Providing graduated corrective feedback in an intelligent computer-assisted language learning environment

Haiyang Ai

University of Cincinnati

Accepted for publication in *ReCALL*, February 28, 2017

Correspondence:

Haiyang Ai
School of Education
University of Cincinnati
615P Teachers College
2610 McMicken Circle, ML 0022
Cincinnati, OH 45220
United States

haiyang.ai@uc.edu

Abstract

Corrective feedback (CF), a response to linguistic errors made by second language (L2) learners, has received extensive scholarly attention in second language acquisition. While much of previous research on CF has focused on whether CF facilitates or impedes L2 development, few studies have examined the efficacy of gradually modifying the explicitness or specificity of CF as a function of a learner's response to the feedback. Yet, the type and extent of CF needed by a learner, as suggested by Vygotsky (1978), sheds light on whether and the ways in which a learner is developing his or her abilities in a particular area. This paper reports on a study that explores the design, effectiveness and learners' perception to a *graduated* (Aljaafreh & Lantolf, 1994) approach to CF, i.e., feedback that progresses from very general and implicit to very specific and explicit, in an intelligent computer-assisted language learning environment. The results show that the graduated approach to CF is effective in helping learners to self-identify and self-correct a number of grammatical issues, although an onsite tutor provides necessary remedies when the ICALL system occasionally fails to do its part. The paper concludes with a discussion of some challenges of providing graduated CF in a computerized environment.

Keywords: corrective feedback, intelligent computer-assisted language learning, sociocultural theory, the *ba*-construction, L2 Chinese

Providing graduated corrective feedback in an intelligent computer-assisted language learning environment

1. INTRODUCTION

Corrective feedback (CF) refers to the practice whereby a teacher or peer provides formal or informal feedback to learners on their performance that contains linguistic error. CF has received increasingly scholarly attention over the past two decades (Bitchener & Ferris, 2012; Bitchener & Storch, 2016; Mackey, 2012). Previous research on CF has primarily focused on (1) the typology of CF types (R. Ellis, 2009), (2) whether CF facilitates or impedes second language (L2) development (Li, 2010; Lyster & Ranta, 1997; Russell & Spada, 2006; Truscott, 1996, 2007), (3) whether explicit CF is more effective than implicit CF (Bitchener, 2008; Bitchener & Knoch, 2010a; Ferris, 2006; Ferris & Roberts, 2001; Sheen, 2007), (4) whether CF is more effective when it is focused (targeting a few structures at a time) or unfocused (comprehensive) (Bitchener & Knoch, 2010b; R. Ellis, Sheen, Murakami, & Takashima, 2008; Ferris, 1997), and (5) learners' attitudes toward and perceptions of CF (Cornillie, Clarebout, & Desmet, 2012). This body of research has accumulatively contributed to our overall understanding of the mechanisms of CF in facilitating language learning.

However, the CF used in previous studies tends to be static and stationary; that is, it generally does not change in terms of explicitness or specificity as a function of a learner's response to the feedback. The type and extent of CF needed by a learner, as suggested by Vygotsky (1978), sheds important light on whether and the ways in which a learner is developing his or her abilities in a particular area. Aljaafreh and Lantolf (1994) showed that CF, provided in a graduated fashion—i.e., from more implicit (asking learners to read an erroneous sentence) to

more explicit (providing learners with metalinguistic explanations)—can promote L2 learning in a dialogically and collaboratively constructed zone of proximal development (ZPD).

Guided by the same theoretical framework, Poehner and colleagues have developed a web-based formative assessment tool called computerized dynamic assessment (C-DA) to evaluate learners' language proficiency in French, Russian, and Chinese (Poehner & Lantolf, 2013; Poehner, Zhang, & Lu, 2015). The C-DA system documents how many test questions learners answer correctly and how many incorrectly on the first try, as well as tracking how much CF they need in order to complete the assessment task. By offering graduated CF, the C-DA system is able to gauge test-takers' listening and reading comprehension abilities in a more fine-grained way than is possible with traditional tests. This research project has highlighted the usefulness of dynamically adjusting the explicitness of CF depending on a learner's response to it.

The present study continues this line of research and explores how graduated CF can be implemented in an intelligent computer-assisted language learning (ICALL) environment. Technology-mediated CF, as Sauro (2009) noted, holds great promise for the learning of especially complex or low-salient forms. Given marked advances in computational linguistics, ICALL has become a promising avenue for exploring the effects of graduated CF in language learning. Yet scholars have not addressed the implementation and efficacy of graduated CF in an ICALL environment, despite extensive and valuable research on ICALL's potential for facilitating language learning. The C-DA project, which uses multiple-choice questions, has primarily focused on language recognition and comprehension so far. In the present study, we implemented graduated CF in a socioculturally informed ICALL environment in order to assess learners' language production through the use of a more open-ended translation task. We explore the effectiveness of graduated CF in helping American learners learn the Chinese *ba*-

construction, an important aspect of the Chinese language that invariably proves challenging to L2 speakers (Jin, 1992; Wen, 2012).

2. THEORETICAL FRAMEWORK

2.1 Corrective feedback in second language acquisition

A central question pertaining to the value of CF in the context of second language acquisition centers on whether CF benefits language learning at all, and if, in fact, it does, then in what ways. Truscott (1996, 2007) has argued that CF does not benefit language learning and advocated for its total abandonment in instructional contexts. Using a combination of qualitative analysis and quantitative meta-analysis, Truscott (2007) has claimed with 95% confidence that CF has a "very small" actual benefit, if any, in regard to having a positive impact on learners' ability to write accurately (p. 255). Such negative effect of CF has also been reported by Kepner (1991) and Polio, Fleck, and Leder (1998). For example, Polio et al. (1998) examined 64 ESL students' writings over seven weeks and found that differences in posttest score for the treatment and control group were not significant. However, it needs to be pointed out that this may be due to the difference in instruments between pretest and posttest (journal entry vs. in-class essay). In response, other researchers have amassed a large volume of empirical evidence showing that CF, in fact, benefits learners both in the short term and in the long term (Li, 2010; Lyster & Ranta, 1997; Russell & Spada, 2006). CF helps learners *notice* (Schmidt, 1990) mismatches between their own language production and target-like forms. Researchers have found that CF can be particularly effective when it targets at specific error types as compared to providing comprehensive CF to all errors (Bitchener & Knoch, 2010a, 2010b; R. Ellis et al., 2008; Ferris,

1997; Han, 2002). Certain approaches to CF have been found to be conducive for pushed output, as evidenced in learners' self- or other-repair (Lyster & Ranta, 1997; Panova & Lyster, 2002) as well as accuracy in repair (Nassaji, 2007). Used as a pedagogical tool, CF has been found to be valuable in increasing learners' accuracy in L2 writing (Ferris, 1999, 2006; Bitchener & Ferris, 2012). There is a general consensus among SLA researchers that CF makes errors more salient and explicit and that it is especially useful for helping adult learners avoid fossilization and continue developing their target language competence.

