Toward Personalized Scaffolding and Fading of Motivational Support in L2 Learner–Dialogue Agent Interactions: An Exploratory Study

Emmanuel Ayedoun, Yuki Hayashi, and Kazuhisa Seta

***Abstract*—This article proposes a computer-based approach to effectively enhance second language learners' willingness to communicate in the target language. To do so, we implemented a conversational agent embedding a dialogue management model based on two conversational strategies (i.e., communication strategies and affective backchannels), serving as scaffolds for enhancing learners' willingness to communicate in the target language. Here, we report on differences observed among second language learners' preferences for both conversational strategies according to their initial level of willingness to communicate and on variations of their willingness with respect to such differences. Although we found that most students generally preferred a combination of both strategies, learners' preferences and the effects of the support provided by these strategies varied according to their level of willingness to communicate. Learners with lower willingness to communicate tended to prefer affective backchannels, whereas those with higher willingness to communicate seemed to favor communication strategies. These results were consistent with post-test results, which showed that learners' expected willingness to communicate tended to be higher after interacting with systems embedding their preferred strategies. In sum, these results are preliminary evidence of the meaningfulness of accounting for such learners' preferences in adaptively using and fading the strategies employed by conversational agents to motivate second language learners to communicate in the target language.**

***Index Terms*—Motivationally intelligent educational systems, scaffolding and fading, willingness to communicate in second language.**

I. Introduction

SCAFFOLDING has been identified as one of the most effective instructional procedures [1], [2], especially when the learner needs assistance to perform a task [3]. In recent decades, scaffolding has gained widespread popularity in several education-related fields [4] and has been considered a key attribute of intelligent tutoring systems. Despite the potential of such systems, they have often been limited in their ability to produce skills that are transferable to the real world [5], [6]. For instance, one area where traditional educational systems are struggling to make a significant impact on learning outcomes is second language acquisition [7]. One approach in overcoming this problem may be a stepwise reduction of the amount of scaffolding provided by the system to learners so as to develop learners' ability to solve problems on their own (self-reliance); fading is a means of achieving this [8]. The idea and importance of combining scaffolding with fading in instructional settings has been stressed in the literature, and instructional theories such as the modeling-scaffolding-fading [9] paradigm have been proposed. This paradigm describes a three steps instructional process where the tutor first models a desired skill or solution, then has the learners try out the skill while the tutor provides feedback and scaffolds; finally, as the learners become more and more able to perform unaided, the tutor eventually fades from the process. In this article, we especially focus on the last two concepts, "scaffolding and fading" as these two directly address learners' actions within the task, whereas modeling solely relies on the learner observing and the design of learning support.

The main purpose of second language learning is to provide learners with the ability to autonomously convey their intended meaning effectively in the target language and, by extension, to facilitate exchanges between people from different countries. However, many second language learners do not display such ability even after studying the target language for several years. Research [10] suggests that the key factor for ensuring a spontaneous and sustained use of a second language is the willingness to communicate (WTC), which is defined as a "readiness to enter into discourse at a particular time with a specific person or persons, using a second language" [10, p. 547]. WTC studies have shown that learners displaying high WTC are more likely to show greater improvement in their communication skills [11] and acquire higher levels of language fluency [12] than those with low WTC.

In our previous work [13], [14], a conversational agent enhanced with a set of specific conversational strategies, i.e., communication strategies (CS) and affective backchannels (AB) dedicated to carrying out WTC-friendly conversations with learners in an English-as-a-foreign-language context was proposed. CS are techniques, such as *Repetition* of a previous utterance, that are used by the dialogue system to help overcome communication pitfalls (e.g., difficulty in understanding the agent's utterances or answering the agent's questions) that learners may encounter during interactions. AB are expressions

Manuscript received May 20, 2019; revised November 6, 2019 and March 23, 2020; accepted April 12, 2020. Date of publication April 23, 2020; date of current version September 16, 2020. This work was supported in part by JSPS KAKENHI under Grant 17K18017, Grant 18H03345, and Grant 18J22445. (*Corresponding author: Emmanuel Ayedoun.*)

The authors are with the Graduate School of Humanities and Sustainable Systems Science, Osaka Prefecture University, Osaka 599-8531, Japan (e-mail: eayedoun@ksm.kis.osakafu-u.ac.jp; hayashi@kis.osakafu-u.ac.jp; seta@mi.s.osakafu-u.ac.jp).

Digital Object Identifier 10.1109/TLT.2020.2989776

such as "*Come on, don't give up!*" or "*Great, well done!*" used to explicitly convey thoughtful support to learners by congratulating, cheering up, or empathizing in accordance with the interaction state. An experimental evaluation of the system suggested that combining these communicative and affective conversational strategies may be effective in fostering L2 learners' WTC [14].

However, it also well-acknowledged that willingness to communicate in a second language (L2 WTC) is influenced by a complex interplay of learner-specific internal and external variables [15]–[17]. Thus, while the combined effect of CS and AB may predict an increase in WTC among L2 learners, it might be erroneous to conclude that all L2 learners would perceive or benefit similarly from the effects of such scaffolds. For instance, it is conceivable that some learners might benefit more from AB, while others might benefit more from CS, based on their developmental level. Gradual removal of the scaffolds provided by the system is necessary to enable learners to autonomously maintain their WTC at a suitable level and engage in language production on their own.

In this article, we take a closer look at learners' perceptions of the support provided by each of the above-mentioned strategies or combination of strategies and investigate differences in preferred strategies according to learners' WTC levels. Then, we analyze how these differences are connected to the scaffolds' outcomes in terms of an increase in expected WTC among learners. Finally, we discuss the feasibility of and propose approaches for achieving a tailored deployment and fading of these scaffolds according to learners' preferences and WTC level.

II. Literature Review

A. Scaffolding and Fading in Learning Environments

Scaffolding has traditionally been characterized as a process in which an expert helps a learner perform a specific task or achieve a specific objective [3] by providing, for example, hints or pumps when necessary during a problem-solving task. It has proven a particularly interesting and promising way to support teaching and learning practices [18]. Several studies have investigated and proposed design guidelines for scaffolding learning in traditional environments. Some of these works focused on experts or teachers' strategies [19], [20] and others on learners [21], [22].

Notably, the concept of scaffolding is often associated with the idea of fading [23], which makes it "conceptually and operationally distinguishable from other types of assistance" [18]. Fading is defined as the gradual reduction and eventual elimination of scaffolds [24]. Scaffolding combined with fading supports the learning and development of independent skills by facilitating successive levels of competence through which scaffolds are withdrawn to promote the learner's independent functioning [25]. In other words, effective learning support interactions and can be considered a two-step process of providing and then withdrawing expert support as learner competence increases. It follows that fading is essential to learning support in educational or learning contexts.

In the area of technology-enhanced learning, the role of scaffolding has attracted the attention of researchers [26], [27]. Several studies suggested the meaningfulness of employing scaffolding tools to support various aspects of learning such as argumentation in ill-structured problem solving [28], [29], reasoning in social sciences [30], searching and sorting of science information [31], critical thinking [32], and mathematical problem solving [33]. Nevertheless, to be effective, scaffolding tools need to be carefully crafted; otherwise, they may inadvertently misdirect learners [18], obstruct effective learning [34], or over-instruct learners [35]. Hence, research [18] suggests that technological scaffolding must be designed according to learners' developmental and cognitive needs for the specifically targeted learning context. In other words, it is desirable for the scaffold design to be consistent with the target learning context, as well as learners' characteristics [36]. However, logistical and conceptual difficulties prevent most learning support systems from embedding scaffolds that change dynamically as individual circumstances evolve or from achieving fading that is predicated on a learner's needs or performance [37]. The question of how best to balance providing and withholding of support has been framed within the field of educational technology as the "assistance dilemma" [38]. Research [39] also highlights the absence of reference to empirical accounts of fading in most scaffolding systems, concluding that such systems appear to function not as scaffolds-with-fading, as is desirable, but rather as scaffolds-for-performance.

Interestingly, it has been suggested that computer-based scaffolds might be designed with a capacity to be faded, but this does not necessarily need to be a system function that happens dynamically through successive iterations by the system itself [40]. Such scaffolds can be embedded within a learning context, and fading can be determined heuristically by a human expert based on their assessment of learner performance, or learners could determine this for themselves. For instance, in the context of the SQL-Tutor system, fading was implemented to enhance students' ability to independently select appropriate problems to work on [7]. In their experiments, the researchers used three different versions of the SQL-Tutor: one in which problems were always selected by students; another in which problems were selected by the system; and finally, a faded one in which the control over the selection of problems was progressively shifted to students as their level increased over time. The results of their investigations suggested that the students in the faded group improved their selection accuracy and performed better at selection than other students.

