteristics beneficial for L2 comprehension, including the repetition, paraphrasing and simplification of original utterances (Jensen & Vinther, 2003) with

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

slower speech rate (Zhao 1997). To avoid and/or repair communication breakdowns during L2 interaction, NNSs obtain successful comprehension by searching for the most logical possibilities via the effective use of context and prior knowledge (e.g., topic familiarity, cultural background) (Goh, 2002). Through this, these learners are assumed to become more aware of the importance of using various comprehension strategies, such as selective attention, problem solving, planning, evaluation and monitoring, all of which are essential for the development of listening ability (Vandergrift & Tafaghodtari, 2010).

Over the past 50 years, a number of empirical studies have extensively expounded the effect of interaction on L2 comprehension development with pre- and post-test designs. There is ample evidence showing that L2 learners who negotiate for meaning through actual interaction with NSs tend to show relatively immediate and large gains in language ability, especially compared to learners who are merely exposed to pre-modified and simplified input (for a comprehensive review, Mackey, 2012). Though revealing, these previous studies have brought to light several methodological problems which make further investigation worthwhile. According to Mackey and Goo's (2007) research synthesis, for example, most of studies have involved a very brief amount of interactional treatment (< one hour), corresponding to a general lack of longitudinal work in the field of SLA (Ortega & Byrnes, 2008). Furthermore, these studies have exclusively focused on the acquisition of specific L2 vocabulary and grammar features, without giving much attention to the development of pronunciation, fluency and listening comprehension skills. Given that L2 speech research has examined *intentional* (rather than *incidental*) focus on form via explicit instruction, form-focused tasks and interactional feedback (Saito, 2012; Vandergrift & Tafaghodtari, 2010), it remains unclear how L2 learners can improve their global

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

oral and listening proficiency through negotiation for meaning during natural conversations with NSs.

To advance the existing literature on this topic, we conducted an experimental research with a pre/post-test design focusing on inexperienced Japanese college students' L2 English speech learning (comprehension, production) in a foreign language context, where L2 use is extremely limited outside of classrooms. To create communicatively authentic conversational opportunities under such restricted L2 learning conditions, the participants engaged in weekly, dyadic conversation exchanges with NSs in the US by way of a video-conferencing tool (i.e., *Google Hangouts*) beyond the regular curriculum over one academic semester (12 weeks). Unlike naturalistic environments, where L2 learners have access to a great deal of input and interaction on a daily basis (e.g., study-abroad), this specific research setting—video-based interaction in foreign language classrooms—could be considered as an interesting testing ground, particularly for longitudinal analysis of L2 interaction, as it allowed us to monitor the quality and quantity of their conversational experience throughout the experiments (see below).

In our precursor research (Saito & Akiyama, 2016), we reported that one academic semester of video-based L2 interaction activity was facilitative of various dimensions of the Japanese learners' spontaneous *production* ability development (e.g., comprehensibility, fluency and vocabulary). In this paper, we aimed to revisit the dataset to examine the effects of long-term interaction on the development of L2 *comprehension* ability. Following the interactionist account of L2 comprehension ability development (e.g., Ellis et al., 1994), we predicted that longitudinal interaction would enhance Japanese learners' comprehension (measured via a general listening proficiency test), as it could provide opportunities for comprehensible input and output during meaning-oriented discourse.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Method

Participants

A total of 30 freshman and sophomore Japanese students at a university in Japan (NNS learners), and 15 native speakers of English (NS interlocutors) participated in the current study.

NNS learners. For the purpose of recruitment, we created two kinds of flyers: (a) one on conversational activities through conversational exchanges with college students in the US; and (b) the other on vocabulary/grammar activities with the goal of attaining higher scores in Test of English for International Communication (TOEIC)—the same test format as the pre-/post-test measures. Among the interested participants, a total of 30 students were selected based on their relatively homogeneous L2 English backgrounds and status as conversationally inexperienced L2 learners.