Different types of CF may generate different types of responses, which may, in turn, produce different levels of processing. While explicit CF (e.g., metalinguistic explanation) can be effective in promoting acquisition of specific grammatical features and may be more valuable for L2 learners than unlabeled one (Bitchener, 2008; Bitchener & Knoch, 2010a; Bitchener, Young, & Cameron, 2005; R. Ellis, et al., 2008; Ferris, 2006; Ferris & Roberts, 2001; Sheen, 2007). This may be due to L2 learners receiving extensive formal grammar instruction and explicit CF may elicit their prior knowledge. However, explicit CF has a disadvantage in that it requires minimal processing on the part of the learner, and thus, considered not as beneficial for long-term learning. By contrast, implicit (or indirect) CF requires more work on the part of the learner than explicit CF does and is, therefore, thought to facilitate long-term language learning (Ferris & Roberts, 2001). Finally, scholars have emphasized the importance of considering individual student responses to CF in addition to cross-group comparisons (Bitchener & Ferris, 2012; Ferris, 2006, 2010; Ferris, Liu, Sinha, & Senna, 2013; Hyland & Hyland, 2006). While individual differences may serve as a "useful direction for future second language writing research'' (Reynolds, 2010, p. 167), few studies have attempted to individualize CF for different student writers.

2.2 Sociocultural theory-informed corrective feedback

CF has been studied through a diverse range of theoretical lenses. The guiding principles underlying the design of an ICALL system for Chinese, the focal language in the present study, were drawn from Vygotskian sociocultural theory (Vygotsky, 1978). A fundamental tenet of sociocultural theory is that the human mind is mediated by culturally constructed artifacts, the most pervasive of which is language, considered the most powerful auxiliary means for intentionally controlling and reorganizing social life and psychological processing (Lantolf & Thorne, 2006). A key theoretical construct in sociocultural theory is mediation, which refers to the use of material and symbolic tools or signs in regulating, including influencing and changing, our relationships with others and with ourselves. Of particular interest to the present study is the extent to which graduated mediation can be implemented in an ICALL environment in order to promote L2 development.

One of the best-known theoretical constructs in sociocultural theory, also relevant to the discussion of corrective feedback, is ZPD, which refers to "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). To put it another way, what a person can do today with corrective feedback from others is what he or she will be able to do independently tomorrow, because feedback from others triggers the internalization process by which one's ability to control the mind is enhanced (Lantolf, 2000). As such, the ZPD construct not only evaluates a learner's past performance, but also predicts the learner's potential. For Vygotsky (1978), learning and development are not the same: the developmental process does not coincide

with—but lags behind—the learning process. The difference between the two is what constitutes a learner's ZPD at any given time. The essential characteristic of learning is that it creates the ZPD. Intentionally designed learning instructions organized to be sensitive to a learner's ZPD are highly effective in stimulating qualitative mental development and can lead, therefore, to qualitative changes in development. The present study explores how ZPD can be created between the ICALL system and the learner by developing computer algorithms that provide graduated CF that, in turn, creates language learning opportunities.

The transformation of learners' abilities in the ZPD through dialogic collaboration between the learner and his or her mediator constitutes much work behind the pedagogical approach known as Dynamic Assessment (Lantolf & Poehner, 2004; Poehner, 2008, 2011). Through providing appropriate mediation (i.e., corrective feedback) to both understand and to intervene in development, Dynamic Assessment dialogically linked assessment and instruction as a single activity. Studies employing the Dynamic Assessment approach have examined microgenetic growth both in a range of learning contexts, including traditional classroom-based environments (Lantolf & Poehner, 2011; Poehner & Ableeva, 2011), and computer-based learning environments (Poehner & Lantolf, 2013; Poehner, Zhang, & Lu, 2015).

Of particular interest to L2 development is ontogenetic (longitudinal) analysis and microgenetic analysis, the former of which focuses on developmental processes over a person's lifetime, whereas the latter of which focuses on developmental processes that occur in a relatively short period of time. Wertsch (1985) has noted that microgenetic analysis can be thought of as "a very short-term longitudinal study" (p. 55). The microgenetic approach has been used by SLA researchers to document L2 growth in various contexts. Through detailed transcription of oral interactions, Ohta (2000) has documented two university-level L2 Japanese

learners' microgenetic developmental process on the Japanese desiderative construction in a translation task. By using a range of verbal cues (vowel elongation, filled pauses, intonation contours), the learners provided and responded to developmentally appropriate corrective feedback that facilitated learning and internalization of this grammatical feature. Particularly relevant to the present study is the ICALL system's ability, through the use of relational database technology, to track a learner's microgenetic changes as he/she works through iterations of graduated CF in an effort to complete an English to Chinese translation task.

2.3 Feedback in intelligent computer-assisted language learning

Intelligent computer-assisted language learning (ICALL) has benefited from an evolving theoretical understanding of SLA processes and from rapid advances in computational linguistics and NLP technologies (e.g., lemmatization, part-of-speech annotation, syntactic parsing). Typically, ICALL systems can automatically enhance textual input, analyze a learner's language production, and provide immediate and individualized CF (Dickinson, Eom, Kang, Lee, & Sachs, 2008; Heift, 2002, 2004, 2010a; Heift & Schulze, 2007; Schulze, 2008). To date, a number of ICALL systems have been created, including E-Tutor for German (Heift, 2010a, 2010b), TAGARELA for Portuguese (Amaral, Meurers, & Ziai, 2011), ROBO-SENSEI for Japanese (Nagata, 2009), and WERTi for English (Meurers et al., 2010). E-Tutor provides individualized interactions between the learner and the computer system by emulating a learner–teacher interaction. Through the use of an error-checking system, E-Tutor provides corrective feedback to the learner "one error at a time (Heift, 2010b, p. 448). In this way, the system tracks a learner's performance history on specific activities. Similarly, TAGARELA uses a learner and

a teacher model to select the best feedback strategy to use with each learner based on the level of the activity, the type of task, the characteristics of the errors, and the learner's profile (Amaral et al., 2011). By comparison, while WERTi does not provide individualized CF to learners, it offers various types of supplementary language-learning activities (e.g., colorize, click, and practice) and supports the practice of a wide range of grammatical forms and functions (e.g., articles, gerunds/infinitives, phrasal verbs).

The theoretical frameworks guiding the development of E-Tutor, TAGARELA and WERTi are not related to Vygotskian sociocultural theory, at least not explicitly. However, they have in various ways inspired the development of the Chinese ICALL system reported in this study (e.g., textual enhancement, tracking functions). Whereas ICALL researchers have appropriated various SLA theories, few studies have drawn on theoretical insights of language development from the perspective of Vygotskian sociocultural theory, although CALL researchers have examined the Vygotskian perspective on technology-rich language learning environments (e.g., Blin, 2004). In contrast to the C-DA project, which uses multiple-choice questions, we explore a more open-ended question format to study how graduated CF can help learners to learn the Chinese *ba*-construction. In this article, we consider two research questions: What effect, if any, does graduated CF in an ICALL environment have on language learning? What are the participants' perceptions of this type of CF? Next, we briefly describe the design of the Chinese ICALL system (core algorithms, tracking capabilities, and system architecture), participants and context, and data collection methods and data analysis.