B. Individual Differences in Language Learning

Language learning styles characterize the consistent and rather enduring traits, tendencies, or preferences that may differentiate one learner from another [41]. Language teachers need to be aware of such differences as a single L2 methodology cannot reasonably fit an entire class filled with students who have a range of stylistic and strategic preferences [42]. In

addition, in order to maximize learning outcomes, it is important for teachers to modify their teaching methodology according to the factors related to the individual tendencies, preferences, learning strategies, personality, etc. of their students [43], [44]. All these positions tend to confirm a link between learners' styles and preferences and effectiveness of the teaching methodology. This further suggests the need to take a personalized approach to language teaching to help learners reach higher learning outcomes [45], [46]. In a similar vein, one can assume that depending on the level of WTC reached by L2 learners, it might be desirable to provide them with support that better fits their preferences.

However, in the field of second language acquisition, research on individual differences has generally been focused on explaining the processes of learning and acquisition, rather than looking for practical ways to improve them [47]. As a result, even though all language teachers would quickly agree that learners differ from one another, most language teaching materials have assumed that all learners are the same [48]. This is emphasized in the following terms: "While the learner has not been ignored in second language acquisition research, more attention has been paid to characterizing an acquisition process that is common to all learners" [49, p. 12].

More generally, the major pedagogical implication of learners' differences (i.e., personalization of learning contents) is difficult to realize in usual classroom settings because it would result in additional workload for teachers [47]. This would also require teachers to have the necessary motivation for such work as well as enough pedagogical resources to help each learner receive such personalized instruction. The need to overcome such temporal and contextual exigencies became one of the key factors of the development of so-called adaptive learning support systems.

C. Adaptive Language Learning Support Systems

The purpose and promise of adaptive learning support technologies are to facilitate online instruction that is personalized to the needs of individual learners [50], [51]. The adaptive nature of such systems is not necessarily determined by their level of sophistication. For instance, although complex systems can use artificial intelligence and machine learning algorithms to analyze and find patterns in the large amounts of data being captured, the simplest adaptive systems are based on predefined models describing how the data from an individual learner can lead to personalized learning pathways.

In the intelligent tutoring systems research field, representation of student behaviors and knowledge (i.e., student model) is used to select problems not yet mastered and to provide feedback for each individual learner. Such adaptive learning content selection and assistance is regarded as an efficient way to improve learning achievement [52]–[55].

Although the use of adaptive technologies in learning is growing rapidly, they are still in their infancy as far as language learning is concerned [56]. Furthermore, a major research problem with adaptive language learning is that such technologies have not been deployed for a sufficiently long time to enable reliable longitudinal findings. Moreover, there are many challenges in developing fully adaptive learning systems since it is very difficult to restrict the number of variables to account for the personalization of instructional contents, especially in language learning, which is not a linear process [57], [58]. Yet, research on adaptive language learning has been based on the models for other academic subjects such as mathematics, where adaptive technology is used to determine paths through predetermined content [56]. As a result, successful systems in the area of language learning tend to personalize learning contents by focusing on learners' performance [59]–[61]. For example, some vocabulary or grammar apps provide automated spaced repetition, which tailors the order and frequency with which learning contents are presented to the learner by using the rate of correct or incorrect answers [62]. In this sense, such systems are personalized and adaptive learning tools, albeit of an admittedly fairly simple nature [56]. Nevertheless, the field has received increasing attention from researchers, and recently developed systems hint at the substantial contribution that carefully designed systems can make to the computer-assisted language learning field. For instance, CSAL AutoTutor is an interactive intelligent tutoring system that employs three-party conversations or trialogues between the learner, a tutor agent, and a peer agent to foster both shallow and deep comprehension of text in low-literacy readers. The system has yielded promising preliminary results according to a recent study [63]. CBA is a conversation-based English assessment system that allows learners to converse in natural language with animated agents, and a recent study has proven that the system is able to assess students' English capabilities comparably to human interviewers [64]. Note that both systems dynamically adapt to students by providing aid to student's based on their assessed knowledge level. In other terms, within these systems, scaffolding is provided to students when needed and less scaffolding to students who display higher levels of knowledge.

Recent studies also point out that educational systems dedicated to language learning should go beyond learners' performance and incorporate observation of students' opinions, preferences, as well as contextual information to provide more accurate learning support [65], [66].

To sum up, the focus of adaptive language learning systems has been mostly on leveraging the technology to support areas such as grammar, vocabulary, and reading. Furthermore, available adaptive language support systems tend to rely exclusively on quantitative learning data, while a more complete picture of student learning may also necessitate an account for qualitative data, especially when it comes to dealing with language production or learner's motivation toward communication, which is precisely the main focus of this article.

D. Conversational Agents to Enhance L2 Learners' WTC

L2 WTC is believed to play a major role in learners' decision to use their second language for communication when given the opportunity to do so. Researchers [10] proposed a heuristic model of variables influencing L2 WTC through

Fig. 1. System interface featuring Peter in a restaurant context.

which they argued that L2 learners' intention to engage in L2 communication, rather than linguistic competence, is determined by an interplay of numerous psychological, affective, contextual, and communicative factors. Their model has inspired several empirical studies that shed light on the consistent influence of contextual and emotional variables on L2 WTC [67]–[69]. For instance, some of these studies have shown that to enhance L2 learners' intention to interact in the target language, it is important to increase their self-confidence and reduce their anxiety since there is much evidence that if communication apprehension recedes, an individual's perceived communicative competence is likely to be higher, leading to a higher level of WTC [70], [71]. To achieve such affective, and motivational support, we proposed the system shown in Fig. 1 [14]. The system is a conversational agent that uses two conversational strategies in realistic daily conversation scenarios, namely task-based CS and affect-based AB, as scaffolds to motivate L2 learners to communicate in English. The originality of such an approach is related to its use of a pair of two conversational strategies which enable the system to take care of both aspects related to communicative breakdowns that often occur in L2 learners–agent interactions and those related to affective variables influencing L2 WTC, in accordance with MacIntyre's WTC model [10]. By enabling the conversational agent to make use of CS, our idea is to enhance the agent's own strategic competence to release learners from the challenging and WTC-inhibiting burden of resolving communication pitfalls by themselves. By identifying a novel category of backchannels (i.e., AB), our aim is to foster the agent's ability to convey empathetic and WTC-friendly support to learners. In other words, the rationale of implementing such strategies is to increase L2 learners' confidence via CS and reduce their level of anxiety toward communication via AB. Some examples of how CS and AB are implemented in the proposed system are shown in Table I. Detailed information on dialogue management, as well as firing policies of CS and AB, can be found in [14]. In general, when a communication pitfall occurs, the system first selects a given category of AB to either reassure, encourage, or empathize with the learner; then, it targets an appropriate category of CS to help move the conversation forward according to the nature of the communication pitfall or dialogue state. When a given category of AB (e.g., *Encouraging AB*) is selected, a corresponding AB (e.g., *Come on! Don't be shy.*) is chosen stochastically from the options in that category. On the other hand, instances of CS are triggered in a heuristically predefined order. For example, when the conversation state, *The learner is NUNA (i.e., Nor able to Understand, Nor to Answer)*, is detected, the system first makes use of *Repetition*. If the same state is detected on the next turn, then *Simplification* is applied. If the same state is detected on the following turn, *Code-switching* is used. The purpose of this is to make it progressively easier for learners to overcome their current difficulty when the conversation is stuck in a given state.

In [14], learners' expected WTC were examined after interacting with one of the following versions of the system: 1) an agent featuring both CS and AB; 2) an agent featuring only CS; and 3) an agent featuring only AB. The results suggested that the system combining CS and AB was the most effective in terms of WTC outcomes and also the most preferred by learners.

However, the reasons behind such experimental results may not be homogeneous among all learners. We cannot rule out the hypothesis that although most learners preferred the combination of CS and AB, the relative effect of one strategy or another may depend on complex personal factors specific to each learner (e.g., preference of CS over AB, level of WTC). A first step toward implementing an L2 learner-adapted WTC support in our current system may be to explore ways to enable the dialogue agent to account for such learners' differences in dialogue management.

E. Contribution and Novelty of this Article

In light of the abovementioned studies, although the importance of scaffolding is well acknowledged, proposing systems that address variations in student expectations and interests and providing an adaptive level of support remain major challenges in the field of technology-enhanced education. For instance, most learning support systems, with a few exceptions such as Affective AutoTutor [72], address this matter from a cognitive standpoint and tend to rely exclusively on learners' performance data, which may distort learning activities through a lack of consideration of motivational and contextual factors affecting learning.