First, they had studied English for six years only through foreign language education (typically with grammar-translation methods) since Grade 7 before entering the university without any extensive experience overseas (< 1 month). Second, the learners' exposure to L2 English was highly limited. All of them belonged to the same program (business and marketing) and were required to take three hours of language-focused lessons per week. As specified in the department syllabus, our casual classroom observations confirmed that the content of these lessons mainly consisted of reading and listening activities. Finally, all participants reported lacking any experience at private, conversational English language schools during the project, indicating that their L2 use with NSs was highly limited on a daily basis.

NS interlocutors. The NS interlocutors were L2 Japanese learners at several universities in the US who volunteered for conversational exchanges. Their L2 Japanese proficiency widely varied (beginner to advanced).

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Research Design

First, the NNS learners individually took the pre-test in the researchers' office (Week 1), and were assigned to either the experimental group ($n = 15$) or the comparison group ($n = 15$) based on their preference (videoconferencing for intercultural interaction with NSs vs. vocabulary/grammar exercise as prep for TOEIC). Next, they proceeded to an orientation session (Week 2) during which they received explanation on the procedure for the video-based conversation activities (for the experimental group) and the vocabulary/grammar exercise activities (for the comparison group) (Weeks 3-11). After finishing all sessions, they revisited the researchers' office to take the post-test (Week 12).

Experimental Group

The experimental treatment was organized as a language-exchange program. Each session lasted for 60 minutes with the first part in English (30 min) and the second part in Japanese (30 min). For both parts, a two-way information exchange task was used: NNSs prepared two visuals corresponding to a different theme each week (e.g., education, food) which they thought represented Japan and the US, and prepared two discussion questions for each visual. This task was chosen following Lee's (2002) suggestion that two-way exchanges of information on real-life topics that are thematic and minimally structured allow L2 learners to recycle ideas and reinforce language skills. Respecting the principle of learner autonomy in language exchange, we did not provide pre-determined visuals. Instead, NNSs were responsible for exploring cultural differences/similarities via the autonomous selection of the visuals.

Due to the time difference between Japan and the US, the participants were allowed to arrange the sessions according to their individual schedules, and engaged in the activities using their own computers. To keep track of their attendance and participation, the participants were

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

required to record and submit their own sessions to the researchers (by using a function of *Google Hangouts*) upon the completion of every session.

As operationalized in previous L2 interaction research (e.g., Mackey, 2012), NS interlocutors were trained to provide interactional feedback in the form of recasts, but only when they perceived the NNSs' errors to hinder the comprehensibility of their L2 speech. To maintain the communicative nature of the interaction, the interlocutors were told to pay primary attention to completing the tasks successfully, providing interactional feedback where natural and appropriate.

The quality of the L2 interaction treatment during the project was analyzed at the onset (T1: the second session/Week 4) and endpoint (T2: the eighth session/Week 8) by tallying three key elements of interaction: (a) the number of linguistic errors made by the NNSs; (b) the amount of feedback provided by the NS interlocutors in the form of negotiation strategies (after communication breakdowns) and recasts (following communicatively harmful errors); and (c) the number of attempts made by NNSs to correct their own errors (i.e., self-modified output) (e.g., Mackey, 2012).

Comparison Group

The 15 NNS learners in the comparison group were asked to complete weekly take-home assignments which consisted of a variety of vocabulary/grammar exercise activities, such as vocabulary recall tests (i.e., comprehension practice) and fill-in-the-blank grammar questions (i.e., production practice). The materials were piloted prior to the project, and each assignment took approximately 30 minutes. The learners submitted the assignment to the researchers for grading every week.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

The purpose of including the comparison group in the current study was two-fold. First, given that similar tests were used during the pre/post-test sessions (see below), examining the comparison group's performance was expected to reveal any test-retest effects. Second, since all of the NNS learners were enrolled in three-hour English lessons during the project, the comparison group served as a baseline to reveal the gains which Japanese learners typically exhibit after one semester of foreign language learning without any opportunities for L2 interaction.