3. METHOD

3.1 Participants and context

This study is part of a larger research project that examines whether and how a concept-based approach to language instruction can promote L2 development in regard to the acquisition of the Chinese *ba*-construction (Ai, 2015). As the full study from which this paper is drawn is quite extensive in the scope of its findings, we limit the analysis here to some of the data that illustrate how graduated CF is implemented in an ICALL environment and how CF promotes language learning in this environment. This study involved six participants enrolled in a one-on-one eight-week enrichment program with the researcher (tutor) to learn the Chinese *ba*-construction. The participants were college students taking third-semester Chinese courses at a large public university in the US. Students at this level had the prerequisite vocabulary and grammatical knowledge (e.g., resultative verb compound) necessary to learn the Chinese *ba*-construction. This study reports the results of the ICALL session of the enrichment program in which the participants spent 30–45 minutes completing an English–Chinese translation task that required them to negotiate various syntactic components of the *ba*-construction.

3.2 Designing a Chinese ICALL system

In this study, we developed a socioculturally informed ICALL system that assesses learners' language production through the use of a relatively open-ended translation task. Figure 1 depicts the core algorithm of the Chinese ICALL program. The ICALL system was designed to provide a series of graduated CFs to the participants whereby the CF progresses from implicit and general to explicit and specific. For instance, if a learner does not provide the correct answer on the first try, the system will start with a very implicit CF: "Hmm, can you take a look at it

again?" This creates an opportunity for the learner to identify and correct the answer him/herself. If, however, the learner still cannot produce a correct answer, then the system provides CF that is slightly more explicit and specific (e.g., "OK. So can you take a look at the grammatical object of the verb phrase?"). The system compares the learner's answer to a set of pre-constructed acceptable ones. In the event that the answer provided by the learner does not match any of the pre-defined answers, the ICALL system then subjects the answer to a series of NLP processes (e.g., Chinese-word segmentation, syntactic parsing) in order to determine the location and nature of the problematic areas and provide relevant CF based on the result of the analysis.

<INSERT FIGURE 1 AND FIGURE 2 ABOUT HERE>

Aljaafreh and Lantolf (1994) have proposed a 13-level regulatory CF system, ranging from the most implicit to the most explicit. In the C-DA project (Poehner & Lantolf, 2013; Poehner, Zhang, & Lu, 2015), four levels of CF are provided, regardless of the learner's response. In contrast, the algorithm designed in this study do not have a predetermined number of levels of graduated CF (Figure 1). This is because the CF provided by the ICALL system depends on the type of error the learner makes, which varies across learners. However, when a learner fails to produce a correct answer on the first attempt, the ICALL system always starts with the most implicit CF. The subsequent CF seeks to target aspects of the syntactic elements of the *ba*-construction, which include (1) the *ba*-particle, i.e., whether it is correctly present or incorrectly absent, (2) the perfective *–le*, i.e., whether it is correctly present or incorrectly absent,[1] (3) the word order, i.e., whether the *ba*-NP correctly occurs before the *ba*-VP or incorrectly occurs after it, (4) the grammatical object, i.e., whether the *ba*-NP is correctly

translated, and (5) the verb complement, i.e., whether a verb complement exists and the complement collocates well with the main verb.

The web-based ICALL system was implemented in Python and used the Django Web Framework. The system utilized the Java-based open-source Stanford Parser[2] to parse the participants' language production in Chinese, and used Tregex[3] to traverse the parse tree in order to identify important structural arrangements of grammatical elements of the *ba*-construction. For instance, a well-formed *ba*-construction stipulates that the *ba*-NP must occur before the *ba*-VP. This idea can be expressed in a Tregex pattern: NP > (IP $ BA) & $ (VP < VRD|VP|VV).[4] This means that an NP must be dominated by an IP that contains the *ba*-particle and that the NP must be located to the left of a VP in which a VRD, VP, or VV is included.

A crucial aspect of CALL and ICALL design is that of tracking learners' interactions with the system (Heift, 2010a; Heift & Schulze, 2007; Park & Kinginger, 2010). Heift (2010a) pointed out that "for an ICALL system to individualize instruction by providing a unique set of system responses and interactions, it must keep a record of each user and exchange that information among all system components" (p. 446). In this study, we used MySQL, an open-source relational database management system, to keep a record of the following information pertaining to the participants' interactions with the system: Who is answering the question? What is the question being answered? What is the participant's answer? What is the participant's confidence level for his or her answer? How long does it take for the participant to provide an answer? What is the IP address of the computer from which the participant provided the answer? Because each user was assigned an independent account, the ICALL system can document in detail the revisions made by each participant to each answer during the ICALL activity.

<INSERT TABLE 1 ABOUT HERE>

3.3 Data collection and analysis

The data analyzed in this study were collected from video screen recordings, website logs, audio and video recordings, and post-enrichment program interviews. The participants' complete interactions with the ICALL system were recorded using a video screen recording software called Camtasia. The post-enrichment interviews were conducted, recorded, and transcribed. The tutor was present when the participants performed the tasks. To analyze the data, we first viewed the video and audio recordings in order to identify instances in which the ICALL system identified (or failed to identify) the participants' problematic areas. We then examined the participants' interactions with the system as captured by the website's logging function, which allowed us to reconstruct a moment-by-moment edit made by the participants as they completed the English–Chinese translation task. Finally, we transcribed the semi-structured interview and used a top-down approach to analyze the transcripts of the interview focusing on illustrative episodes where the participants expressed their views on the ICALL system's pedagogical value, including its occasional break-down, in regard to helping them navigate the various aspects of the *ba*-construction.

4. RESULTS

4.1 Effectiveness of graduated CF

The analysis of the data shows that overall the graduated CF provided by the ICALL system was effective in identifying the participants' problems in regard to various syntactic elements of the *ba*-construction and in providing pertinent and meaningful CF for them to revise their answers. In total, the ICALL system identified 54 errors across the five translation questions by the six participants. Of these, about 15% turned out to be false positives, instances where the ICALL system mistakenly identified the participants' answer as erroneous but nevertheless were grammatically acceptable. In such situations, the tutor intervened and provided needed CF. In general, however, the graduated CF provided by the ICALL system was effective in helping the participants to identify and correct a number of grammatical issues (e.g., punctuation, grammatical objects, and verb complement) related to the *ba*-construction. Derrick, for instance, was able to identify and self-correct a punctuation error based on the graduated CF from the ICALL system. Table 2 shows Derrick's moment-by-moment interaction with the ICALL system on this point.