As previously discussed, the literature is rich in studies that have actually focused on enhancing systems' abilities to assess and tailor the presentation of scaffolds according to student profiles [30]–[33]. On the other hand, aspects related to fading or scaffold withdrawal have not been sufficiently discussed in the literature. Fading seems particularly difficult to achieve in computer-mediated learning since it requires the system to identify a suitable timing and pace of withdrawal of the scaffolds, while making sure that the learner can still autonomously perform the target learning task. In particular, when it comes to providing support for motivational variables affecting (language) learning, the literature has been quite elusive on scaffolding tools that are sensitive to learners' needs, differences, characteristics, etc.

TABLE I
EXAMPLES OF CS AND AB IMPLEMENTED IN THE CONVERSATIONAL AGENT

	Strategy	Description	Example
CS	Simplification	Use an alternative or a shorter term, which express the meaning of the target lexical term	Agent: May I have your order? Learner: ... (silent) Agent: Order please.
	Asking clarification	Request an explanation of an unfamiliar meaning structure	Learner: One xxx please. Agent: What do you mean?
	Suggesting AP (Answer Pattern)	Provide an example of an answer that could fit the current discourse context	Agent: What would you like to drink? Learner: ... (silent) Agent: For example, you may say, "one beer please" to order a beer.
AB	Encouraging AB	Employed when the learner seems to hesitate to the extent that he/she remains silent	• Come on! Don't be shy. • You can do it. • ...
	Sympathetic AB	Employed when the learner's utterance does not match the agent's expectations	• Sorry, I didn't get you... • Oops, I'm afraid I missed something... • ...
	Reassuring AB	Employed when the learner seems to face many difficulties in the conversation	• Don't worry, dear. • You'll be fine. • ...

To the best of our knowledge, no previous work has proposed a concrete approach on how traditional intelligent tutoring systems should implement scaffolding and fading of support when targeting affective and motivational aspects of learning. The major contribution of this article is that we extend the discussion on provision and withdrawal of scaffolds from the traditional (meta) cognitive aspects to the motivational dimension of learning. Concretely, through an analysis of the results of experimental studies, we show that it might be possible and meaningful to consider L2 learners' preferences of conversational strategies toward providing them an effective scaffolding and fading that are adapted to their level of WTC in the target language (i.e., WTC level).

Furthermore, we particularly argue that there might be a conceptual difference between fading of support directed to the cognitive aspects of learning on one hand and those targeting motivational variables of learning on the other. For instance, according to previous research [73] on phenomena such as "gaming the system" behavior within intelligent tutoring systems that targets cognitive variables of learning, learners' inclination or preference for a given category of scaffold is not warranted for obtaining desired learning outcomes. Hence, in such systems, the design of scaffolds and their fading are somewhat constrained by the necessity to prevent learners from taking advantage of the existing scaffolds to achieve systematic task completion instead of mastering learning contents as is desired. However, in the case of support directed at motivational variables of learning (i.e., the aim of this article), there seems to be concordance between use of preferred scaffolds by learners and achievement of desired learning outcomes, as we will explain later on in this article. This is an interesting finding since it hints at the meaningfulness of referring to learners' preferences to obtain both scaffolding and fading of motivational support. More concretely, when dealing with motivational variables affecting learning, fading could consist of removing or lowering scaffolds that do not correspond to learners' preferences, in order to make them feel as comfortable as possible with the learning environment. In the last sections of this article, we discuss a way to achieve such gradual removal of motivational scaffolds (i.e., conversational strategies) as learners' WTC increases.

Along these lines, we aim to provide novel paths of discussion on the feasibility and challenges of scaffolding as well as fading in terms of motivational factors that affect language learning with a computer-based system. Finally, the qualitative approach of our work is consistent with the recommendations of notable studies in second language acquisition research [74]–[76] which have identified the need for greater emphasis on social, affective, and conceptual dimensions of the language learning process.

III. OBJECTIVE AND RESEARCH QUESTIONS

This article investigates the following research questions.

RQ1 What are the differences in L2 (second language) learners' perceptions or preferences of the WTC support provided by the system?

RQ2 How do WTC outcomes vary according to such differences in L2 learners' preferences?

RQ3 Which approach could be suitable to achieve an adaptive WTC scaffolding that accounts for such L2 learners' differences?

Hence, this article aims to examine differences in L2 learners' perception of the support provided by the dialogue agent via CS and AB, as well as the implications of such

TABLE II
OVERVIEW OF THE EXPERIMENT FLOW

Phase		*Steps*	*Group 1* (n=10)	*Group 2* (n=10)	*Group 3* (n=10)	*Group 4* (n=10)	*Group 5* (n=10)	*Group 6* (n=10)
Phase 1: Pre and Post WTC Evaluation		Step 0	First WTC questionnaire (Pretest)					
		Step 1	Warm-up interaction with the system					
Phase 2: Evaluation of System Version Preference		Step 2	CS+AB	CS+AB	CS	AB	CS	AB
		Step 3	Second WTC questionnaire (Posttest)					
		Step 4	CS	AB	CS+AB	CS+AB	AB	CS
		Step 5	System preference survey					

perception on the feasibility of providing those learners with more adapted WTC support.

Since WTC is believed to be influenced by an interplay of affective factors (mainly *anxiety* and *self-perceived communicative competence*) specific to each learner, we hypothesize that L2 learners' developmental level of WTC may influence the preference for a given category of motivational scaffolds. Such learners' preference tendencies, if established, may further help us tailor or balance the use of CS and AB to maximize outcomes in terms of WTC. Furthermore, we discuss the following.

RQ4 Which approach to follow for fading the support provided by the system and making learners more independent in terms of WTC?

IV. METHODOLOGY

A. Conversational Agent

We used the same embodied conversational agent as in [14], which allows for spoken dialogues between the conversational agent, personified as Peter, and a learner in a restaurant context. The conversation scenario begins with an entrance scene where learners are welcomed by Peter. After checking whether they have a reservation, they are guided to a table in their preferred area (i.e., smoking, non-smoking). From there, learners can call Peter anytime, ask for the menu, order drinks and dishes, and request the bill. During the interaction, learners are able to answer Peter's questions or take the initiative to ask questions or place orders.

B. Study Design

We designed two-phase experiments to compare learners' WTC results across different versions of the system (Phase 1), and to examine their preference after interacting with different versions of the system (Phase 2), as shown in Table II.

During Phase 1, we gauged learners' WTC by administering a widely used survey (Cronbach $\alpha = .88$) [77], before (Step 0) and after (Step 3) their first interaction with the system (Step 2). The WTC surveys consisted of a set of self-report questions targeting three variables: *confidence*, *anxiety*, and *desire to communicate (desire)*, which are considered to be the immediate precursors of WTC [78]. Participants were asked to rate 30 scenarios (e.g., *making a telephone call to make a reservation at a hotel in an English-speaking country*) related to using English in various circumstances on a four-point Likert scale (0–3). Note that the two WTC questionnaires (i.e., pretest and posttest), although asking similar questions, were different in that the first asked about learners' actual WTC, whereas the second asked about learners' expected WTC if they were given the opportunity to interact with the system as much as they wanted on a regular basis. In Step 2, three versions of the system—the normal version featuring both CS and AB (CS+AB), a second version featuring only CS (Communication Strategies), and a third version featuring only AB—were employed in the interactions to examine how the outcomes on participants' WTC varied with the system version. Participants interacted individually with the system in a room specially prepared for the evaluation and were given as much time as they wished to enjoy the conversation with the agent, until the end of the interaction. They were also informed that they were free to interrupt the interaction at any time they desired, but were requested to let us know beforehand.

To complete Phase 2, we allowed all participants to have a second round of interactions with another version of the system (i.e., different from the one used for their first interaction) (Step 4). We then conducted a survey to obtain feedback concerning their preference for the system's versions (Step 5). All participants were asked to choose which one of the two interactions (i.e., which version of the system) they preferred the most, as well as the reason for their preference. For example, participants in Group 1 had to choose between the CS+AB and the CS versions, those in Group 2 between the CS+AB and the AB versions, and so on for participants in other groups. To minimize the eventuality that learners' preference would be based on the order in which they interacted with different versions (i.e., order effect), their interactions with the system in each group were designed by applying the counterbalancing method [79].

C. Participants and Data Collection

Data from 60 university students (39 males and 21 females with at least 2 females in each group) were collected and used

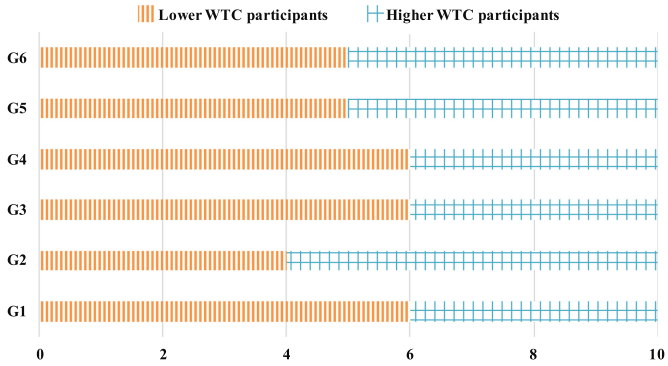

Fig. 2. Distribution of participants in the six groups according to their WTC level.