Outcome Measures

Materials. In line with the L2 listening literature (e.g., Vandergrift & Tafaghodtari, 2010), the participants' comprehension was measured via a composite proficiency test (TOEIC). This type of assessment is assumed to tap into L2 learners' ability to process various kinds of realistic spoken language in order to "understand linguistic information unequivocally included in the text and to make inferences implicated by the content of the text" (Vandergrift & Tafaghodtari, 2010, p. 477), and is thus are as a good fit for the current project, whose main objective was to examine the impact of interaction as a whole (comprehensible input and output) on L2 learners' overall listening proficiency.

Two versions (A, B) of the TOEIC test were chosen from the New Official Workbook (Educational Testing Service, Vol.4), with Version A used for the pre-tests and Version B for the post-tests.¹ The participants marked their answers on a score sheet. Each test lasted for approximately 50 minutes. Each version consisted of three components—Part 2 Question-Response, Part 3 Conversations, and Part 4 Talks.

¹ The materials were reprinted by permission of Educational Testing Service. No endorsement of any kind by the copyright owner should be inferred.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

1. **Question-Response (30 questions):** The participants selected the best response (out of three options) for a single-sentence question (5-10 words). Their performance on this section was assumed to reflect basic-level L2 comprehension proficiency (understanding linguistically and semantically simple input).
2. **Conversations (30 questions):** The participants listened to a dialogue between a male and a female speaker (80-100 words), and selected the best response (out of four options) to three comprehension questions, respectively. This section was assumed to measure L2 learners' comprehension ability of interactional speech with frequent turn taking (20-25 words per turn).
3. **Talks (30 questions):** The participants listened to a business announcement spoken by a single person (80-100 words), and selected the best response (out of four options) to three comprehension questions, respectively. This section was assumed to tap into the NNS learners' advanced comprehension proficiency (understanding of linguistically and semantically complex input).

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Table 1. *Lexical, Grammatical and Discoursal Characteristics of Aural Texts in the Comprehension Test*

Task type	Part 2: Question-Response		Part 3: Conversations		Part 4: Talks	
	Comprehension of a short question (5-10 words)		Comprehension of a dialogue (80-100 words)		Comprehension of a talk (80-100 words)	
Test version	A (pre)	B (post)	A (pre)	B (post)	A (pre)	B (post)
A. Vocabulary						
Diversity (Measure of Textual Lexical Diversity)	112.1	113.2	94.2	94.2	124.5	113.3
Concreteness	372.1	364.3	373.1	363.0	379.5	385.8
Familiarity	579.6	584.5	581.6	584.5	570.2	575.7
Sophistication						
• 3000 word-families + proper nouns (%) ^a	97.8	98.0	98.8	98.5	98.1	97.4
• 6000 word-families + proper nouns (%) ^a	99.7	100	99.9	99.7	99.7	99.1
• Frequency (CELEX Log)	3.03	3.05	3.09	3.10	2.97	2.99
B. Grammar						
No. of words per sentence	5.79	5.15	12.19	11.36	14.87	15.89
Left embeddedness (the number of words before main verbs)	0.95	0.88	1.24	1.86	3.01	2.81
No. of modifiers per noun phrase	0.65	0.59	0.68	0.57	0.83	0.90
C. Discourse						
Connectives (incidence)	<i>N.A.</i>	<i>N.A.</i>	73.8	72.4	81.5	74.5
Cohesion						
• Stem overlap in adjacent sentences	<i>N.A.</i>	<i>N.A.</i>	0.11	0.15	0.31	0.31
• Semantic overlap in adjacent sentences	<i>N.A.</i>	<i>N.A.</i>	0.06	0.10	0.14	0.14

Note. ^a Both 3000- and 6000-word families were measured based on Vocab Profiler (Cobb, 2012); the other vocabulary, grammar and discourse factors were analyzed via Coh-Metrix (McNamara et al., 2014).