<INSERT TABLE 2 ABOUT HERE>

Derrick was quite confident [5] (i.e., level 5) about his initial answer. He said "Okay, I feel good about that" and clicked on the submit button. However, his initial answer was rejected by the ICALL system as incorrect, and Derrick received implicit CF: "Hmm, can you take a look at it again?" Derrick then read his initial answer out loud and immediately removed the adverb 很快地 ("quickly"). At that point, he realized that the English sentence specified "those" questions, so he changed the demonstrative 那个 ("that") to 那些 ("those"), and restored the deleted adverb to its original place. He then selected the highest confidence level and submitted this second

answer. Unfortunately, his second attempt was still unsuccessful. Although Derrick had corrected his error in regard to the use of the demonstrative, his error in regard to missing punctuation still remained. In response, the ICALL system provided a second CF that was more explicit than the first: "A complete sentence should have a …?" Derrick thought for a moment, pointed the cursor to the end of his answer and entered a period (i.e., 。in Chinese), and smiled. Choosing again the highest confidence level and saying "Picky, Picky," Derrick submitted his third answer. At this point, the system accepted his answer as correct and displayed—"Congratulations! That's exactly right!" Derrick smiled and said to himself, "Alright! Third time is the charm!"

The above brief episode shows that the two rounds of CF provided by the ICALL system helped Derrick to iteratively revise his answer until it was accepted as correct. This indicates that Derrick was, in fact, able to self-identify and self-correct the mistake without the need of extensive explicit CF. In other words, he already had the relevant knowledge in his repertoire—he just needed a little external mediation in order to correct the errors himself. The interaction between Derrick and the ICALL system collaboratively and dialogically—through a written rather than a spoken mode—created a ZPD that promoted Derrick's development in regard to understanding minor aspects (i.e., demonstrative and punctuation) relating to the *ba*-construction.

In addition to guidance on punctuation, the graduated CF was also effective in drawing the participants' attention to a more difficult area of the *ba*-construction, i.e., the omission or incorrect form of the verb complement. In translating "My roommate fixed my bicycle yesterday afternoon," Larry produced 昨天下午我的室友把我的自行车修了 ("Yesterday afternoon my roommate fix my bicycle –*le*"). His answer was rejected by the ICALL system as incorrect, because the *ba*-construction calls for an explicit description of the results of the verb action. The

graduated CF moved from the most implicit option ("Hmm, so can you take a look at it again?") to the next more explicit option ("Okay, so what's the result of the verbal action?"). Based on this gradually more specific feedback, Larry was able to correct his own error by revising the predicate from 修了 ("fix perfective –*le*") to 修好了 ("fix good perfective –*le*").

The analysis of the data also shows that, in a number of cases, in their efforts to identify and correct various syntactic aspects of the *ba*-construction, the participants benefited from the graduated CF jointly provided by the ICALL system and the tutor, who was present during the ICALL session. Excerpt 1 documents Stacy's interaction with both the ICALL system and the tutor during the first translation task.

<INSERT TABLE 3 ABOUT HERE>

Excerpt 1

Stacy: ((Upon finishing and reviewing her first answer))
I'm pretty confident about this ((chooses the highest confidence level and submits her answer)).
T(utor): ((Smile)).
Stacy: ((The computer rejects her answer and offers an implicit prompt: "Hmm, can you take a look at it again?")) Oh, my gosh, haha ((Stacy shows a little disappointment)).
Stacy: ((Highlighted 昨天下午 "yesterday afternoon")).
Um (+++) do I need to move this here?
T: (++) You could try.

Stacy: 我的同屋 “My roommate,” maybe 我的同屋昨天下午

“my roommate yesterday afternoon.”

Stacy: ((Highlights 修完了 “fix-complete-perfective”)).

Um, oh, it’s probably with the verb again ((laughs)).

This is the part I usually mess up ((looking toward the tutor)).

Stacy: I can probably take this 的 “possessive de” out.

T: No, here you can’t.

Stacy: Okay.

Stacy: ((Stacy chooses confidence level 4 and submits her second answer.

However, the computer rejects her answer as incorrect and offered

feedback: “Okay, so what’s the result of the verbal action?”)).

Okay. So the result is not right.

T: (++) What could it be?

Stacy: > Oh, could it be 好:: “good.”

((Stacy changes 修完 “fix-complete” to 修好 “fix-good”))?

T: Why, why do you think so?

Stacy: Because it’s like, fixed it, so that, it’s (++) good, like …

T: It’s working?

Stacy: Yeah. (+) Let me try that.

((Stacy chooses confidence level 4, and submits her answer)).

Stacy: ((The computer accepts Stacy’s answer as correct and displays “Congratulations!

That’s exactly right!” Stacy is very happy and smiles broadly.))

T: That’s exactly right! It’s not finished. It’s fixed well.

34 Stacy: Okay. So my resultative verb is a little bit shaky at the moment

35 ((then Stacy moves on to the next question)).

Stacy's translation of the *ba*-construction—昨天下午我的同屋把我的自行车修完了 ("Yesterday afternoon, my roommate fixed my bicycle")—was in many regards grammatically correct. It satisfied the major syntactic requirements of the *ba*-construction: the word order was correct, the predicate included a resultative verb compound (RVC), and the use of the perfective marker *–le* was also correct. However, the real issue related to the RVC 修完 ("fix-complete"), which does not express the idea that the bicycle has been fully fixed and restored to good working condition. Stacy's usage merely indicated that the roommate has finished working on the bike. The correct RVC in this context is 修好 ("fix-good"). The most implicit CF provided by the ICALL system (line 6) was not enough to enable Stacy to identify her error. In contrast, the second CF provided by the ICALL system was more explicit and specifically focused on the issue: "So what is the result of the verbal action?" (line 21). At this point, Stacy was fully convinced that the real issue was related to the RVC structure: "Okay, so the result is not right" (line 22). Sensing that Stacy was on the right track in regard to finding the correct resultative word for the RVC structure, the tutor followed up with a leading question: "What could it be?" (line 23). In response, Stacy produced the correct answer, 好 ("good"), and revised her answer accordingly (lines 24–25). In addition to producing a correct answer, more critically, Stacy was able to explain why her choice of the resultative word was appropriate: "The bicycle is fixed so that it's good now." Based on this understanding, Stacy submitted her third revision, which the ICALL system accepted as correct. A congratulatory message appeared on the computer screen, and Stacy expressed great pleasure in regard to her interaction with the ICALL system and in

regard to the fact that she had persisted and ultimately worked out the correct answer based on a more complete understanding of the resultative component of RVC in the *ba*-construction.