TABLE III
PARTICIPANTS' PREFERENCE FOR CONVERSATIONAL STRATEGIES ACCORDING TO THEIR WTC

Section	*Groups*	*Strategies*	*Lower WTC*	*Higher WTC*
A	Groups 1 and 3	CS	1	4
		CS+AB	11	4
B	Groups 2 and 4	AB	3	1
		CS+AB	7	9
C	Groups 5 and 6	AB	8	2
		CS	2	8

in this article. Experiment data for participants in groups 1 to 4 were obtained from [14]. We then ran additional experiments to collect data for those in groups 5 and 6. Identical to groups 1– 4, participants in groups 5 and 6 were also quite homogeneous in terms of language background; all of them were native Japanese speakers and none had lived in an English-speaking country. Participants were university students who had learned English as a second language for at least six years during their university and/or secondary education (i.e., junior and senior high school) years. They were informed that their participation in the study was voluntary and that the experiment results would be anonymized. To preserve uniformity of conditions across the two studies, we also rigorously ensured that the same experimental settings used in our previous work [14] were used for this new round. Moreover, a one-way ANOVA was conducted, which confirmed the homogeneity of initial WTC conditions (first WTC questionnaire) among the six groups. Specifically, the tests revealed no statistically significant differences among the groups in terms of initial confidence [$F(5,54) = 1.85$, $p = .12$], anxiety [$F(5,54) = 0.44$, $p = .81$], and desire [$F(5,54) = 1.36$, $p = .25$].

Note that whereas the focus of our previous work [14] was mainly to evaluate the effectiveness of the combination of CS and AB in fostering learners' WTC, our current study rather focuses on analysis of differences in learners' preferences of these conversational strategies and explores the feasibility of accounting for such differences to achieve personalized support to increase L2 learners' WTC.

V. RESULTS

A. Differences in Learners' Preference for Scaffolds by WTC

To investigate the differences in participants' preference for CS and/or AB, we analyzed the results of the system preference survey (Step 5) with respect to learners' initial WTC level. To that extent, all participants were labeled as either lower or higher WTC according to the results of their first WTC questionnaire (Step 0): participants who had better than average scores on all of their initial WTC precursors (*confidence*, *anxiety*, and *desire*) were labeled as higher WTC, and the others were categorized as lower WTC. The resulting distribution of participants by WTC level was relatively uniform across the six groups, as shown in Fig. 2.

Table III (A) shows the results of the system preference survey for groups 1 and 3, who interacted with both the CS and CS+AB versions of the system. Barnard's test for independence was conducted, indicating a relationship between learners' WTC level and their preference for CS or CS+AB ($p = .04$), with a medium (Cramer's $V = .47$) effect size according to Cohen's conventions for Cramer's V [80].

Table III (B) shows the results for groups 2 and 4, who interacted with both the AB and CS+AB versions. Barnard's test indicated a trend towards the relationship between learners' WTC level and their preference for AB or CS+AB ($p = .09$), with a medium (Cramer's $V = .42$) effect size.

Table III (C) shows the results for groups 5 and 6, who interacted with both the AB and CS versions of the system. Barnard's test confirmed a relationship between learners' WTC level and their preference for AB or CS ($p = .01$), with a medium (Cramer's $V = .6$) effect size.

In sum, these results indicated that participants' preference tendencies of the different system versions (i.e., CS, AB, or CS+AB) seem to be related to their WTC level.

B. Relationships Between Learners' WTC Level, WTC Outcomes, and Scaffolds

To investigate whether and how WTC outcomes vary according to learners' WTC level, we analyzed differences among lower- and higher WTC learners' results after interactions with the three versions of the system. To that end, a one-way ANCOVA was conducted to determine whether there were statistically significant differences between WTC posttest results, while controlling for pretest results. Post-hoc Tukey–Kramer tests were additionally conducted to further investigate the differences when they were statistically significant, as described below.

Fig. 3 shows results for pairwise comparisons of WTC scores across lower WTC participants, with an emphasis on observed differences according to the system version used by learners. There was a significant difference in lower WTC participants' expected *confidence* [$F(2,28) = 3.55$, $p<.05$], *anxiety* [$F(2,28) = 3.40$, $p<.05$] and *desire* [$F(2,28) = 3.39$,

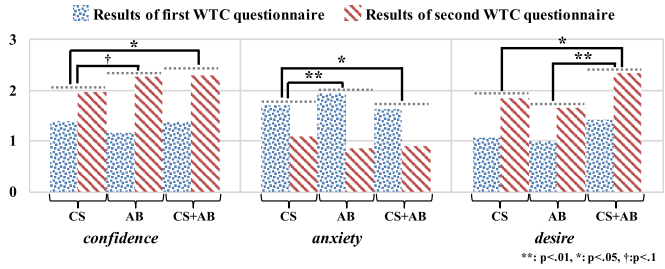

Fig. 3. WTC outcomes for lower WTC participants.

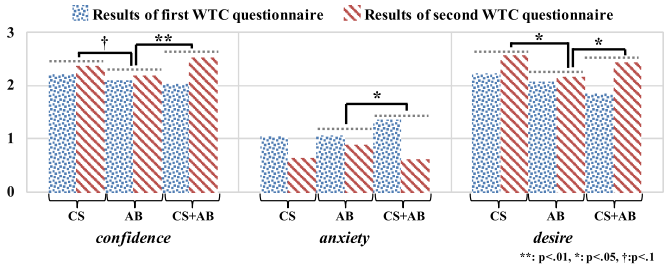

Fig. 4. WTC outcomes for higher WTC participants.

$p<.05$] among the three versions (i.e., CS, AB, and CS+AB). The post-hoc Tukey–Kramer tests showed that the expected *confidence* of lower WTC participants who interacted with the CS+AB and AB versions was higher than for those who used the CS version; $p = .04$, Cohen's $d = 1.11$ for CS+AB versus CS, and $p = .06$, Cohen's $d = 1.01$ for AB versus CS, respectively.

In addition, in terms of expected *anxiety*, results of lower WTC participants who interacted with the CS+AB and AB versions were better than those who interacted with the CS version; $p = .04$, Cohen's $d = 1.11$ for CS+AB versus CS, and $p = .006$, Cohen's $d = 1.43$ for AB versus CS, respectively.

Finally, in terms of expected *desire*, results of lower WTC participants who interacted with the CS+AB version were better than those who used the CS version ($p = .04$, Cohen's $d = 1.08$) and AB version ($p = .004$, Cohen's $d = 1.51$).

These results indicate that as far as lower WTC participants are concerned, the CS+AB and AB versions are more promising than the CS version in enhancing their WTC.

Regarding higher WTC participants, the one-way ANCOVA tests revealed that there was a significant difference in their expected *confidence* [$F(2,24) = 3.48$, $p<.05$] and *desire* [$F(2,24) = 4.97$, $p<.05$], and a trend toward a significant difference for *anxiety* [$F(2,24) = 2.88$, $p<.1$]. Post-hoc Tukey–Kramer tests were additionally conducted to further investigate the differences when they were statistically significant, as described below.

Fig. 4 shows results for pairwise comparisons of WTC scores across higher WTC participants with an emphasis on observed differences according to the system version used by learners. The post-hoc Tukey–Kramer tests showed that the expected *confidence* of higher WTC participants who interacted with the CS+AB and CS versions was higher than for those who used the AB version; $p = .004$, Cohen's $d = 1.63$ for CS+AB versus AB, and $p = .09$, Cohen's $d = 1.05$ for AB versus CS, respectively. In addition, in terms of expected *anxiety*, results of participants who interacted with the CS+AB were significantly better than those who interacted with the AB version ($p = .01$, Cohen's $d = 1.40$). Finally, in terms of expected *desire*, results of lower WTC participants who interacted with the CS+AB and CS versions were better than for those who interacted with the AB version; $p = .04$, Cohen's $d = 1.20$ for CS+AB versus AB, and $p = .03$, Cohen's $d = 1.28$ for AB versus CS, respectively.

These results suggest that as far as higher WTC participants are concerned, the CS+AB and CS versions are in most cases more promising than the AB version in enhancing learners' WTC.

To sum up, the analysis of WTC outcomes with respect to participants' WTC level suggests that: CS+AB and AB versions seem to work better for lower WTC participants, while for higher WTC participants, the most effective system versions seem to be the CS+AB and CS versions.