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Following the research framework of Révész and Brunfaut (2013), the processing difficulty of the aural text (versions A, B) was analyzed via Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014) and Vocab Profiler (Cobb, 2015). As summarized in Table 1, the three tasks—Question-Response, Conversations, Talks—noted somewhat similar lexical profiles of spontaneous spoken in English especially in terms of lexical frequency: 97-98% lexical coverage by 3000 word-families and proper nouns and 98-99% lexical coverage by 6000-word families and proper nouns (Nation, 2006).

At the same time, the three tasks differed in terms of the complexity of grammar and discourse. Question-Response featured less complex grammar structures than Conversations and Talks according to the number of words per sentence (5-6 vs. 12-16) and the number of words before main verbs (0.9-1 vs. 1-3). Conversations and Talks were also different in several respects. Not only did Conversations contain more frequent words than Talks, but the aural text of the former was less complex than that of the latter according to all grammatical complexity and discourse connective/cohesion factors.

In terms of the different levels of difficulty between the two test versions, Version A used slightly less familiar words than Version B did in Parts 1, 2 and 3 (familiar ratings: $M = 570.2-581.6$ vs. $M = 575.7-584.5$). However, no other consistent patterns were observed in terms of the other domains of the aural texts (lexical, grammatical features and discourse complexity).

Taken together, the three components (Question-Response, Conversations, Talks) had a lexical frequency range similar to that typically found in English conversational interactions (3000-6000 word families) (Nation, 2006), but at the same time were ranked by cognitive demand as follows: Talks > Conversations > Question-Response. The results of the text analyses indicated that Version A may have been slightly more difficult than Version B. To follow up on

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

the tentative pattern of test difficulty (Version A > Version B), we used the comparison group's performance as baseline data (see below).

Results

Based on the video-coded data at the outset and end of the project, we first explored the nature of the interaction treatment that the experimental group received. According to the descriptive statistics (summarized in Table 2), negotiation for meaning episodes (as a result of communication breakdowns) occurred only a few times per session ($M = 2.3$ times [5.3%] for T1, 2.2 times [4.1%] for T2), and the NS interlocutors selectively recasted only salient linguistic errors ($M = 7.8$ errors [17.6%], 4.9 errors [9.3%] per session). In response to such feedback moves, the NNSs attempted to modify their own errors with relatively high uptake ratio (47.9-71.4% for negotiation; 36.4-65.4 for recasts). The results here indicated that the NNS learners processed a certain amount of comprehensible input and output while maintaining their primary focus on meaning throughout the sessions.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Table 2

Overall Interaction Patterns of Total Errors, Negotiation Strategies and Recasts, and Attempts to Self-Repair

Errors		Feedback		Uptake				
		T1	T2	T1	T2			
All linguistic domains	<i>n</i> = 664	<i>n</i> = 795	No feedback	<i>n</i> = 34.1 (77.1%)	<i>n</i> = 45.8 (86.5%)			
			Recasts	<i>n</i> = 7.8 (17.6%)	<i>n</i> = 4.9 (9.3%)	Uptake <i>n</i> = 2.8 (36.4%)	<i>n</i> = 3.2 (65.4%)	
						No uptake	<i>n</i> = 4.9 (63.6%)	<i>n</i> = 1.7 (34.6%)
			Negotiation	<i>n</i> = 2.3 (5.3%)	<i>n</i> = 2.2 (4.1%)	Uptake	<i>n</i> = 1.1 (47.9%)	<i>n</i> = 1.5 (71.4%)
				No uptake	<i>n</i> = 1.2 (52.1%)	<i>n</i> = 0.6 (28.5%)		

Note. T1 = 2nd session, T2 = 8th session out of 9 sessions over one academic semester