As noted previously, there were instances in which the ICALL system was not able to locate the source of the error such that it failed to provide effective graduated CF. The data analysis shows that the participants seemed to have difficulty determining the correct location of the perfective marker –*le*, particularly in post-verbal positions. In translating the first sentence, Larry produced 写了在黑板上 ("write –*le* on the blackboard"). In response, the graduated CF provided by the ICALL system was not particularly relevant or useful: "You might be right already, but the translation you provided is not exactly what I have on file. Can you please try it one more time?" This CF option is typically shown to the participants when the ICALL system has exhausted all the possible potential answers and still can't find a match. As such, it serves as a catch-all clause for all the cases that the system cannot handle by itself. Seeing that the CF series provided by the ICALL system was ineffective, the tutor intervened by informing Larry that the issue at hand pertained to the placement of the perfective marker –*le*. With this information, Larry ventured a guess that the perfective marker –*le* could be placed at the end of the sentence, just before the 吗 *ma* question word, to express the notion that the whole idea is in the past. He, therefore, promptly revised his answer to 她把她的钥匙又锁在车里了吗？ ("Did she lock her keys in the car again?"), which was correct.

4.2 Participants' reflections

The second research question considers the participants' perceptions of the ICALL system in general and the graduated CF in particular in helping them learn the *ba*-construction. The analysis of the transcripts of the post-enrichment interview data shows that the participants

generally expressed positive views on the effectiveness of the ICALL system in helping them learn the various aspects of the *ba*-construction. For example, one aspect of the ICALL system that Larry liked most was that he felt the feedback provided was very "personal" and functioned like "a teacher on the Internet" with whom he could communicate by typing "back and forth."

Excerpt 2

T: So you, you definitely like the computer-based exercises=

Larry: =The most, yes.

T: >Can, can you elaborate on that,< How-, Why, why do you like it the most?

Larry: Um, I think it (++) um, it was more personal, ah, perhaps the messages that came in the dialog box, when you didn't do it right, it was like, "Hmm, that's not it, look at the verb phrase." I think it was almost like having a teacher on the Internet ((laughing)) that you're typing back and forth. And, sure, the program needs like, more work on it, so it can get more answers that are potential. But it was still-, it seems to me, if I wanted to learn Chinese (+) on my own, and I didn't want to take [a] cla:ss, I would use the program, and it could really, I think teach me=

T: =Okay.

While Larry pointed out that the program had room for improvement, he nonetheless acknowledged that the ICALL system has pedagogical value, especially for learners who want to learn Chinese on their own. Some participants mentioned that the graduated CF provided by the ICALL system, particularly the implicit feedback, was useful in that it provided an opportunity

for them to locate and revise the problems on their own. Chris, as shown in Excerpt 3, felt strongly about the pedagogical value of using implicit CF as it afforded him an opportunity to "pick out what is wrong" by himself, a practice that is likely to help him recognize the same type of error in different contexts in their future learning.

Excerpt 3

> If you say, look at it again, even just looking at it the second time, rings something. You can, half the time, you'll find your mistakes: "Oh, I forgot the –*le* at the end, I'm so dumb." And, then, if you really don't get it, you really can't see it, having the computer point to it, and say "This is where you should look. What's wrong here?" And, you say, "Oh, what is wrong here?" "Oh::, and I got it. " But it is, I think it is good not to point to it right away, because instead of kind of giving you the answer like: "Look at the end of the sentence, do you have –*le*?" ((laughing)), like, let me look at it again, and let me figure out it myself. If I can figure [it] out myself, that's gonna be more beneficial than having it pointed out to me.

On the other hand, the participants did comment on the ICALL system's inability to provide pertinent CF in some occasions. For instance, Derrick's translation 她又把钥匙锁在车里面了吗？ ("Did she lock the keys in the car again?") was rejected by the ICALL system as incorrect, because it was looking for 她的钥匙 ("her keys") for the grammatical object slot (i.e., Derrick only provided 钥匙 ("keys")). The prompt Derrick received from the ICALL system for this particular iteration was not particularly helpful: "You might be right already, but the translation you provided is not exactly what I have on file. Can you please try it one more time?"

Derrick looked at the feedback for a while, and eventually said "I have no idea how to translate this." However, the translation provided by Derrick was, in fact, correct because the pronoun "her" can be inferred and thus omitted from the context—a practice that is not uncommon in Chinese.

5. DISCUSSION

The graduated approach to providing CF in the ICALL system was found to be an effective pedagogical tool in mediating the participants' progress in using the *ba*-construction. When the CF became more specific, the participants were more likely to locate the issue and self-correct the error in question. For instance, when the ICALL system provided a more specific feedback on Stacy's problematic area (i.e., resultative), Stacy was able to act on this information, and worked out an acceptable answer that include the correct components of the grammatical construction. This finding corroborates Heift's (2004) study on learner uptake of CF in German ICALL system E-Tutor. She reported that when the CF was more explicit and prominent, her students were more likely to correct their errors in grammar and vocabulary exercises. Similar results have been reported in non-ICALL studies. For example, Han (2002) noted that when targeted at specific L2 forms, CF can be especially useful in helping learners notice mismatches between their own language production and target-like forms.

The mechanism of CF designed in this study differs from previous work in several important aspects. It differs from the CF provided in the C-DA project (Poehner & Lantolf, 2013; Poehner, Zhang, & Lu, 2015) in that the type of CF provided in this study was contingent on learners' language production, not recognition. Another difference pertains to the selection of

which linguistic feature to focus on. The CF provided in the ICALL system reported in this study takes an iterative process and focuses on one error at a time, a feature that is similar to the German ICALL system E-Tutor (Heift, 2010b). Yet it differs from the E-Tutor in that CF reported in this study was *graduated* in nature (moving from more general to more specific), whereas in E-Tutor, it was prioritized according to frequency and error type (e.g., first focus on word order, then subject, then object, and finally prepositional phrase).

L2 microgenetic development in previous research has primarily been studied in the context of moment-to-moment interactions between language learners and mediators in face-to-face scenarios (Aljaafreh & Lantolf, 1994; Lantolf & Poehner, 2011; Poehner, 2008; van Compernolle, 2011). Using conversational analytic method, van Compernolle (2011) illustrated how one learner progressed in cognitive function in sociopragmatic concepts of French second person pronouns in one-hour one-on-one concept-based instruction tutorial. In the present study, the majority of the CF or mediation was realized through computer-mediated means. However, the microgenetic development documented in the ICALL environment, as the results show, parallel to those in the more traditional face-to-face context. In addition, the provision of graduated CF by the ICALL system to understand and intervene L2 development resonates with practices advocated by Dynamic Assessment (Lantolf & Poehner, 2004; Poehner, 2007, 2008). For instance, Chris expressed that figuring out something (not just language) by himself is more valuable because “now that I know that is, I can see it the next time I do it.” In other words, the amount of support provided to Chris by the ICALL system enabled him to stretch beyond their independent performance. Yet at the same time, it also provided insights for his emergent abilities. The graduated approach to providing CF affords an opportunity for the participants to take on as much responsibility for task completion as possible, and the ICALL system (and at

times, the tutor) remain ready to intervene when the participants "slip over the edge" of their abilities (Newman, Griffin, & Cole, 1989, p. 87).