In addition, note that even though no time constraints were placed on the participants, we could not find any significant differences between groups regarding the amount of time they spent on task in the different steps during the experiments [$F(5, 54) = 0.58$, $p = 0.54$].

VI. Discussion and Limitations

A. Discussion

1) RQ1: Differences in Learners' Preference of Scaffolds by WTC: The above described results allow us to draw a number of preliminary conclusions. First, we found that learners' preferences for the scaffolds embedded in the system tend to vary by their level of WTC. Although the combination of both scaffolds was the most preferred by learners, we observed that learners with a lower WTC tended to prefer AB over CS, whereas their higher WTC counterparts tended to favor CS over AB (*RQ1*). This is an interesting finding because it gives us novel insights on the existence of a relationship between second language learners' developmental level and perception of the scaffolds. Although, we acknowledge that such learners' perceptions or preferences depend on their metacognitive skills [81] and may therefore not be necessarily correct, the following are our conjectures on the reasons behind the results.

As mentioned earlier (see Section II-D), CS and AB are intended to play different roles in interactions between second language (L2) learners and the dialogue agent. As detailed in [14], CS are dedicated to allowing relatively smooth interactions between learners and dialogue agents by providing hints and suggestions when pitfalls occur in conversations. By contrast, AB are employed to achieve warm interactions where learners feel anxious about communication, via empathetic backchannels. With this in mind, we assume that L2 learners, especially at lower WTC levels, face greater anxiety and less self-confidence, which often inhibit attempts to use the target language. This further implies that such learners may display a greater need for empathetic care to reassure or encourage them to begin, rather than direct support toward overcoming

pitfalls in the conversation itself, which may explain such learners' inclination for AB.

On the other hand, we think that higher WTC learners feel less anxiety and are more open to conversation than their lower WTC counterparts, and thus they display a need for more direct communicative support to overcome the pitfalls they encounter in their quest for dialogue task completion. This may explain their relative preference for CS.

Additionally, the finding that affective backchannels are especially popular among lower WTC learners suggests that our approach to conveying empathetic support via such scaffolds could be an essential feature contributing in creation of a comfortable learning environment for L2 learners, as they take their first steps in the challenging activity of second language use. Such an empathetic feature, which is often missing in traditional spoken systems, may be especially necessary in systems targeting novice or low-motivated L2 learners. Let us mention here although most learners did not provide in-depth detailed reasons behind their preferences, explanations offered by some of them during the system preference survey tend to support the above hypotheses. Especially among group of learners that interacted with the AB-only and CS-only versions, some lower WTC learners pointed out they felt more at ease with the AB-only agent, while some of their higher WTC counterparts praised the availability of hints provided by the CS-version.

Such assumptions are also consistent to some extent with a previous study on the balance of motivational scaffolding in tutorial dialogue tutoring, which revealed, among other results, that "direct standalone encouragement" helped students of low self-efficacy but not those with high self-efficacy [82].

2) RQ2: Relationships Between Learners' WTC Level, WTC Outcomes, and Scaffolds: Second, we found that the effectiveness of the employed scaffolds toward increasing WTC was related to learners' WTC level, and consequently to their preferences regarding CS and AB, since preferences were related to WTC level (RQ1). *The CS+AB and CS versions tended to work better for learners with higher WTC, whereas their lower WTC counterparts tended to benefit more from their interactions with the CS+AB and AB versions. In other words, by combining these results with the results of RQ1 above, we observe that learners' preferred versions and effective versions toward enhancing their WTC tended to be consistent, irrespective of their current level of motivation toward communication (RQ2).*

To explain such results, we assume that the gradual development of L2 WTC leads to some important shifts within learners' internal affective states and communicative postures toward the dialogue activity. More concretely, as learners' WTC rises from lower to higher levels, it is conceivable that they gradually become less inhibited by their initial anxiety and fear of making mistakes; instead, they become more open to taking risks and speaking the target language. This could explain why lower WTC learners tended to perform better with a version embedding at least some affective scaffolds (i.e., AB) designed to provide affective support, whereas higher WTC learners perform better with a version including at least some communicative scaffolds (i.e., CS) that are designed to catalyze the dialogue flow between L2 learners and the system.

Altogether, obtained results through the current study indicate that learners' preferences for scaffolds embedded in the different system versions vary according to the stage of development of their WTC, and more importantly, such learners' preference tendencies tend to be in line with the effectiveness of the system in enhancing their WTC. This supports the idea that while scaffolds embedded in dialogue agents may provide a powerful mechanism for enhancing learners' engagement toward conversation, their design must be strengthened by an account of learners' preferences for scaffolds according to their WTC level.

3) RQ3: Accounting for Learners' Differences to Achieve Adaptive L2 WTC Support: Our findings beg the broader question of which approach to follow for achieving an L2 learner-tailored WTC support by accounting for their preferences (i.e., RQ3). From the results above, we understand that a *carefully balanced use of communicative (i.e., CS) and affective (i.e., AB) scaffolds according to learners' level of communicative motivation may be quite beneficial for second language learners. For example, learners with a lower WTC (i.e., who tended to prefer AB over CS) could be presented with a system combining both scaffolds where affective scaffolds are more frequently triggered than communicative ones, whereas for learners with higher WTC (i.e., who tend to prefer CS over AB), a version of the system where communicative scaffolds are more frequently used than affective scaffolds could be employed (RQ3).* In this way, a learner's preference-aware dialogue management could be implemented so that as a given second language learner's WTC increases, the dialogue agent uses an adapted balance of conversational strategies that corresponds to each learner's preference according to their WTC level.

More generally, the essence of such conception of second language acquisition support, which should go beyond performance to encompass learners' preferences, is well highlighted as follows: "people tend to choose and consequently do what they feel comfortable with and get better at the given skills" [83, p. 101]. Also, as stated in previous works [11], [15], the key to making personalization in language learning a practical possibility could be the use of technology, since practical constraints limit the extent to which learners' individual preferences can be accommodated in traditional classroom settings. We feel that this article is in line with this view because our study suggests the feasibility of providing language learners with a computer-based language support system that could account for their preferences.

4) RQ4: Accounting for Learners' Differences to Achieve Fading of L2 WTC Support: A desirable attribute of any educational system is to provide support that "should fade over time, to allow the learner to resume control over the process, become independent, and acquire metacognitive skills" [7]. In other words, such systems should not only provide learners with support that ideally fits their characteristics, but also ensure that such support is dismantled in a timely fashion to

increase learning gains. Ideally, it would be interesting to be able to define boundaries for the quantity of scaffold to be removed and under what circumstances. However, the question of optimal pacing and timing for scaffold withdrawal is a challenging one. Although the lowering or withdrawing of assistance is certainly necessary to avoid over-support, it might also have disadvantages when not carried out cautiously [38]. In the same vein, the excessive use of AB and CS strategies might be perceived as "heavy-handed," whereas infrequent use of such strategies may not provide L2 learners with the support and encouragement they need to attain a higher WTC.

In traditional intelligent tutoring systems, fading is necessary to help learners go beyond simple task achievement and enable actual transfer of (meta) cognitive skills to learners. However, in the context of the current work, the concept of fading is relatively different, as mentioned earlier in Section II. First, the scaffolds implemented in the current study mainly target motivational aspects rather than cognitive aspects of language learning. Second, the removal of scaffolds is intended to make learners feel as comfortable as possible with the learning environment by removing or lowering scaffolds that do not correspond to their preferences. In other words, here, scaffold removal is not directly intended to enable transfer of cognitive skills, nor to prevent learners from "gaming the system" behavior (i.e., taking advantage of the existing scaffolds to achieve systematic task completion rather than by attempting to learn the material [73]). Instead, it is intended to provide learners with support that best fits their preferences, to create a conversational environment where learners can enjoy feelings of well-being, and to enhance their WTC in the target language.

We assume that as learners' WTC increases, they will likely become less comfortable with some of the scaffolds employed by the conversational agent. Such an assumption is supported by our results, as we found that there were different tendencies among higher- and lower WTC learners in terms of preferred conversational strategies (i.e., lower WTC learners prefer AB whereas their higher WTC counterparts favor CS), and preferred strategies seem to be more promising towards enhancing learners' WTC. *Therefore, a reasonable approach toward implementing fading here may consist in smoothly reducing or even withdrawing the scaffolds that do not correspond to learners' preferences, thereby preventing them from experiencing sensations of "heavy-handedness" that may make them disengage from interacting with the system (RQ4).*

5) Limitations and Future Works: The results presented above suggest the feasibility of accounting for L2 learners' preferences toward achieving a tailored WTC effective support. Nevertheless, our work has some limitations.