Next, we investigated the longitudinal development of comprehension skills by the participating students in the experimental group, and compared it with those in the comparison group. Due to the relatively small size of the dataset ($n = 15$ per group), a series of nonparametric tests were conducted. The alpha level was set at $p < .05$ and adjusted to $p < .025$ via Bonferroni correction. With respect to pre-existing differences between the two groups, the results of Mann-Whitney tests showed that they were found comparable for Question-Response ($z = -1.33, p = .187$), Conversations ($z = -1.91, p = .056$), and Talks ($z = -1.97, p = .202$) at the beginning of the project.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Table 3

Descriptive Results of the Comprehension Test Scores over Time

		Pre-test (30 points)		Post-test (30 points)		Improvement (pre → post)		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>z</i>	<i>p</i>	<i>d</i>
Question-Response	Experimental Group	19.6	3.45	23.1	2.79	-2.85	.004*	1.15
	Comparison Group	18.0	3.76	19.6	3.77	-1.96	.049	0.42
Conversations	Experimental Group	18.6	4.80	23.5	3.44	-3.30	.001*	1.17
	Comparison Group	15.4	4.23	19.9	5.00	-3.01	.003*	0.97
Talks	Experimental Group	14.8	5.09	21.6	4.86	-3.18	.001*	1.36
	Comparison Group	13.0	3.96	16.8	6.63	-2.51	.011*	0.63

Note. * stands for a statistically significant improvement at a $p < .025$ level.

To examine the presence/absence of any significant improvement over time, a set of nonparametric Wilcoxon Signed Ranks tests were then performed for the experimental and comparison groups, respectively. The magnitude of their improvement over time (pre → post) was measured by Cohen's d analysis, as suggested by Plonsky and Oswald (2014). As summarized in Table 3, the comparison group noted somewhat limited improvement in their comprehension scores from pre- to post-tests (a significant gain was found in Conversations and Talks but not in Question-Response). In contrast, the experimental group significantly increased their scores in all of three sections ($p < .025$), and, more importantly, the amount of their improvement could be interpreted as relatively large ($d > 1.00$) in keeping with the research standards in instructed SLA research (Plonsky & Oswald, 2014).

Discussion and Conclusion

The current study took a first step towards examining the effectiveness of video-based interaction on the longitudinal development (one academic semester) of Japanese college students' L2 English comprehension skills in a foreign language setting. Building on recent L2 listening research (e.g., Révész & Brunfaut, 2013; Vandergrift & Tafaghodtari, 2010), participants' performance was analyzed via a composite test tapping into various dimensions of

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

L2 comprehension proficiency, such as the processing of short and simple input (Question-Response), interactional input with frequent turn taking (Conversations), and long and complex input (Talks). According to the results of the pre-/post-test data, both the experimental and comparison groups equally developed their listening skills over time. Yet, analysis of the effect sizes revealed that the amount of the experimental group's improvement was equally large under all task conditions ($d > 1$), although the comparison group's improvement was limited in two out of three tasks (Conversations, Talks) with small-to-medium effects ($d < 1$) (Plonsky & Oswald, 2014).

The results presented here allow us to assume that the extracurricular drill activities and classroom listening activities received by the comparison group may be effective for L2 lexicogrammar learning, which could in turn help develop L2 listening skills to some degree, even without any opportunities to interact with NSs. Other interpretations could be methodological: (a) the gains resulting from the comparison group may simply indicate test-retest effects (taking the similar tests twice can result in improved comprehension scores), as frequently reported in many English proficiency test settings (Liao & Qu, 2010)²; and/or (b) the pre-test materials (Version A) could have been more difficult than the post-test materials (Version B), as suggested by our text analysis presented earlier.

Compared to the comparison group's performance as a baseline, it is crucial to emphasize that the participants in the experimental group attained significant and robust comprehension skill acquisition with large effects regardless of task condition (i.e., different levels of processing

² It is important to remember here that the comparison group did not improve on the simplest part of the listening comprehension test (Question-Response), but improved on the two more difficult sections (Conversations, Talks). This in turn corresponds to previous research evidence that the magnitude of test-retest effects is relatively large especially when L2 learners participate in repeated, cognitively demanding tasks (e.g., Kim & Tracy-Ventura, 2013).