The analysis also found instances in which the ICALL system failed to locate the source of the error (e.g., the placement of the perfective marker –*le*) and thus failed to provide effective graduated CF. At such junctions, the tutor intervened and remedied the situation by providing necessary CF to the participants. This was not an intended feature in the original research design. The presence of the human tutor in an ICALL environment may have hindered the effectiveness of the graduated CF of the ICALL system being used as a stand-alone computer program. While a completely independent ICALL system may be desirable, in technology-mediated learning environments, it is not entirely unusual to have an instructor available to help the learners navigate the various technology and non-technology related hurdles in L2 learning. The interview data from this study showed that the participants generally acknowledged the pedagogical value of this tool and appreciated the tutor's remedial CF when the ICALL system broke down and failed to do its part.

The interview data revealed a preference towards the implicit CF by some participants. For example, Chris indicated that implicit CF provides him opportunities to figure out the issue by himself, and regarded it as more beneficial than having the correct answer provided to him. Even the most implicit CF of having him looked at his answer one more time "rings something." This finding corroborates results in Panova and Lyster (2002), who found that students prefer implicit type of CF (e.g., recasts). Ferris et al.'s (2013) reported that students appear to learn more with indirect/implicit CF. One plausible explanation is that implicit CF requires more language processing (noticing or rehearsing in short-term memory) on the part of the learner in order to turn explicit linguistic knowledge into implicit knowledge (see DeKeyser, 2003; N.

Ellis, 2005; R. Ellis, Loewen, & Erlam, 2006; R. Ellis & Sheen, 2006; Long, 2007 for discussion on explicit/implicit learning/knowledge).

Another finding emerged from the interview pertains to the quality of the CF provided by the ICALL system. By design, the graduated approach to the provision of CF in the ICALL system varies as a function of each participant's actual language production, a practice that can be considered as "individualized" to each learner. For Larry, this type of CF was valuable because it was "personal" and resembled "a teacher on the Internet," with whom he communicated back and forth through typing. This finding is consistent with Ferris et al.'s (2013) study in which they reported that "the students found the individualized and interactive teaching and learning provided through the feedback and interview cycles in the research project to be extremely valuable" (p. 322). Similarly, other scholars have recognized the need to consider individual student responses when providing CF (Bitchener & Ferris, 2012; Ferris, 2006, 2010; Hyland & Hyland, 2006).

Clearly, the ICALL system in its current condition needs further improvement. Derrick's struggle with the unhelpful CF highlighted the system's failure to account for all potentially correct answers, particularly answers to open-ended questions. Although the core algorithm developed in the ICALL system was able to analyze a range of linguistic features (e.g., the *ba*-particle, word order, verb complement), the system still needed to determine whether the answers provided by the participants were grammatically correct. Because there is usually more than one way to express more or less the same meaning, the combination of different lexical and syntactic arrangements can dramatically increase the total number of potential correct answers. This challenge is well recognized in the ICALL literature (Heift, 2010a; Meurers, 2012; Nagata, 2009). Nagata (2009) showed that in order to provide a direct response to a simple question, one

could obtain 6,048 correct sentences by considering possible well-formed lexical, orthographical, and word-order variants. However, that number jumps to a staggering one million if incorrect options restricted only to incorrect particles and conjugation choices were to be included. Heift (2010a) concurred that "it is simply not feasible to anticipate every mistake a student might make" (p. 445). To address this issue, Meurers (2012) suggested that it is necessary to "abstract away from the specific string entered by the learner to more general classes of properties by automatically analyzing the learner input using NLP algorithms and resources" (p. 4194). This appears to be a major hurdle and more research is needed in this area in order for the field to move forward.

6. CONCLUSION

Levy and Stockwell (2006) noted that CALL designers and researchers frequently turn to SLA theories in order to make more principled and sound decisions. WERTi draws on the notion of focus on form and textual enhancement; E-Tutor supports discovery learning; TAGARELA adopts a general scaffolding methodology to help learners develop self-editing skills. The present study turned to Vygotskian sociocultural theory for theoretical guidance in designing a web-based Chinese ICALL system. A distinctive feature of this system is that it analyzes learners' language production from a relatively open-ended question format (i.e., a translation task) and provides a series of graduated CF that moves from more general and implicit to more specific and explicit. The microgenetic analysis showed that the graduated CF, provided by the ICALL system, and supplemented by the tutor, when needed, was effective in helping the participants to

self-identify and self-correct a number of grammatical issues (e.g., punctuation, verb complement).

A technical challenge encountered in this study pertained to cross-programming language interoperability. The challenge lied in the difficulties associated with programmatic communication between the Java-based Stanford NLP software and the Python-based ICALL system in a web-based environment that calls for a relatively short response time. We resolved this issue by using a bridging software called PY4J, a software library written in Python and Java that allows Python to utilize Java functions running natively in Java Virtual Machine (JVM). This, in effect, eliminated the need for the NLP tools to load the time-consuming large dictionaries into the computer memory each time a learner submits an answer to the system, thereby shortening the response time from 15+ seconds to only a few seconds and creating a viable web-based environment. From a technological point of view, the use of such bridging software packages has broader significance to the field of ICALL, as it opens up the possibility of leveraging a wide variety of existing NLP software packages that may or may not be available in the particular programming language and the related environments in which an ICALL system is developed.

The conventional understanding of the notion of "intelligence" in ICALL research leans more toward technology than toward language learning. The idea of intelligence in ICALL comes from the field of artificial intelligence, particularly in relation to NLP techniques (Schulze, 2008). To echo a call Oxford (1993) made a quarter century ago that that "ICALL must devote as much attention to its language learning/teaching principles as it does to its exciting technology" (p. 174), in this study, we proposed an alternative understanding on the notion of intelligence in ICALL: what makes an ICALL system intelligent is not simply the use of state-

of-the-art technologies—although they are necessary—but how such technologies, NLP or otherwise, are creatively used to develop language-learning opportunities and to provide immediate, meaningful, and graduated CF in order to facilitate language development.

This study is not without limitations. In some cases, the ICALL system failed to identify problematic areas in some of the participants' answers and, therefore, could not provide useful graduated CF due to the difficulty of predicting all the potential correct answers. A partial solution might be to inventorize potential errors by combining instructors' knowledge about common errors and errors from large learner corpora. Another approach might be to design activities that allow learners to drag and drop words from a randomized but definitive set of words in order to circumvent the "unlimited number of answers" challenge. Additionally, the ICALL system will benefit from more fine-grained hinting such as highlighting only the problematic segment of a learner's answer. One of the challenges, also room for future improvement, as commented by many participants in this study in post-enrichment interviews, is to "fine-tune" (Poehner, 2007, p. 325) the mediations generated by the ICALL system as learners improve their language abilities and become more agentive.