First, since WTC, of course, does not increase overnight, we are aware that mid- to long-term empirical investigations including monitoring of the WTC development of same learners from lower to higher WTC levels is necessary to draw firm conclusions on the practical impact of our approach to L2 learners' WTC; the lack of such aspects at this stage of our work, as well as the relatively small number of samples can be viewed as a limitation. We also acknowledge that self-report of affective states, as conducted in our study, may present some shortcomings in terms of accuracy [84].

Furthermore, although even a very small amount of system personalization may have significantly positive effects on learners [47], higher degrees of personalization might be more beneficial, especially in the long term. This could be achieved by accounting for learners' previous interactions in dialogue management as well as devising a method to automatically and continuously estimate learners' WTC level from their interactions data to properly balance the conversational strategies. However, one of the pressing issues that still require clarification here is the pace and timing for learners' WTC level estimation. For instance, whether a frequent updating of learners' WTC level is desirable or even necessary is an aspect that should be discussed carefully since levels of WTC might not fluctuate so often. Along this line, we will also keep looking for eventual differences for example in the amount of time learners spend on task depending on their WTC level and the employed system version. This might help us further understand the nature of the interaction between time spent on task and level of WTC.

Besides, to maximize the learning outcomes of the system, we would certainly need to propose an approach toward combining the current motivational scaffolds with cognitive scaffolds since both dimensions play an important role in learning.

Moreover, the tentative nature of our discussion on how to implement the fading of the WTC scaffolds can also be viewed as a shortcoming. To overcome this, we would need to implement a gradual removal of the scaffolds and collect experimental evidence on the effectiveness of the proposed approach. To achieve this, by referring to the results presented in this article, we could for example, stick to observed learners' preference tendencies and progressively fade out AB and fade in CS as learners' WTC increases from lower to higher levels, and further progressively fade out CS, so as to entirely dismantle both scaffolds when learners display a sufficient level of WTC. The validity of such fine-grained modelization of fading would certainly require mid-long-term investigation.

Finally, an interesting research direction we are interested in delving into would be clarifying the degree of consistency of the present findings with relevant prior works such as with: Affective AutoTutor [72] in which low prior-knowledge students benefited more from emotion-sensitive support than high prior-knowledge students; and UNC-ITSPOKE [85], where students with greater uncertainty benefited more from the uncertainty-sensitive support than those with greater certainty.

VII. Conclusion

The promise of scaffolding in learning is to make possible constructive interactions between experts and learners as they work collaboratively to shift the locus of responsibility for task completion and learning from the expert to the learner. A key to achieving this is a careful design of fading or withdrawal of scaffolds according to learners' developmental needs, in order to make unaided performance possible. However, discussion on how to implement such fading is often missing in the

literature, especially when it comes to supporting motivational variables of learning.

In this article, based on results of interactions between L2 learners and a dialogue agent, we discussed the ways scaffolding and fading could be achieved in dialogue management according to learners' preferences in order to effectively enhance their willingness to communicate in a second language. We found that depending on the level of motivation toward communication (i.e., WTC level) reached by L2 learners, they tend to prefer certain types of scaffolds, which in turn seem to work better to stimulate their readiness for communication in the target language. Such results suggest the feasibility and meaningfulness of progressive scaffolding and fading of employed motivational scaffolds (i.e., CS and AB), by taking as reference learners' preference for scaffolds as well as their WTC level. We also shed light, especially in Sections II and VI, on the novelty of the conception of fading of motivational scaffolds, which seems to be relatively different from the notion of fading as conceived in traditional (meta) cognitively intelligent tutoring systems.

Developing systems that address variations in student expectations and interests to provide an adaptive level of support is a fruitful avenue for development and research in the field of technology-enhanced education. We hope that the present work will contribute to opening new perspectives toward achieving effective technology-mediated scaffolding and fading of motivational variables affecting second language learning.

References

[1] C. Cazden, *Whole Language Plus: Essays on Literacy in the United States and New Zealand*. New York, NY, USA: Teachers College Press, 1992.

[2] A. P. Sweet, *State of the Art: Transforming Ideas for Teaching and Learning to Read*," Washington, DC, USA: U.S. Dept. Educ., Office Educational Res. Improvement, Office Res., 1993.

[3] D. Wood, J. S. Bruner, and G. Ross, "The role of tutoring in problem solving," *J. Child Psychol. Psychiatry*, vol. 13, no. 2, pp. 89–100, Apr. 1976, doi: 10.1111/j.1469-7610.1976.tb00381.x.

[4] T. Gonulal and S. Loewen, "Scaffolding technique," in *The TESOL Encyclopedia English Language Teaching*. Hoboken, NJ, USA: Wiley, 2018, pp. 1–5.

[5] N. Enyedy, *Personalized Instruction: New Interest, Old Rhetoric, Limited Results, and the Need for a New Direction for Computer-Mediated Learning*. Boulder, CO, USA: Nat. Educ. Policy Center, 2014.

[6] J. Pane, B. McCaffrey, and R. Karam, "Effectiveness of cognitive tutor algebra i at scale," *Eval. Policy Anal.*, vol. 36 no. 2, pp. 127–144, Jun. 2014, doi: 10.3102/0162373713507480.

[7] S. Thornbury, "Educational technology: Assessing its fitness for purpose," in *The Cambridge Guide to Blended Learning for Language Teaching*, M. McCarthy, Ed. Cambridge, U.K.: Cambridge Univ. Press, 2016, pp. 25–35.

[8] A. Mitrovic and B. Martin, "Scaffolding and fading problem selection in SQL-Tutor," in *Proc. 11th Int. Conf. Artif. Intell. Educ.*, 2003, pp. 479–481.

[9] B. Rogoff and W. Gardner, "Adult guidance of cognitive development," in *Everyday Cognition: Its Development in Social Context*, B. Rogoff and J. Lave, Eds. Cambridge, MA, USA: Harvard Univ. Press, 1984, pp. 95–116.

[10] P. D. MacIntyre, R. Clément, Z. Dörnyei, and K. A. Noels, "Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation," *Modern Lang. J.*, vol. 82, no. 4, pp. 545–562, 1998, doi: 10.2307/330224.

[11] T. Yashima, L. Zenuk-Nishide, and K. Shimizu, "The influence of attitudes and affect on willingness to communicate and second language communication," *Lang. Learn.*, vol. 54, no. 1, pp. 119–152, Mar. 2004, doi: 10.1111/j.1467-9922.2004.00250.x.

[12] T. M. Derwing, M. J. Munro, and R. I. Thomson, "A longitudinal study of ESL learners' fluency and comprehensibility development," *Appl. Linguistics*, vol. 29, no. 3, pp. 359–380, Sep. 2008, doi: 10.1093/applin/amm041.

[13] E. Ayedoun, Y. Hayashi, and K. Seta, "Web-services based conversational agent to encourage willingness to communicate in the EFL context," *J. Inf. Syst. Educ.*, vol. 15, no. 1, pp. 5–27, Aug. 2016, doi: 10.12937/ejsise.15.15.

[14] E. Ayedoun, Y. Hayashi, and K. Seta, "Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' Willingness to Communicate," *Int. J. Artif. Intell. Educ.*, vol. 29, no. 1, pp. 29–57, Mar. 2019, doi: 10.1007/s40593-018-0171-6.

[15] J. M. Dewaele and L. Dewaele, "Learner-internal and learner-external predictors of willingness to communicate in the FL Classroom," *J. Eur. Second Lang. Assoc.*, vol. 2, no. 1, pp. 24–37, Aug. 2018, doi: 10.22599/jesla.37.

[16] Y. K. Cao, "Exploring dynamism in willingness to communicate," *Aust. Rev. Appl. Linguistics*, vol. 36, no. 2, pp. 160–176, Jan. 2013, doi: 10.1075/aral.36.2.03cao.

[17] A. Henry, Z. Dörnyei, and P. D. Macintyre, "The dynamics of possible selves," in *Motivational Dynamics Language Learning*, Z. Dornyei, P.D. MacIntyre, A. Henry, Eds. Bristol, U.K.: Multilingual Matters, 2015, pp. 83–94.

[18] P. Sharma and M. J. Hannafin, "Scaffolding in technology-enhanced learning environments," *Interactive Learn. Environ.*, vol. 15, no. 1, pp. 27–46, Apr. 2007, doi: 10.1080/10494820600996972.

[19] A. N. Applebee and J. A. Langer, "Instructional scaffolding: Reading and writing as natural language activities," *Lang. Arts*, vol. 60, no. 2, pp. 168–175, Feb. 1983.

[20] A. Walqui, "Scaffolding instruction for English language learners: A conceptual framework," *Int. J. Bilingual Educ. Bilingualism*, vol. 9, no. 2, pp. 159–180, Dec. 2008, doi: 10.1080/13670050608668639.