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

difficulty: Question-Response < Conversations < Talks). As shown in the video-coded data of the experimental group, the participants received feedback (recasts, negotiation) approximately 10 times per session (30 min). The nature of interaction could be comparable to other observation studies which descriptively looked at the feedback frequency during meaning-oriented interaction (e.g., Mackey, 2012). This in turn suggests that the NNS learners could work on the development of their L2 listening skills with constant and immediate assistance from their NS partners for prolonged periods of time. That is, the NS interlocutors occasionally led the NNS learners to attain better comprehension via negotiation strategies in the case of communication breakdowns (i.e., comprehensible input). Different from existing L2 listening studies (e.g., Vandergrift & Tafaghodtari, 2010), where L2 learners focus on incoming input in a receptive mode, the NS interlocutors also encouraged the NNS learners to repair certain communicatively-harmful errors in production via recasts (i.e., comprehensible output). Self-modified output is believed to push NNSs to align their linguistic representations more closely with native speakers' models (Swain, 2005), which will, in theory, stimulate the development of more advanced and robust comprehension ability in a complementary fashion (Field, 2008).

Whereas our study shed some light on the acquisitional value of video-based interaction as a whole from a longitudinal perspective, the findings should be considered as tentative in nature, and thus need to be replicated with a larger sample size in different L1/L2 contexts. In particular, future studies need to scrutinize precisely which aspects of L2 interaction treatment are relatively important for L2 speech learning in the long run by controlling a range of affecting factors, such as different types of interactional feedback (recasts, negotiation), opportunities for repair, quality and quantity of turn taking, NSs' specific linguistic use, and NNS learners' proficiency levels (see Mackey, 2012). Additionally, the experimental group's gain could be

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

related to L2 learners' individual differences in motivation, willingness to communicate, aptitude, cognition (working and phonological memory), and personality traits (extroversion vs. introversion). Finally, more qualitative analyses may be needed to examine whether certain students particularly benefit from L2 interaction of this kind, because they have practiced, or prepared earnestly for each conversation session with their NS partners.

References

- Cobb, T. (2015). *The compleat lexical tutor*. Retrieved from <http://www.lextutor.ca/vp/>
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension and the acquisition of word meanings. *Language Learning*, 44, 449–491.
- Field, J. (2008). Emergent and divergent: A view of second language listening research. *System*, 36, 2–9.
- Goh, C. (2002). Exploring listening comprehension tactics and their interaction patterns. *System* 30, 185–206.
- Jensen, E., & Vinther, T. (2003). Exact repetition as input enhancement in second language acquisition. *Language Learning*, 53, 373–428.
- Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41, 829-840.
- Lee, L. (2002). Enhancing learners' communication skills through synchronous electronic interaction and task-based instruction. *Foreign Language Annals*, 35, 16-24.
- Liao, C. W., & Qu, Y. (2010). Alternate forms test-retest reliability and test score changes for the TOEIC speaking and writing tests. *TOEIC Compendium*.
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126-141.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition. Vol. 2: Second language acquisition* (pp. 413-468). New York: Academic Press.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

- Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford, UK: Oxford University Press.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82.
- Ortega, L., & Byrnes, H. (Eds.). (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.
- Plonsky, L., & Oswald, F. L. (2014). How big is ‘big’? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878-912.
- Révész, A., & Brunfaut, T. (2013). Text characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition*, 35, 31-65.
- Rost, M. (2011). *Teaching and researching listening*. London: Longman.
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 842-854.
- Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, 67(1).
- Vandergrift, L., & Tafaghodtari, M. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, 60, 470-497.

ROLE OF INTERACTION IN L2 LISTENING COMPREHENSION

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18, 49–68.