Notes

1. Not all *ba*-constructions require the use of the perfective marker *–le*. However, the perfective marker *–le* does occur frequently in the *ba*-construction. For the purpose of this study, all the translation sentences involved the use of the perfective marker *–le*.
2. The Stanford Parser is an open-source statistical parser developed by the Stanford NLP Group. See http://nlp.stanford.edu/software/lex-parser.shtml for more information.
3. Tregex is a utility for matching patterns in parse trees. This Java-based software is available from http://nlp.stanford.edu/software/tregex.shtml.
4. The Stanford Parser uses the Penn Chinese Treebank. In the cited example "NP > (IP $ BA) & $ (VP < VRD|VP|VV)", NP stands for noun phrase, VP for verb phrase, VRD for resultative compound, BA for the *ba* particle, VV for other verbs (e.g., modals, action verbs), and IP for inflectional phrase (clause).
5. A five-point scale was used for the participants to self-identify their confidence level about each answer they provide, with 1 being the least confident, and 5 the most confident.

References

Ai, H. (2015). *Promoting second language development with concept-based instruction and intelligent computer-assisted language learning* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal, 78*, 465–483.

Amaral, L., Meurers, D., & Ziai, R. (2011). Analyzing learner language: Towards a flexible NLP architecture for intelligent language tutor. *Computer Assisted Language Learning, 1*, 1–16.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*, 102–118.

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. New York, NY: Routledge.

Bitchener, J., & Knoch, U. (2010a). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*, 207–217.

Bitchener, J., & Knoch, U. (2010b). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31*, 193–214.

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK: Multilingual Matters.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL students. *Journal of Second Language Writing, 12*(3), 191–205.

Blin, F. (2004). CALL and the development of learner autonomy: Towards an activity-theoretical perspective. *ReCALL*, *16*(02), 377–395.

Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. *ReCALL*, *24*(3), 257–278.

Dekeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford, England: Blackwell.

Dickinson, M., Eom, S., Kang, Y., Leeb, C. M., Sachs, R., & Lee, C. M. (2008). A balancing act: How can intelligent computer-generated feedback be provided in learner-to-learner interactions? *Computer Assisted Language Learning*, *21*(4), 369–382.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit knowledge. *Studies in Second Language Acquisition*, *27*(2), 305–352.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, *63*(2), 97–107.

Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, *28*(4), 575–600.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, *36*, 353–371.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*(2), 339–368.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, *31*(2), 315–339.

Ferris, D. R. (1999). One size does not fit all: Response and revision issues for immigrant student writers. In L. Harklau, K. Losey, & M. Siegal (Eds.), *Generation 1.5 meets college composition* (pp. 143–157). Mahwah, NJ: Erlbaum.

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge, England: Cambridge University Press.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*, 181–201.

Ferris, D. R., & Roberts, B. J. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing, 10*, 161–184.

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, *22*(3), 307–329.

Han, Z. H. (2002). Rethinking of corrective feedback in communicative language teaching. *RELC Journal*, *33*(1), 1–34.

Heift, T. (2002). Learner control and error correction in ICALL: Browsers, peekers, and adamants. *CALICO Journal*, *19*(2), 295–313.

Heift, T. (2004). Corrective feedback and learner uptake in CALL. *ReCALL*, *16*(2), 416–431.

Heift, T. (2010a). Developing an intelligent language tutor. *CALICO Journal*, *27*(3), 443–459.

Heift, T. (2010b). Prompting in CALL: A longitudinal study of learner uptake. *The Modern Language Journal*, *94*(2), 198–216.

Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. New York, NY: Routledge.

Hyland, K., & Hyland, F. (2006). Context and issues in feedback on L2 writing: An introduction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 1–20). Cambridge, England: Cambridge University Press.

Jin, H. G. (1992). Pragmaticization and the L2 acquisition of Chinese ba constructions. *Journal of the Chinese Language Teachers Association*, *28*(3), 33–52.

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second language writing skills. *Modern Language Journal, 75,* 305–313.

Lantolf, J. P. (Ed.) (2000). *Sociocultural theory and second language learning*. Oxford, England: Oxford University Press.

Lantolf, J. P. & Poehner, M. E. (2004). Dynamic assessment: Bringing the past into the future. *Journal of Applied Linguistics 1*, 49–74.

Lantolf, J. P. & Poehner, M. E. (2011). Dynamic Assessment in the classroom: Vygotskian praxis for L2 development. *Language Teaching Research, 15,* 11–33.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford, England: Oxford University Press.

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309–365.

Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.

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

Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford, England: Oxford University Press.

Meurers, D. (2012). Natural language processing and language learning. In Carol A. Chapelle (Ed.), *Encyclopedia of Applied Linguistics* (pp. 4193–4205). Oxford, England: Blackwell.

Meurers, D., Ziai, R., Amaral, L., Boyd, A., Mimitrov, A., Metcalf, V., et al. (2010). *Enhancing authentic web pages for language learners*. Paper presented at the Proceedings of the 5th Workshop on Innovative Use of NLP for Building Educational Applications, NAACL-HLT 2010, Los Angeles, CA.

Nagata, N. (2009). Robo-Sensei's NLP-based error detection and feedback generation. *CALICO Journal*, *26*(3), 562–579.

Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning*, *57*(4), 511–548.

Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school*. Cambridge, England: Cambridge University Press.

Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51–78). Oxford, England: Oxford University Press.

Oxford, R. L. (1993). Intelligent computers for learning languages: The view for language acquisition and instructional methodology. *Computer Assisted Language Learning*, *6*(2), 173–188.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36*(4), 573–595.

Park, K., & Kinginger, C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning & Technology*, *14*(3), 31–50.

Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal, 91*, 323–340.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Berlin, Germany: Springer.

Poehner, M. E. (2011). Validity and interaction in the ZPD: Interpreting learner development through L2 Dynamic Assessment. *International Journal of Applied Linguistics, 21*(2), 244–263.

Poehner, M. E. & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment. *Language Teaching Research 17*(3): 323–342.

Poehner, M. E. & Ableeva, R. (2011). Dynamic Assessment. From display of knowledge to engagement in the activity of development. In D. Tsagari & I. Csepes (Eds.), *Classroom-based language assessment*. Frankfurt, Germany: Peter Lang.

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing, 32*(3), 337–357.

Polio, C., Fleck, N., & Leder, N. (1998). ‘‘If only I had more time’’: ESL learners’ changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing, 7*, 43–68.

Reynolds, D. (2010). Beyond texts: A research agenda for quantitative research on second language writers and readers. In T. Silva & P. K. Matsuda (Eds.), *Practicing theory in second language writing* (pp. 159–175). West Lafayette, IN: Parlor Press.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam: John Benjamins.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*(1), 96–120.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 219–258.

Schulze, M. (2008). AI in CALL-artificially inflated or almost imminent? *CALICO Journal*, *25*(3), 510–527.