[21] B. Rosenshine and C. Meister, "The use of scaffolds for teaching higher-level cognitive strategies," *Educ. Leadership*, vol. 49, no. 7, pp. 26–33, Apr. 1992. [Online]. Available: http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_199204_rosenshine.pdf. Accessed on: Mar. 15, 2020.

[22] M. F. Graves and S. Braaten, "Scaffolded reading experiences: bridges to success," *Preventing School Failure*, vol. 40, no. 4, pp. 169–173, 1996.

[23] A. Collins, J. S. Brown, and S. E. Newman, "Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics," Univ. Illinois, Champaign, IL, USA, Tech. Rep. 403, 1987.

[24] M. R. Lepper, M. F. Drake, and T. O'Donnell-Johnson, "Scaffolding techniques of expert human tutors," in *Scaffolding Student Learning: Instructional Approaches and Issues*, K. Hogan and M. Pressley, Eds. Cambridge, MA, USA: Brookline Books, 1997.

[25] A. S. Palincsar, "The role of dialogue in providing scaffolded instruction," *Educational Psychologist*, vol. 21, nos. 1/2, pp. 73–98, 1986, doi: 10.1080/00461520.1986.9653025.

[26] P. Beed, M. Hawkins, and C. Roller, "Moving learners towards independence the power of scaffolded instruction," *Reading Teacher*, vol. 44, no. 9, pp. 648–655, May, 1991.

[27] J. Wollman-Bonilla and B. Werchadlo, "Teacher and peer roles in scaffolding first graders' responses to literature," *Reading Teacher*, vol. 52, no. 6, pp. 598–607, Mar. 1999.

[28] K. L. Cho and D. H. Jonassen, "The Effects of argumentation scaffolds on argumentation and problem solving," *Educational Technol. Res. Develop.*, vol. 50, no. 3, pp. 5–22, 2002.

[29] C. Quintana *et al.*, "A scaffolding design framework for software to support science inquiry," *J. Learn. Sci.*, vol. 13, no. 3, pp. 337–386, 2004, doi: 10.1207/s15327809jls1303_4

[30] J. W. Saye and T. Brush, "Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments," *Educ. Technol. Res. Develop.*, vol. 50, no. 3, pp. 77–96, Sep. 2002, doi: 10.1007/BF02505026.

[31] A. T. Lumpe and K. Butler, "The information seeking strategies of high school science students," *Res. Sci. Educ.*, vol. 32, no. 4, pp. 549–566, Dec. 2002, doi: 10.1023/A:1022415924768.

[32] P. Sharma and M. Hannafin, "Scaffolding critical thinking in an online course: An exploratory study," *J. Educ. Comput. Res.*, vol. 31, no. 2, pp. 181–208, May 2004, doi: 10.1080/10494820600996972.

[33] B. Rittle-Johnson and K. R. Koedinger, "Designing knowledge scaffolds to support mathematical problem solving," *Cognition Instruction*, vol. 23, no. 3, pp. 313–349, 2005, doi: 10.1207/s1532690xci2303_1.

[34] K. VanLehn *et al.*, "The andes physics tutoring system: Lessons learned," *Int. J. Artif. Intell. Educ.*, vol. 15, no. 3, pp. 147–204, Aug. 2005.

[35] M. Ueno and Y. Miyazawa, "IRT-based adaptive hints to scaffold learning in programming," *IEEE Trans. Learn. Technol.*, vol. 11, no. 4, pp. 415–428, Oct.–Dec. 2018.

[36] I. Tabak, "Synergy: A complement to emerging patterns of distributed scaffolding," *J. Learn. Sci.*, vol. 13, no. 3, pp. 305–335, 2004, doi: doi.org/10.1207/s15327809jls1303_3.

[37] J. D. Farquhar and M. A. Orey, "Intelligent tutoring systems: Toward knowledge representation," in *Instructional Development Paradigms*, C. R. Dolls and A. J. Romiszowski, Eds. Englewood Cliffs, NJ, USA: Educational Technol. Publications, Inc, 1997, pp. 805–814.

[38] K. R. Koedinger and V. Aleven, "Exploring the assistance dilemma in experiments with cognitive tutors," *Educ. Psychol. Rev.*, vol. 19, pp. 239–264, Sep. 2007, doi: 10.1007/s10648-007-9049-0.

[39] R. D. Pea, "The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity," *J. Learn. Sci.*, vol. 13, no. 3, pp. 423–451, Jul. 2004, doi: 10.1207/s15327809jls1303_6.

[40] E. B. Fretz, H. K. Wu, B. Zhang, E. A. Davis, J. S. Krajcik, and E. Soloway, "An investigation of software scaffolds supporting modeling practices," *Res. Sci. Educ.*, vol. 32, no. 4, pp. 567–589, Mar. 2002, doi: 10.1023/A:1022400817926.

[41] M. Lombaard, "Task-based assessment for specific purpose Sesotho for personnel in the small business corporation," doctoral dissertation, Univ. Stellenbosch, Stellenbosch, South Africa, 2006.

[42] R. L. Oxford, *Language Learning Styles and Strategies: An Overview*. Oxford, U.K.: GALA, 2003.

[43] S. Zafar and K. Meenakshi, "Individual learner differences and second language acquisition: A Review," *J. Lang. Teaching Res.*, vol. 3, no. 4, pp. 639–646, Jul. 2012, doi: 10.4304/jltr.3.4.639-646.

[44] R. Ellis, "Principles of instructed language learning," *System*, vol. 33, no. 2, pp. 209–224, Jun. 2005, doi: doi.org/10.1016/j.system.2004.12.006.

[45] M. E. Ehrman, "Understanding the learner at the superior-distinguished threshold," in *From Superior to Distinguished: Training Language Learners at Very High Proficiency Levels*, B. L. Leaver and B. Shekhtman, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2002, pp. 245–259.

[46] B. L. Leaver and B. Shekhtman, "Principles and practices in teaching superior-level language skills: Not just more of the same," in *From Superior to Distinguished: Training Language Learners at Very High Proficiency Levels*, B. L. Leaver and B. Shekhtman, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2002, pp. 3–33.

[47] M. Hinton, "Aptitude as a cognitive/affective construct and its role in the language classroom," *Konin Lang. Stud. (Konińskie Studia Jzykowe)*, vol. 2, no. 2, pp. 125–141, 2014. [Online] Available: http://www.ksj.pwsz.konin.edu.pl/wp-content/uploads/2015/03/KSJ-2-2_-125-141.pdf. Accessed on: Mar. 15, 2020.

[48] Z. Dörnyei and P. Skehan, *Individual Differences in Second Language Learning*. Evanston, IL, USA: Routledge, 2003, pp. 589–630.

[49] D. Larsen-Freeman, "Individual cognitive/affective learner contributions and differential success in second language acquisition," in *Learner Contributions to Language Learning*. Harlow, U.K.: Pearson, Jan. 2001, pp. 12–24.

[50] P. Kerr, "Adaptive learning," *ELT J.*, vol. 70, no. 1, pp. 88–93, Jan. 2016, doi: 10.1093/elt/ccv055.

[51] O. Conlan, I. O'Keeffe, A. Brady, and V. Wade, "Principles for designing activity-based personalized eLearning," in *Proc. IEEE Int. Conf. Adv. Learn. Technol.*, 2007, pp. 642–644.

[52] Y. Gong, J. E. Beck, and N. T. Hefferman, "How to construct more accurate student models: comparing and optimizing knowledge tracing and performance factor analysis," *Int. J. Artif. Intell. Educ.*, vol. 21, nos. 1/2, pp. 27–46, 2011, doi: 10.3233/JAI-2011-016.

[53] C. Conati, A. S. Gertner, K. VanLehn, and M. J. Druzdzel, "Online student modeling for coached problem-solving using Bayesian networks," in *Proc. 6th Int. Conf. User Model.*, 1997, pp. 231–242.

[54] E. Millán and J. L. Pérez-De-La-Cruz, "A bayesian diagnostic algorithm for student modeling and its evaluation," *User Model. User-Adapted Interact.*, vol. 12, nos. 2/3, pp. 281–330, Jun. 2002, doi: 10.1023/A:1015027822614.

[55] R. Sottilare, A. Graessaer, X. Hu, and H. Holden, *Design Recommendations for Intelligent Tutoring Systems*. Adelphi, MD, USA: U.S. Army Res. Lab., 2013.

[56] P. Kerr, *Personalization of Language Learning Through Adaptive Technology. Part of the Cambridge Papers in ELT Series*. Cambridge, U.K.: Cambridge Univ. Press, 2016, pp. 1–15.

[57] D. Larsen-Freeman, "A complexity theory approach to second language development/acquisition," in *Alternative Approaches to Second Language Acquisition*. Evanston, IL, USA: Routledge, 2011, pp. 60–84.