Sheen, Y. (2007). The effect of focused written corrective feedback of contrastive analysis on EFL learners' acquisition of verb tenses. *TESOL Quarterly*, *41*(2), 255–283.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*, 327–369.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, *16*(4), 255–272.

van Compernolle, R. A. (2011). Developing second language sociopragmatic knowledge through concept-based instruction: A microgenetic case study. *Journal of Pragmatics*, *43*(13), 3267–3283.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wen, X. (2012). A daunting task? The acquisition of the Chinese *ba*-construction by nonnative speakers of Chinese. *Journal of Chinese Linguistics, 40*(1), 216–240.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

Appendix

```
If the learner's answer is already correct:
    Set CF as "Congratulations! That's exactly right!"
    Log meta-information into relational database
    Display CF to the learner
Else:
    Perform Chinese Word Segmentation on submitted answer
    Perform Syntactic Parsing on segmented answer
    If this is the first answer by the participant:
        Set CF as "Hmm, can you take a look at it again?"
    Else:
        If it does not have the ba-particle:
            Set CF as "Okay, so what particle do you think that we might need here?"
        If it does not have the full-stop punctuation:
            Set CF as "A complete sentence should have a ...?"
        If it does not have the perfective marker -le:
            Set CF as "Hmm, so has the action completed yet?"
        If it does not have the correct verb complement:
            Set CF as "Okay, so what's the result of the verbal action?"
        If the ba-VP is placed before the ba-NP:
            Set CF as "Are you sure that the word order is okay?"
        If the ba-NP is correctly translated:
            Set CF as "Can you take a closer look at the grammatical object?"
        Else:
            Set CF as "You might be right already, but the translation you provided is not
                       exactly what I have on file. Can you please try it one more time?"
    Log meta-information into relational database
    Display CF to the learner via web interface
```

Figure 1. Pseudo code for the core algorithms implemented in the ICALL system.

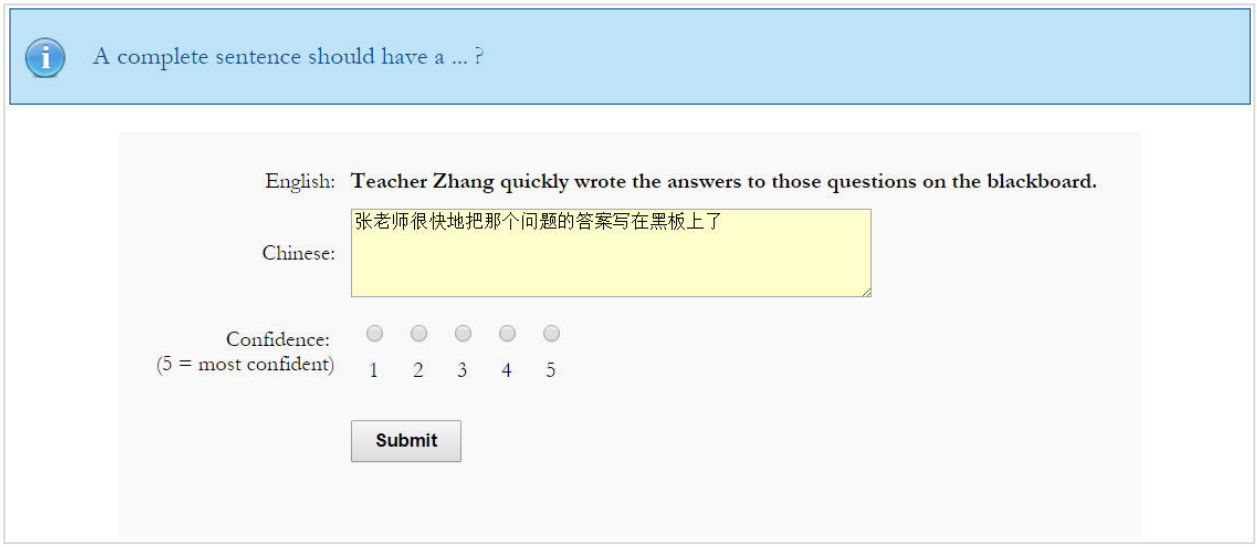

Figure 2. Screenshot of feedback provided by the ICALL system.

Table 1. *Participant Information*

Name	**Gender**	**Major**	**Year/ Program**	**Previous Studies in Chinese**	**Other Language Background**
Derrick	Male	Teaching English to speakers of other Languages (TESL)	Master	2–3 semesters; Spent nine years in Taiwan	Russian (native-like); Spanish (2 years); French (2 years)
Megan	Female	Chinese	Sophomore	2 semesters	Spanish (3 years) French (4 years)
Stacy	Female	Chinese	Freshman	3 years using Rosetta Stone Program in home school setting	N/A
Elaine	Female	Public relations	Sophomore	2 semesters	French (high school)
Chris	Male	History and political science	Sophomore	2 semesters in college; 5 years in high school	N/A
Larry	Male	Classics and ancient Mediterranean studies and ancient languages	Junior	3 semesters (CHNS 001, CHNS 002, CHNS 003)	Spanish (6 years) Modern Greek (one semester of immersion program)

Note. Pseudonyms are used in place of the participants' names.

Table 2. *Moment-by-Moment Changes via Derrick's Interaction with the ICALL System*

#	**Answer**	**Issue**	**ICALL CF**	**Tutor CF**	**Final changes**	**Time elapsed in seconds**
1	张老师很快地把那个问题的答案写在黑板上了	Missing punctuation; incorrect determiner	Hmm, can you take a look at it again?	N/A	(1) Deleting adverb 很快地 "quickly", but later restored it; (2) Changed 那个 "that" to 那些 "those."	71
2	张老师很快地把那些问题的答案写在黑板上了	Missing punctuation	A complete sentence should have a ... ?	N/A	Added period at sentence-final position	35
3	张老师很快地把那些问题的答案写在黑板上了。	N/A	Congratulations! That's exactly right!	That's right!	N/A	9

Table 3. *Moment-by-Moment Changes via Stacy's Interaction with the ICALL System*

#	**Answer**	**Issue**	**ICALL CF**	**Tutor CF**	**Intermediate changes**	**Final changes**	**Time elapsed in seconds**
1	昨天下午我的同屋把我的自行车修完了。	Proper choice of RVC	Hmm, can you take a look at it again?	Answering Stacy's clarification question regarding vocabulary 下午 ("afternoon")	Clarifying 下午	N/A	114
2	我的同屋昨天下午把我的自行车修完了。	Proper choice of RVC	Okay, so what's the result of the verbal action?	(1) You can't take out "de"; (2) what could it be (for the resultative component)?	Tried to take out possessive 的	Moved temporal adverbial 昨天下午 ("Yesterday afternoon") after subject.	59
3	我的同屋昨天下午把我的自行车修好了。	N/A	Congratulations! That's exactly right!	N/A	N/A	N/A	33