[58] R. Godwin-Jones, "Scaling up and zooming in: Big data and personalization in language learning," *Lang. Learn. Technol.*, vol. 21, no. 1, pp. 4–15, Feb. 2017, doi: 10125/44592.

[59] C. M. Chen, C. J. Chung, and M. C. Chen, "Personalized intelligent English vocabulary learning system based on item response theory and learning memory cycle," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2006, pp. 3828–3833.

[60] H. P. C. P. A. Nascimento, L. B. Marques, D. G. de Souza, F. M. Salgado, and R. Q. Bessa, "A AIED Game to help children with learning disabilities in literacy in the Portuguese language," *in Proc. SBGames*, 2012, pp. 134–143.

[61] M. Carey, "CALL visual feedback for pronunciation of vowels: Kay sona-match," *CALICO J.*, vol. 21, no. 3, pp. 571–601, 2004, doi: 10.1558/cj.v21i3.571-601.

[62] T. Nakata, "Computer-assisted second language vocabulary learning in a paired-associate paradigm: A critical investigation of flashcard software," *Comput. Assisted Lang. Learn.*, vol. 24, no. 1, pp. 17–38, Jan. 2011, doi: 10.1080/09588221.2010.520675.

[63] A. Lippert *et al.*, "The effect of CSAL autotutor on deep comprehension of text in low-literacy adult readers," in *Proc. Annu. Meeting CogSci*, 2017, vol. 1, p. 3766.

[64] C. M. Forsyth, C. Luce, D. Zapata-Rivera, G.T. Jackson, K. Evanini, and Y. So, "Evaluating English language learners' conversations: Man vs. Machine," *Comput. Assisted Lang. Learn.*, vol. 32, no. 4, pp. 398–417, 2019, doi: 10.1080/09588221.2018.1517126.

[65] B. Youngs, S. Moss-Horwitz, and E. Snyder, "Educational data mining for elementary French online: A descriptive study," in *Researching Language Learner Interactions online: From social media to MOOCs* (CALICO Monograph Series). San Marcos, TX, USA: Texas State Univ., 2015, pp. 347–368.

[66] S. Bull and J. Kay, "SMILI ☺: A framework for interfaces to learning data in open learner models, learning analytics and related fields," *Int. J. Artif. Intell. Educ.*, vol. 26, no. 1, pp. 293–331, Mar. 2016, doi: 10.1007/s40593-015-0090-8.

[67] R. Clément, S.C. Baker, and P.D. MacIntyre, "Willingness to communicate in a second language: The effects of context, norms, and vitality," *J. Lang. Social Psychol.*, vol. 22, no. 2, pp. 190–209, Jun. 2003, doi: 10.1177/0261927X03022002003.

[68] B. Ghonsooly, G. H. Khajavy, and S. F. Asadpour, "Willingness to communicate in English among Iranian non–English major university students," *J. Lang. Social Psychol.*, vol. 31, no. 2, pp. 197–211, Jun. 2012, doi: 10.1177/0261927X12438538.

[69] S. Ryan, "Self and identity in L2 motivation in Japan: The ideal L2 self and Japanese learners of English," in *Motivation, Language Identity and L2 Self*, Z. Dornyei and E. Ushioda, Eds. Bristol, U.K.: Multilingual Matters, 2009, pp. 120–143.

[70] A. Mesgarshahr and E. Abdollahzadeh, "The impact of teaching communication strategies on EFL learners' willingness to communicate," *Stud. Second Lang. Learn. Teaching*, vol. 4, no. 1, pp. 51–76, 2014, doi: 10.14746/ssllt.2014.4.1.4.

[71] L. Compton, "Using text chat to improve willingness to communicate," in *Computer-Assisted Language Learning: Concepts, Contexts and Practices*. New York, NY, USA: iUniverse, Inc., 2004, pp. 123–144.

[72] S. K. D'Mello and A. C. Graesser, "AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," *ACM Trans. Interactive Intell. Syst.*, vol. 2, no. 4, pp. 1–39, Jan. 2012, doi: 10.1145/2395123.2395128.

[73] R. Baker, J. Walonoski, N. Heffernan, I. Roll, A. Corbett, and K. Koedinger, "Why students engage in "gaming the system" behavior in interactive learning environments," *J. Interactive Learn. Res.*, vol. 19, no. 2, pp. 185–224, Apr. 2008.

[74] J. Lantolf and A. Pavlenko, "(S) econd (L) anguage (A) ctivity theory: Understanding second language learners as people," in *Learner Contributions to Language Learning: New directions in research*, M. P. Breen Ed. Evanston, IL, USA: Routledge, 2001, pp. 141–158.

[75] B. Norton and K. Toohey, "Changing perspectives on good language learners," *TESOL Quart.*, vol. 35, no. 2, pp. 307–322, 2001, doi: 10.2307/3587650.

[76] D. Block, The Social Turn in Second Language Acquisition. Washington, DC, USA: Georgetown Univ. Press, 2003.

[77] R. Matsuoka, "Willingness to communicate in English among Japanese college students," in *Proc. 11th Conf. Pan-Pac. Assoc. Appl. Linguistics*, 2004, pp. 165–176.

[78] P. D. MacIntyre and C. Charos, "Personality, attitudes, and affect as predictors of second language communication," *J. Lang. Social Psychol.*, vol. 15, no. 1, pp. 3–26, Mar. 1996, doi: 10.1177/0261927X960151001.
[79] D. Howitt and D. Cramer, *Introduction to Research Methods in Psychology*. 3rd ed. London, U.K.: Pearson Education, 2011, pp. 179–181.
[80] A. Aron, E. Coups, and E. J. Aron, *Statistics for Psychology*. 6th ed. London, U.K.: Pearson Education Limited, 2013.
[81] J. Dunlosky and J. Metcalfe, *Metacognition*. Thousand Oaks, CA, USA: Sage Publications, 2008.
[82] K. E. Boyer, R. Phillips, M. Wallis, M. Vouk, and J. Lester, "Balancing cognitive and motivational scaffolding in tutorial dialogue," in *Proc. Int. Conf. Intell. Tutoring Syst.*, 2008, pp. 239–249.
[83] M. E. Ehrman, *Understanding Second Language Learning Difficulties*. Thousand Oaks, CA, USA: SAGE Publications, 1996.
[84] S. K. D'Mello, S. D. Craig, and A.C. Graesser, "Multimethod assessment of affective experience and expression during deep learning," *Int. J. Learn. Technol.*, vol. 4, no. 3/4, pp. 165–187, Oct. 2009, doi: 10.1504/IJLT.2009.028805.
[85] K. Forbes-Riley and D. Litman, "Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor," *Speech Commun.*, vol. 53, nos. 9/10, pp. 1115–1136, Nov. 2011, doi: 10.1016/j.specom.2011.02.006.

Emmanuel Ayedoun received the B.S., M.S., and Ph.D. degrees in informatics from Osaka Prefecture University, Osaka, Japan, in 2016, 2018, and 2020, respectively.

From 2018 to 2020, he was a recipient of the JSPS Research Fellowship for Young Scientists (DC1). He is currently an Assistant Professor with the Faculty of Engineering Science, Kansai University, Suita, Japan. His research interests include computer-supported second language acquisition, embodied conversational agents and human-computer interaction.

Dr. Ayedoun was the recipient of a Poster Presentation Award at the 18th International Conference of Artificial Intelligence in Education (AIED2017). He is a member of JSAI and IAIED.

Yuki Hayashi received the Ph.D. degree in information science from Nagoya University, Nagoya, Japan, in 2012.

From 2009 to 2012, he was a recipient of the JSPS Research Fellowship for Young Scientists (DC1). From 2012 to 2014, he was an Assistant Professor at Seikei University. He is currently an Associate Professor with the College of Sustainable System Sciences, Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Osaka, Japan. His research interests include computer-supported collaborative learning and human–computer interaction.

Dr. Hayashi was the recipient of the Best Overall Paper Award of the 25th International Conference on Computers in Education (ICCE2017). He is a member of IPSJ, JSAI, JSiSE, HIS, and ACM.

Kazuhisa Seta received the B.E. and M.E. degrees from Ryukoku University, Kyoto, Japan, in 1993 and 1995, respectively, and the Ph.D. degree from Osaka University, Suita, Japan, and IAIED, in 1998, all in engineering.

He is currently a Professor with the College of Sustainable System Sciences, Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Osaka, Japan. His research interests include software engineering, intelligent tutoring systems, human resource management, and ontological engineering.

Dr. Seta was the recipient of best paper awards from the Japanese Society for Information and Systems in Education in 2012 and 2015, as well as the Best Overall Paper Award of the 25th International Conference on Computers in Education (ICCE2017). He is a member of JSAI, IEICE, IPSJ, JSiSE, JCSS, and IAIED.